Magnetism

Yury Holubeu *

November 23, 2025

The note is not intended for distribution. Magnetism is discussed in detail. Links below show contents of problems and summary or special topics.

Contents

1 Preface and main motivation **9**

I —— Magnetism in a Nutshell —— 11

2 Typical magnetism in a nutshell **11**
2.1 Main parameters and overview of models . . . 11
2.2 Main models and properties of typical magnetic media . . . 15
2.2.1 Dia- and paramagnetism of typical atoms and molecules, Langevin theory . . . 15
2.2.2 Magnetism of two-level systems and quantum gases . . . 21
2.2.3 Dia-, paramagnetism of conducting electrons, Pauli and Landau th. . . . 25
2.2.4 Ferromagnetism, exchange interaction, Weiss theory (!) . . . 27
2.2.5 Antiferromagnetism, ferrimagnetism (!!) . . . 32
2.2.6 Typical Phase Transitions in Magnetism . . . 35
2.2.7 Magnetic Symmetries, Magnetic Classes, and Spatial Groups . . . 35
2.2.8 Atoms, matter, electrons, parameters, units background . . . 36
2.3 Typical Constant Magnetic Fields and Their Interactions with Matter . . . 39
2.3.1 Magnetic field of typical direct currents and current systems . . . 39
2.3.2 Typical magnetostatics in a nutshell . . . 39
2.3.3 Thermodynamic Relations and the Free Energy of a Magnet . . . 39
2.3.4 Self-Induction, Forces, and Gyromagnetic Phenomena . . . 42
2.4 Typical Spin Models and Methods for Them . . . 42
2.4.1 Exchange Interaction, Heisenberg Model, Fundamentals of Spin Waves . . . 42
2.4.2 1D Ising Model (!) . . . 45
2.4.3 2D Ising Model (!) . . . 45
2.4.4 Hubbard, SSH, tight-binding models . . . 46
2.4.5 Spin-Wave Theory and Magnons (!!!!) . . . 47
2.4.6 Jordan-Wigner, Holstein-Primakoff Transformation and Similar (??) . . . 47
2.5 Typical Effects of Magnetism . . . 47
2.5.1 de Haas-von Alphen effect in a Nutshell . . . 47
2.5.2 Superparamagnetism, Magnetocrystalline Anisotropy . . . 47
2.5.3 Energy of the Ferromagnetic State . . . 49
2.5.4 Typical magnetoresistance . . . 49
2.5.5 Magnetic Properties of Atomic Nuclei . . . 50

II —— Special Magnetism in a Nutshell —— 51

3 Famous special properties and models of magnets **51**
3.1 Ferromagnetics, anitferromagnetics special phenomena and properties . . . 51
3.1.1 The Domain Structure of Ferromagnets . . . 51
3.1.2 Interaction of Magnons . . . 51
3.1.3 Phenomena with Ferromagnetism . . . 52
3.1.4 Properties of Ferromagnetic Domains . . . 53
3.1.5 Antiferromagnets Under Special Conditions (?) . . . 55
3.1.6 Dynamics of Magnetic Domains . . . 56

*https://yuriholubeu.github.io/, yuri.holubev@gmail.com

3.1.7 Ferromagnets in Unsteady Magnetic Fields . . . 56
3.1.8 Ferromagnetic Resonance . . . 56
3.1.9 Nuclear Gamma Resonance (The Mossbauer Effect) . . . 57
3.1.10 Stationary Magnetic Structures and the Influence of an External Field on Them 58
3.2 Magnetism in Special Media . . . 58
3.2.1 Special Models of Magnetism of Free Atoms . . . 58
3.2.2 Magnetism of Molecules and Chemical Compounds (????) . . . 58
3.2.3 Magnetic Impurities . . . 59
3.2.4 Special Statistical, Gas Models of Magnetism (???) . . . 59
3.2.5 Electronic and Nuclear Paramagnetic Resonance . . . 60
3.2.6 Giant Magnetoresistance (!?!) . . . 61
3.3 Special General Theories in a Nutshell . . . 62
3.3.1 Landau-Ginzburg Theory in Magnetism . . . 62
3.3.2 Magneto-Optics . . . 62
3.3.3 Magnetic Hydrodynamics in a Nutshell . . . 62
3.3.4 Magnetoresistance . . . 63
3.3.5 Spin liquid, glass, ice . . . 63

4 Famous device applications and novel materials in a nutshell 64
4.1 Typical device applications . . . 64
4.1.1 Materials, alloys . . . 64
4.1.2 Indirect exchanges: superexchange, Goodenough-Kanamori, Zener, RKKY, D-M 64
4.1.3 Magnetic semiconductors, insulators (??) . . . 67
4.1.4 Magnetic data storage . . . 68

5 Other Theoretical models in a nutshell 68
5.1 Overview of special interactions in magnets . . . 68
5.1.1 RKKY Interaction (???) . . . 68
5.1.2 Nagoaka's Model (???) . . . 69
5.1.3 Other 1D Models by Tasaki (???) . . . 69
5.1.4 Beta Ansatz . . . 69
5.1.5 Anisotropic Magnetic Interactions . . . 69
5.1.6 BKT Transition . . . 69
5.1.7 Frustrated Systems . . . 69
5.1.8 Kitaev Model . . . 69
5.1.9 Dzyaloshinskii-Moriya interaction . . . 70
5.1.10 Condo Effect . . . 70

6 Other effects and applications in a nutshell 70
6.1 Other effects . . . 70
6.1.1 Spin-Flop Transition (??) . . . 70
6.1.2 Other Properties of Magnetization Curves . . . 70
6.1.3 Magnetic Hysteresis . . . 71
6.1.4 Magnetic Nanoparticles . . . 71
6.1.5 Antiferromagnetic Resonance and Afm Interaction with Light . . . 71
6.1.6 Soft Magnetic Materials and Alloys . . . 72
6.1.7 High-Frequency Magnetic Materials and Their Applications . . . 73
6.2 Other applications and magnetic devises . . . 73
6.2.1 Magnetoresistive RAM, magnetic memory materials . . . 73
6.2.2 Magneto-optic recording . . . 73
6.3 Special magnetic field in space . . . 73
6.3.1 On special magnetostatics . . . 73
6.3.2 Magnetism of the Planets and Between Them . . . 74
6.4 Experiments on Magnetism in a Nutshell . . . 74

6.4.1 X-ray magnetic scattering . . . 74
6.4.2 Experimental Measurement of Magnetic Susceptibility . . . 74
6.4.3 Experimental Methods for Studying Magnetic Materials . . . 76

7 Other Pictures **76**

III Theory 80

8 Fundamentals of Magnetism (???) **80**
8.0.1 On Concept of Spin . . . 80

9 Magnetism in Condensed Matter by Blundell **80**
9.0.1 Preface . . . 80
9.1 Introduction . . . 83
9.1.1 1.1 Magnetic moments . . . 83
9.1.2 1.2 Classical mechanics and magnetic moments . . . 91
9.1.3 1.3 Quantum mechanics of spin . . . 94
9.1.4 Further reading . . . 102
9.2 Isolated magnetic moments . . . 106
9.2.1 2.1 An atom in a magnetic field . . . 106
9.2.2 2.2 Magnetic susceptibility . . . 107
9.2.3 2.3 Diamagnetism . . . 108
9.2.4 2.4 Paramagnetism . . . 112
9.2.5 2.5 The ground state of an ion and Hund's rules . . . 121
9.2.6 2.6 Adiabatic demagnetization . . . 128
9.2.7 2.7 Nuclear spins . . . 130
9.2.8 2.8 Hyperfine structure . . . 131
9.2.9 Further reading . . . 136
9.3 Environments . . . 138
9.3.1 3.1 Crystal fields . . . 139
9.3.2 3.2 Magnetic resonance techniques . . . 150
9.3.3 Further reading . . . 171
9.4 Interactions . . . 176
9.4.1 4.1 Magnetic dipolar interaction . . . 176
9.4.2 4.2 Exchange interaction . . . 176
9.4.3 Further reading . . . 186
9.5 Order and magnetic structures . . . 190
9.5.1 5.1 Ferromagnetism . . . 190
9.5.2 5.2 Antiferromagnetism . . . 198
9.5.3 5.3 Ferrimagnetism . . . 207
9.5.4 5.4 Helical order . . . 208
9.5.5 5.5 Spin glasses . . . 211
9.5.6 5.6 Nuclear ordering . . . 212
9.5.7 5.7 Measurement of magnetic order . . . 213
9.5.8 Further reading . . . 219
9.6 Order and broken symmetry . . . 222
9.6.1 6.1 Broken symmetry . . . 222
9.6.2 6.2 Models . . . 227
9.6.3 6.3 Consequences of broken symmetry . . . 231
9.6.4 6.4 Phase transitions . . . 233
9.6.5 6.5 Rigidity . . . 235
9.6.6 6.6 Excitations . . . 235
9.6.7 6.7 Domains . . . 243

9.6.8 Further reading . . . 256
9.6.9 References . . . 256
9.7 Magnetism in metals . . . 261
9.7.1 7.1 The free electron model . . . 261
9.7.2 7.2 Pauli paramagnetism . . . 266
9.7.3 7.3 Spontaneously spin-split bands . . . 268
9.7.4 7.4 Spin-density functional theory . . . 271
9.7.5 7.5 Landau levels . . . 273
9.7.6 7.6 Landau diamagnetism . . . 275
9.7.7 7.7 Magnetism of the electron gas . . . 278
9.7.8 7.8 Excitations in the electron gas . . . 283
9.7.9 7.9 Spin-density waves . . . 285
9.7.10 7.10 The Kondo effect . . . 287
9.7.11 7.11 The Hubbard model . . . 288
9.7.12 7.12 Neutron stars . . . 289
9.7.13 Further reading . . . 290
9.8 Competing interactions and low dimensionality . . . 294
9.8.1 8.1 Frustration . . . 294
9.8.2 8.2 Spin glasses . . . 295
9.8.3 8.3 Superparamagnetism . . . 299
9.8.4 8.4 One-dimensional magnets . . . 300
9.8.5 8.5 Two-dimensional magnets . . . 312
9.8.6 8.6 Quantum phase transitions . . . 315
9.8.7 8.7 Thin films and multilayers . . . 317
9.8.8 8.8 Magneto-optics . . . 319
9.8.9 8.9 Magnetoresistance . . . 320
9.8.10 8.10 Organic and molecular magnets . . . 331
9.8.11 8.11 Spin electronics . . . 332
9.8.12 Further reading . . . 333
9.8.13 References . . . 334
9.9 Units in electromagnetism . . . 334
9.9.1 Further reading . . . 337
9.10 Electromagnetism . . . 337
9.10.1 B. 1 Magnetic moments . . . 337
9.10.2 B. 2 Maxwell's equations in free space . . . 340
9.10.3 B. 3 Free and bound currents . . . 340
9.10.4 B. 4 Maxwell's equations in matter . . . 342
9.10.5 B. 5 Boundary conditions . . . 342
9.10.6 Further reading . . . 343
9.11 Quantum and atomic physics . . . 343
9.11.1 C. 1 Quantum mechanics . . . 343
9.11.2 C. 2 Dirac bra and ket notation . . . 344
9.11.3 C. 3 The Bohr model . . . 346
9.11.4 C. 4 Orbital angular momentum . . . 347
9.11.5 C. 5 The hydrogen atom . . . 348
9.11.6 C. 6 The g-factor . . . 349
9.11.7 C. 7 d orbitals . . . 351
9.11.8 C. 8 The spin-orbit interaction . . . 353
9.11.9 C. 9 Landé g-factor . . . 354
9.11.10 C. 10 Perturbation theory . . . 354
9.11.11 Further reading . . . 356
9.12 Energy in magnetism and demagnetizing fields . . . 356
9.12.1 D. 1 Energy . . . 356

9.12.2 D. 2 Demagnetizing factors . . . 357
9.12.3 D. 3 A ferromagnet of arbitrary shape . . . 358
9.12.4 Further reading . . . 361
9.13 Statistical mechanics . . . 361
9.13.1 E. 1 The partition function and thermodynamic functions . . . 361
9.13.2 E. 2 The equipartition theorem . . . 363
9.13.3 Further reading . . . 364
9.14 Answers and hints to selected problems . . . 364
9.15 Symbols, constants and useful equations . . . 379
9.15.1 Fundamental constants . . . 383
9.15.2 Useful equations . . . 383

10 Quantum Theory of Magnetism by White 386
10.1 Preface . . . 386
10.2 The Magnetic Susceptibility . . . 387
10.2.1 1.1 The Magnetic Moment . . . 388
10.2.2 1.2 The Magnetization . . . 392
10.2.3 1.3 The Generalized Susceptibility . . . 395
10.2.4 1.4 Second Quantization . . . 403
10.3 The Magnetic Hamiltonian . . . 411
10.3.1 2.1 The Dirac Equation . . . 412
10.3.2 2.2 Sources of Fields . . . 413
10.3.3 2.3 The Spin Hamiltonian . . . 440
10.4 The Static Susceptibility of Noninteracting Systems . . . 455
10.4.1 3.1 Localized Moments . . . 455
10.4.2 3.2 Metals . . . 462
10.4.3 3.3 Measurement of the Susceptibility . . . 479
10.4.4 3.4 Local Moments in Metals . . . 481
10.5 The Static Susceptibility of Interacting Systems: Local Moments . . . 495
10.5.1 4.1 High Temperatures . . . 496
10.5.2 4.2 Low Temperatures . . . 506
10.5.3 4.3 Temperatures Near T_c . . . 509
10.5.4 4.4 Landau Theory of Second-Order Transitions . . . 512
10.5.5 4.5 Critical Phenomena . . . 513
10.5.6 4.6 Stoner-Wohlfarth Model . . . 518
10.5.7 4.7 Dynamic Coercivity . . . 521
10.5.8 4.8 Magnetic Viscosity . . . 523
10.6 The Static Susceptibility of Interacting Systems: Metals . . . 525
10.6.1 5.1 Fermi Liquid Theory . . . 525
10.6.2 5.2 Heavy Fermion Systems . . . 531
10.6.3 5.3 Itinerant Magnetism . . . 532
10.7 The Dynamic Susceptibility of Weakly Interacting Systems: Local Moments . . . 545
10.7.1 6.1 Equation of Motion . . . 545
10.7.2 6.2 The Bloch Equations . . . 548
10.7.3 6.3 Resonance Line Shape . . . 553
10.7.4 6.4 Spin Echoes . . . 559
10.8 The Dynamic Susceptibility of Weakly Interacting Systems: Metals . . . 565
10.8.1 7.1 Paramagnons . . . 567
10.8.2 7.2 Fermi Liquid Theory . . . 568
10.8.3 7.3 Conduction-Electron Spin Resonance . . . 572
10.8.4 7.4 Spin Waves . . . 575
10.8.5 7.5 Local Moments in Metals . . . 576
10.8.6 7.6 Faraday Effect . . . 579
10.9 The Dynamic Susceptibility of Strongly Interacting Systems . . . 581

10.9.1 8.1 Broken Symmetry . . . 581
10.9.2 8.2 Insulators . . . 582
10.9.3 8.2.8 Optical Processes . . . 604
10.9.4 8.3 High Temperatures . . . 608
10.9.5 8.4 Micromagnetics . . . 610
10.9.6 8.5 Metals . . . 615
10.10 Thin Film Systems . . . 620
10.10.1 9.1 Interfaces . . . 620
10.10.2 9.2 Trilayers . . . 624
10.11 Neutron Scattering . . . 653
10.11.1 10.1 Neutron Scattering Cross Section . . . 653
10.11.2 10.2 Nuclear Scattering . . . 654
10.11.3 10.3 Magnetic Scattering . . . 660
10.11.4 10.3.1 Bragg Scattering . . . 664
10.11.5 10.5 Example: Quantum Phase Transitions . . . 677
10.12 References . . . 680

IV Problems and Questions 691

11 General questions 691
11.1 Общие вопросы . . . 691
11.1.1 Typical exam questions . . . 691
11.1.2 Typical exam problems . . . 692
11.1.3 Exam questions about Exotic magnetic models . . . 692

12 Problems 693
12.1 Problems about magnetic materials . . . 693
12.1.1 Problems about магнетизм типичных веществ . . . 693
12.2 Problems about magnetic field . . . 702
12.2.1 Problems in magnetostatics . . . 702

13 Problems in spin models 702
13.0.1 Problems about AKLT or similar . . . 702

V Other Topics 704

14 Other phenomena and properties 704
14.1 Other, General Questions . . . 704
14.1.1 How Long Can a Magnet Hold a Load? . . . 704
14.1.2 Critical Entropy of Quantum Heisenberg Magnets on Simple-Cubic Lattices by Stefan Wessel . . . 704
14.2 Topological magentic models . . . 708

15 From Nagaoka's ferromagnetism to flat-band ferromagnetism and beyond by Tasaki 708
15.1 Introduction . . . 708
15.1.1 Hubbard model and the origin of ferromagnetism . . . 709
15.1.2 Rigorous results . . . 710
15.1.3 About the present paper . . . 711
15.2 Definition of the Hubbard model . . . 712
15.2.1 Tight-binding description of a single electron . . . 712
15.2.2 Fermion operators and Hilbert space . . . 714
15.2.3 Coordinate free notation . . . 716
15.2.4 Non-interacting system . . . 718

15.2.5 Interaction . . . 719
15.2.6 Hubbard model . . . 720
15.3 Magnetism in the Hubbard model . . . 720
15.3.1 Spin operators and $SU(2)$ invariance . . . 720
15.3.2 Ferromagnetism . . . 721
15.3.3 Instability of ferromagnetism . . . 722
15.3.4 Two more theorems for the absence of ferromagnetism . . . 724
15.4 Nagaoka's ferromagnetism . . . 724
15.4.1 Weak version of Nagaoka's theorem . . . 724
15.4.2 Nagaoka's theorem . . . 727
15.4.3 Connectivity condition . . . 728
15.4.4 Stability of Nagaoka's ferromagnetism . . . 730
15.5 Ferromagnetism in the Hubbard model with long-range hopping . . . 730
15.5.1 Main statement and preliminary results . . . 730
15.5.2 First proof . . . 732
15.5.3 Second proof . . . 732
15.6 Flat-band ferromagnetism . . . 734
15.6.1 Definition . . . 734
15.6.2 Examples . . . 735
15.6.3 Single-electron problem . . . 736
15.6.4 Proof . . . 737
15.6.5 Mielke's flat-band ferromagnetism . . . 738
15.6.6 Beyond flat-band ferromagnetism . . . 738
15.7 Possible experimental realizations of (nearly-)flat-band ferromagnetism . . . 742
15.7.1 Ferromagnetism in $La_4Ba_2Cu_2O_{10}$ . . . 742
15.7.2 Possible ferromagnetism in quantum wires . . . 743
15.8 Towards metallic ferromagnetism . . . 744
15.8.1 Conjectures and some evidence . . . 744
15.8.2 Ferromagnetic t-J model . . . 746
15.8.3 Hubbard model with correlated hopping . . . 747
15.8.4 Limiting U-V model . . . 748
15.9 Adds . . . 750
15.9.1 Gauge transformation . . . 750
15.9.2 Hole-particle transformation . . . 750
15.9.3 Positive semidefinite operators . . . 750
15.9.4 Explicit construction of the Hilbert space and fermion operators . . . 752
15.9.5 Band structures in single-electron energy spectra . . . 753
15.9.6 Proof of Theorem 15.6 . . . 754

16 Solvable Model of a Spin-Glass by David Sherrington and Scott Kirkpatrick 757

17 Magnetic Monopoles in Spin Ice 761

VI Experiments 770

18 High magnetic fields for fundamental physics 770
18.1 Introduction . . . 770
18.2 Experiments . . . 775
18.2.1 Vacuum birefringence . . . 775
18.2.2 Searches for ultralight bosonic dark matter . . . 780
18.2.3 Neutrino-mass measurements . . . 791
18.3 Ideas for future experiments . . . 795

18.3.1 Parallel spectroscopy of complex systems at high magnetic fields: Spectroscopy 2.0 . . . 795
18.3.2 Positron capture by atoms . . . 798
18.3.3 Search for dark matter using atomic-physics methods . . . 799
18.4 Current and forthcoming magnet technologies . . . 800
18.4.1 DC magnets . . . 800
18.4.2 Pulsed magnets . . . 807
18.5 Techniques for measuring high magnetic fields . . . 811
18.5.1 DC fields . . . 811
18.5.2 Pulsed fields . . . 813
18.6 Conclusions and outlook . . . 816
18.7 Glossary . . . 818

VII Appendix **820**

A Introduction **820**
A.1 Other Motivation . . . 820
A.2 Some General Ideas of Important Discoveries in Magnetism . . . 820
A.3 Literature . . . 820
A.3.1 Main Literature . . . 820
A.3.2 Extra Literature . . . 821
A.4 General Introduction and Overview . . . 822
A.4.1 Popular Introduction . . . 822
A.4.2 General Overview . . . 823
A.5 Puzzles for Different Cases . . . 823
A.6 History . . . 823

B Mathematics, Physics, Programming **823**
B.1 On general physics and mathematics concepts . . . 823

C Bibliography **824**

1 Preface and main motivation

Also, there is Hubbard's and other spin models description. I used the following literature.

Let's discuss some minimum knowledge and motivation that would be good to understand for studying magnetism.

Множество приложений

Укажем приложения

(потом раскрою)

Возможность записывать информацию на магнитные носители

(раскрою, что там и компьютеры по этому работают (?? да??) и нобелевские премии за это даны. очень важное для человечества приложение)

Удивительные факты

(потом раскрою)

Головоломки для мотивации

(потом раскрою)

m.te-1. Почему вершинные модели могут магнетизм одних материалов описывать, а других не описывать?

(Вот сказали на лекции, что нельзя описать, поэтому там какие ркку, другие модели придумываются. А казалось бы, просто константу можно было бы новую ввести? Когда-то пойму.)

m.te-1. Why Bismuth is diamagnetic if (???)

(?? a very good question, I'll see again the video below and write it)
https://www.youtube.com/watch?v=g-77EcyI9iY&ab_channel=TheActionLab

Part I
Magnetism in a Nutshell

2 Typical magnetism in a nutshell

2.1 Main parameters and overview of models

Основные модели магнетизма

Вообще, у нас может быть много разных явлений в магнитной среде, но по крайней мере в пренебрежении ими у нас есть заготовки для типичного диа- и пара-магнетизма.

Рассмотрим свободный атом. Если у атома имеется собственный магнитный момент, то paramagnetic susceptibility for atoms with a permanent magnetic moment $\boldsymbol{\mu}$:

$$\chi = N\mu^2/(3k_BT), \qquad \text{for} \quad \mu B \ll k_BT,$$

уменьшается с температурой. Диамагнитные свойства такого атома сичтаются пренебрежимо малы.

Если у атома собственный магнитный момент нулевой, то параматгнитная часть нулевая, соответственно, атом диамагнитный и

$$\chi_{\text{(SI)}} = -\mu_0 \frac{NZe^2}{6m_\text{e}} \left\langle r^2 \right\rangle_\text{av} < 0 \qquad \chi_{\text{(cgs)}} = -\frac{1}{c^2}\frac{NZe^2}{6m_\text{e}} \left\langle r^2 \right\rangle_\text{av}$$

$$N = \frac{N_A\rho}{A}$$

N is the number of atoms per unit volume, N_A is Avogadro's number (the number of atoms per mole), ρ is the density, and A is the molar mass.) (we say that mass is number of atoms times number of neutrons and protons, so we need to divide to get to the number)

There is no explicit dependence in temperature but $\langle r^2\rangle_\text{av}$ is weakly temperature dependent.

Рассмотрим химическое соединение (молекулу) в свободном пространстве (??? пока ничего не могу сказать.)

Рассмотрим металлический материал. Он парамагнитен, и описывается моделью Паули: (?!!!!).

(??? но тут может быть еще ферро и другие магнетизмы, про них тоже указание)

(тут про то, что вообще-то у нас есть еще обменные взаимодействия такие-то)

Можно смотреть просто абстрактно: просто спиновые модели.

(про них тоже указания)

Diamagnetic susceptibility of N atoms of atomic number Z is

$$\chi = -Ze^2N\left\langle r^2\right\rangle/(6mc^2),$$

where $\langle r^2\rangle$ is the mean square atomic radius. Не зависит от температуры. N is the number of atoms per unit volume ($N = N_\text{A}\rho/A$, where A is molar mass (needs to convert A to kg/mol), Z is the number of electrons per atom.

For example, for carbon, $Z = 6$ and $A = 12$ g/mol.

For a system of spins $S = \frac{1}{2}$, the exact magnetization is

$$M = N\mu\tanh\left(\mu B/k_BT\right),$$

where $\mu = \frac{1}{2}g\mu_B$. (Brillouin)

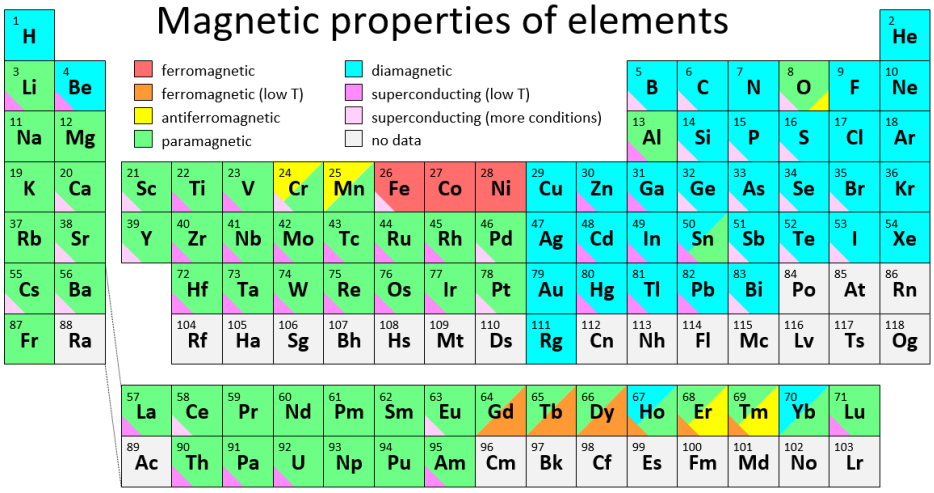
Paramagnetic susceptibility of a Fermi gas of conduction electrons is

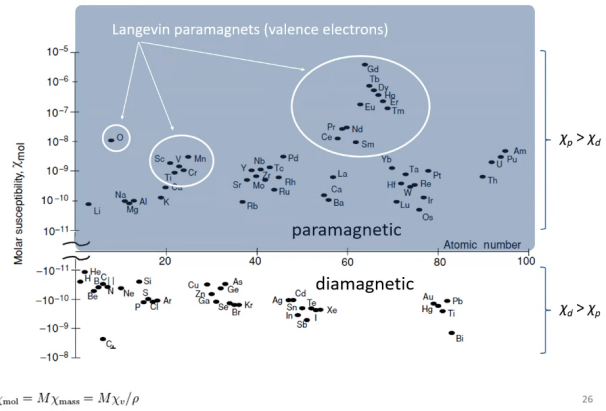
$$\chi = 3N\mu^2/2\epsilon_F,$$

independent of temperature for $k_BT \ll \epsilon_F$ (Pauli)

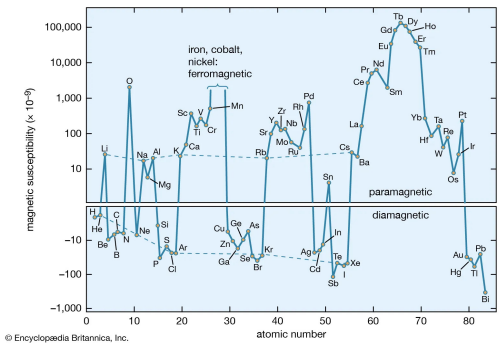
The ground state of electrons in the same shell have the maximum value of S allowed by the Pauli principle and the maximum L consistent with this S. The J value is $L + S$ if the shell is more than half full and $|L - S|$ if the shell is less than half full.

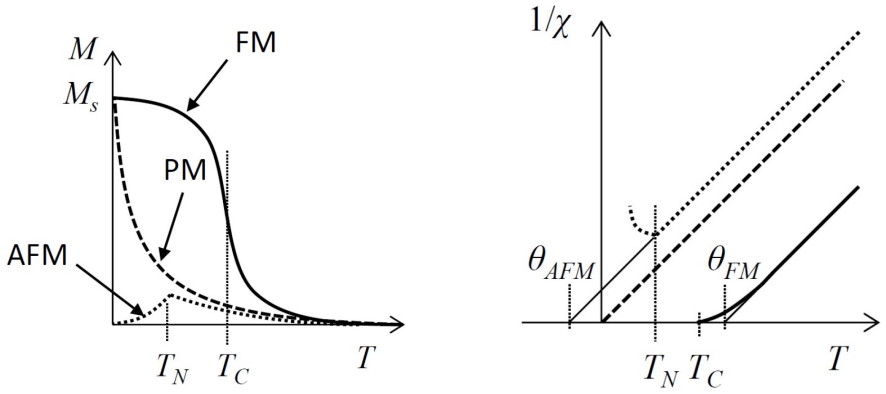
A cooling process operates by demagnetization of a paramagnetic salt at constant entropy. The final temperature reached is of the order of $(B_\Delta/B)\, T_{\text{initial},}$, where B_Δ is the effective local field and B is the initial applied magnetic field.

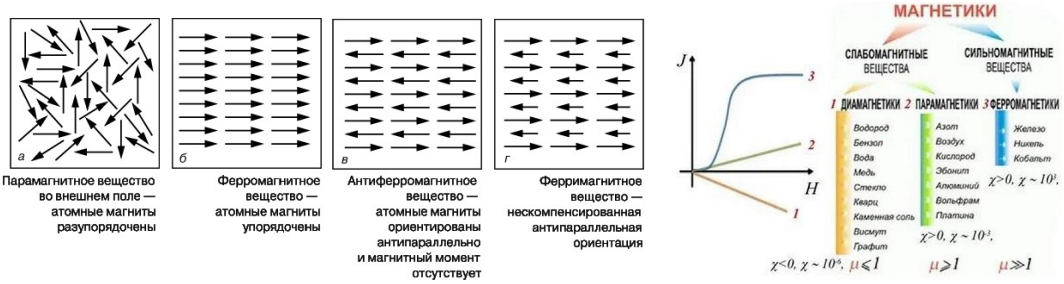




(!! проверю, что действительно тут Ланжевен дает нужные свойства, хорошая тренировка!!!)







Суть законов и параметров электродинамики для магнетизма

(тут определение восприимчивости. Она безразмерная!)

$$\mathbf{B} = \mu_0 \mathbf{H}$$

$$\mu_0 \approx 1,256 \cdot 10^{-6} H \cdot A^{-2}, \text{ (или Гн } \cdot \text{М}^{-1}).$$

(???? а как конвертировать проницаемость из СИ в СГС???)

Для контура площадью S, с током i:

$$\mu := \frac{i}{c} S \qquad\qquad i \equiv \frac{e}{T}$$

$$\frac{\mu}{p} = \frac{e}{2mc}.$$

$$\omega_{\text{L}} := \frac{eH}{2mc},$$

(??? допишу, откуда? из гиромагн соотношения???)

$$\mu_{\text{B}} = \frac{eh}{4\pi mc} = 0,927 \cdot 10^{-27} \frac{\text{erg}}{\text{Э}},$$

(!!! тут общие формулы термодинамического статфиза, которые потом чаще всего используются!!! половина страницы примерно!)

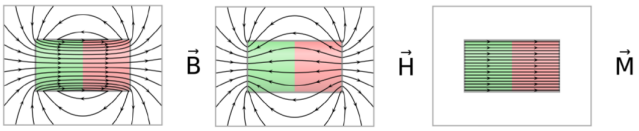
$$\mathbf{H}_{\text{i}} = \mathbf{H}_{\text{a}} - N\mathbf{M},$$

where N is the appropriate demagnetizing factor

'correction term' $\mathbf{H}_{\text{d}} = -N\mathbf{M}$, which you need to add to $\mathbf{H}_{\text{a}}$ to get $\mathbf{H}_{\text{i}}$, is called the demagnetizing field. Similarly

$$\mathbf{B}_{\text{i}} = \mu_0 \left(\mathbf{H}_{\text{i}} + \mathbf{M}\right) = \mathbf{B}_{\text{a}} + \mu_0 (1 - N)\mathbf{M}.$$

$$\vec{H} = \frac{1}{\mu_0}\vec{B} - \vec{M} \text{ (СИ)} \qquad \vec{H} = \vec{B} - 4\pi\vec{M} \text{ (СГС)}.$$

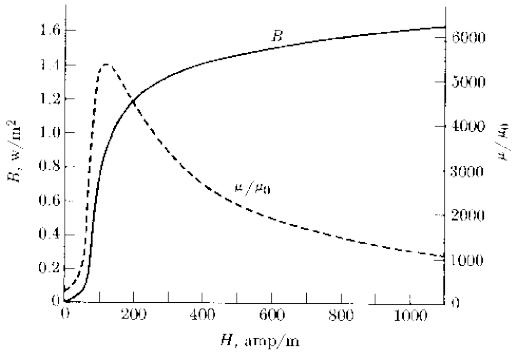


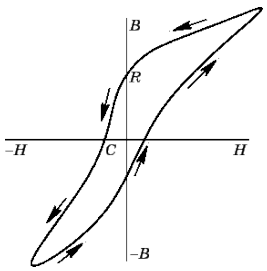
$$\vec{M} = \frac{\chi}{\mu_0(1+\chi)}\vec{B} = \frac{(\mu-1)}{\mu_0\mu}\vec{B} \text{ (СИ)} \qquad \vec{M} = \frac{\chi}{1+4\pi\chi}\vec{B} = \frac{\mu-1}{4\pi\mu}\vec{B} \text{ (СГС)}.$$

В соответствии с определением вектор $\vec{H}$ представляет вклад в магнитную индукцию, обусловленный действием внешних (по отношению к конкретной рассматриваемой точке) причин, создающих поле. Таковыми могут быть токи проводимости $\vec{j}$, переменное во времени электрическое поле (ток смещения $\partial\vec{D}/\partial t$), а также локализованные молекулярные токи $\vec{j}_{\text{mol}}$. Токами $\vec{j}_{\text{mol}}$ создаётся намагниченность, в том числе в областях вне рассматриваемой точки, и эта намагниченность влияет на распределение поля во всём пространстве. Кроме внешних причин, вклад в $\vec{B}$ даёт намагниченность непосредственно в рассматриваемой точке, но этот вклад вычитается. Оперирование вектором $\vec{H}$ не позволяет радикально упростить расчёты. Для нахождения профиля поля (будь то $\vec{B}$ или $\vec{H}$) обычно необходимо решать уравнения Максвелла с учётом соотношений, связывающих $\vec{B}$ и $\vec{H}$. (?? из Вики, потом додумаю!)

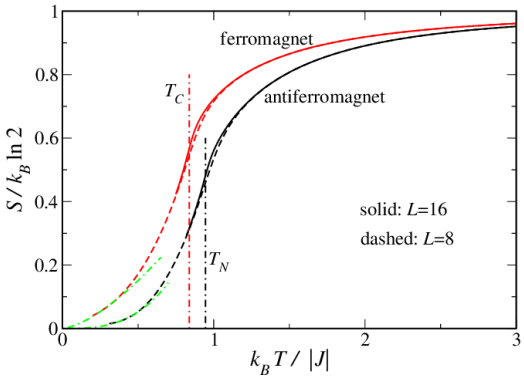
Properties of ferromagnets

Magnetic moment of FM saturates if we increase field, and susceptibility (derivative) has a peak:





(?? а вот про энтропию не знаю, мб после продумывания Изинга пойму ее)



Обзор спиновых моделей

(тут про все модели эти)

Какое поле создают спиновые модели? По

$$\mathbf{B}(\mathbf{r}) = \nabla \times \mathbf{A} = \frac{\mu_0}{4\pi}\left(\frac{3\mathbf{r}(\mu\cdot\mathbf{r})}{r^5} - \frac{\mu}{r^3}\right)$$

по идее нужно просто посчитать от всех спинов. (??? и что будет??? наверняка уже это сделано, когда-то додумаю.)

Обзор общих теоретических свойств магнетизма

(что вообще может быть, а чего нет? пока тут многое не знаю, но наверняка много таких есть общих свойств, все-таки разные теории магнетизма придумывали люди..)

О теореме Бора–ван Леевен

В случае классического (бесспинового электрона) магнетизма быть не может. Действительно:

$$H(\mathbf{r},\mathbf{p}) = \frac{1}{2m}\sum_{i=1}^{N}\left(\mathbf{p}_i - \frac{e}{c}\mathbf{A}\left(\mathbf{r}_i\right)\right)^2 + U(\mathbf{r}),$$

$$Z := \int \frac{d\mathbf{r}d\mathbf{p}}{N!(2\pi\hbar)^{3N}} e^{-\frac{H(\mathbf{r},\mathbf{p})}{T}} = \int \frac{d\mathbf{r}}{N!(2\pi\hbar)^{3N}} e^{-\frac{U(\mathbf{r})}{T}} \left(\int d\mathbf{p}_1 e^{-\frac{1}{2mT}\left(\mathbf{p}_1 - \frac{e}{c}\mathbf{A}(\mathbf{r}_1)\right)^2}\right)^N.$$

Поскольку пределы интегрирования бесконечны, то замена $\mathbf{p} - e\mathbf{A}(\mathbf{r})/c \to \mathbf{p}$ исключает магнитное поле из статистической суммы. Соответственно, магнитный момент системы $\mathbf{M} := -\partial F/\partial \mathbf{B} = 0$. Поэтому чтобы был магнетизм, необходим спин.

(? потом мб чуть додумаю это, как конкретно вставление спина спасает? есть ли еще что-то в этой теореме, кроме этого рассуждения?)

2.2 Main models and properties of typical magnetic media

2.2.1 Dia- and paramagnetism of typical atoms and molecules, Langevin theory

(???? а вообще тут мы считаем ведь много атомов, у нас всегда N есть... так что мб нужно название изменить!)

О диамагнетизме атомов

Диамагнетизм электронной оболочки атома

Рассмотрим свободный покоящийся атом. При отсутствии магнитного поля:

$$\frac{e^2}{r^2} = \frac{mv_0^2}{r} \equiv m\omega_0^2 r.$$

$$\omega_0 = \frac{e}{r^{3/2}m^{1/2}} = \frac{e}{2mc}\frac{2cm^{1/2}}{r^{3/2}} \simeq \frac{e}{2mc}\frac{2\cdot 3\cdot 10^{10}\cdot 3\cdot 10^{-14}}{10^{-12}} \simeq \frac{e}{2mc}\cdot 10^9.$$

В присутствии перпендикулярного магнитного поля $\mathbf{H}$ есть еще сила Лоренца $-\mathbf{f} = (e/c)[\mathbf{vH}]$:

$$\frac{e^2}{r^2} \pm \frac{e}{c}\omega r H = m\omega^2 r \qquad (\pm \text{ - зав. от напр. электрона})$$

$$\omega^2 \pm 2\omega_{\mathrm{L}}\omega - \omega_0^2 = 0, \qquad \omega_{\mathrm{L}} = eH/(2mc); \quad \omega_0 = \frac{e}{r^{3/2}m^{1/2}}.$$

$$\omega = \omega_0 \mp \omega_{\mathrm{L}} \qquad \omega_{\mathrm{L}} \ll \omega_0$$

Изменение энергии:

$$\Delta W = \frac{m}{2}r^2\left[(\omega_0 \mp \omega_{\mathrm{L}})^2 - \omega_0^2\right] = \mp mr^2\omega_0\omega_{\mathrm{L}} + mr^2\omega_{\mathrm{L}}^2 = \mp\mu H + \Delta\mu H.$$

Тут $\mu = \frac{e}{2c}\omega_0 r^2$ и $\Delta\mu := \frac{e}{2c}\omega_{\mathrm{L}} r^2$.

Langevin Diamagnetism Equation

For most materials, the value calculated using the classical Langevin model only shows order-of-magnitude agreement with experiment. Possible sources of the discrepancies between theory and experiment include the following:

- application of Lenz's law (which was determined for electrical circuits) on the atomic scale
- difficulty in calculating or measuring $\langle r^2\rangle_{\mathrm{av}}$; in particular, χ depends on the choice of origin for computing $\langle r^2\rangle_{\mathrm{av}}$
- our assumption that the electrons are orbiting their nucleus; therefore, we expect a poor description of itinerant conduction electrons
- our assumption that the system is spherically symmetric.

We might expect that the use of classical mechanics would introduce errors. However, a full quantum mechanical derivation actually gives the same result.

Quantum Theory of Diamagnetism of Mononuclear Systems

Типичный диамагнетизм

При расчёте диамагнитного момента атома следует учесть, что орбиты его электронов могут быть произвольно ориентированы в пространстве и отличаться по форме от круговой.

$$\Delta\boldsymbol{m} = -\frac{Ze^2\langle r^2\rangle_{\mathrm{av}}\mu_0}{6m_{\mathrm{e}}}\boldsymbol{H}$$

(!! важный абзац про вывод)

For N atoms per unit volume:

$$\chi = \frac{\boldsymbol{M}}{\boldsymbol{H}} = -\frac{N\mu_0 Z e^2}{6m_\mathrm{e}} \left\langle r^2 \right\rangle_\mathrm{av} \equiv -\frac{NZe^2}{6m_\mathrm{e}c^2} \left\langle r^2 \right\rangle_\mathrm{av}$$

Doesn't explicit dependens on temperature, but $\langle r^2 \rangle_\mathrm{av}$ is weakly temperature dependent.
(!! важный абзац про границу применимости!)
(!!! это основная формула!! допишу ее пояснение потом)

Диамагнитная восприимчивость атомов и ионов

(??? в чем идея этой модели, забыл????)

$$\Delta\mu = \frac{e\omega_\mathrm{L}\bar{r}_1^2}{2c} = -\frac{e^2\bar{r}_1^2 H}{4mc^2}.$$

У совокупности из N атомов у каждого атома z электронов и что все ориентации орбит равновероятны, $\bar{r}_1^2 = 2\bar{r}^2/3$, где r - радиус орбиты, поэтому

$$\Delta\bar{M}_\mathrm{A} = -N\frac{e^2 H}{6mc^2}\sum_{i=1}^{l=z}\bar{r}_i^2.$$

$$\chi_\mathrm{a} = \frac{\Delta\bar{M}_\mathrm{A}}{H} = -N\frac{e^2}{6mc^2}\sum_{i=1}^{l=z}\bar{r}_i^2. \qquad \text{(атомная диамагнитная воспр.)}$$

(!! это основная формула)

Вычислим приближённо значение восприимчивости. При этом положим $\sum \bar{r}_i^2 \simeq z\bar{r}_1^2 \simeq z \cdot 10^{-16}$ cm^2. Тогда

$$\varkappa_\mathrm{a} = -6 \cdot 10^{23}\frac{(4,8 \cdot 10^{-10})^2 \cdot 10^{-16} z}{6 \cdot 0,9 \cdot 10^{-27} \cdot 9 \cdot 10^{20}} \simeq -2,8 \cdot 10^{-6} z.$$

Диамагнитная восприимчивость, вычисленная по формуле $\chi_\mathrm{a} = \frac{\Delta\bar{M}_\mathrm{A}}{H} = -N\frac{e^2}{6mc^2}\sum\limits_{i=1}^{l=z}\bar{r}_i^2.$, невелика, и если у атома суммарный орбитальный или спиновый магнитный момент отличен от нуля, то его диамагнитная восприимчивость мала по сравнению с парамагнитной. Действительно, если атом имеет $\mu = \mu_B$, то при комнатной температуре

$$\varkappa_\mathrm{a} = \frac{N\mu_\mathrm{B}^2}{3kT} = 6 \cdot 10^{23}\frac{(0,9 \cdot 10^{-20})^2}{3 \cdot 1,38 \cdot 10^{-16} \cdot 300} \cong 6 \cdot 10^{-4}.$$

Парамагнетизм атомов в рамках теории Ланжевена

(про атомы научусь потом считать так восприимчивость, чтобы описывать эксперимент! отдельное изучение темы требуется)

Классическая теория парамагнетизма Ланжевена

Модель: спин атома как классический вектор статистически распределен в разных направлениях по отношению к магнитному полю, для каждого направления у атома разная энергия взаимодействия с полем, так что повышая поле, будет повышаться энергия и спины будут поворачиваться к полю (?????).

$$\varepsilon = -\mu_0 H \cos\vartheta$$

$$Z_{\text{magn}}(T,H) = \int_0^\pi e^{\mu_0 H\cos\vartheta/(kT)} \cdot 2\pi\sin\vartheta d\vartheta = 4\pi\frac{kT}{\mu_0 H}\frac{e^{\mu_0 H/(kT)} - e^{-\mu_0 H/(kT)}}{2} = 4\pi\frac{kT}{\mu_0 H}\,\text{sh}\,\frac{\mu_0 H}{kT}.$$

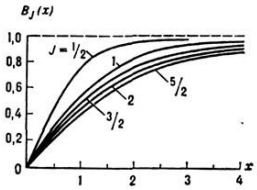
Магнитный момент единицы объёма:

$$m = \frac{N}{V}kT\frac{\partial}{\partial H}\ln Z(T,H) = \frac{N}{V}\mu_0\left[\text{cth}\left(\frac{\mu_0 H}{kT}\right) - \frac{kT}{\mu_0 H}\right].$$

$$L\left(\frac{\mu_0 H}{kT}\right) := \text{cth}\left(\frac{\mu_0 H}{kT}\right) - \frac{kT}{\mu_0 H}$$

In general:

$$B_J(x) = \frac{2J+1}{2J}\coth\left(\frac{2J+1}{2J}x\right) - \frac{1}{2J}\coth\left(\frac{1}{2J}x\right)$$



В малых полях

$$L(x) = \text{cth}(x) - \frac{1}{x} \approx \frac{x}{3} \qquad x := \mu_0 H/(kT) \ll 1.$$

$$m = \frac{N}{V}\mu_0\frac{\mu_0 H}{3kT}. \qquad \varkappa = \frac{I}{H} = \frac{N}{V}\frac{\mu_0^2}{3kT}. \qquad \text{(Закон Кюри)}$$

$$\text{(quantum)} \qquad \varkappa = \frac{N}{V}\frac{J(J+1)g^2\mu_{\text{B}}^2}{3kT}. \qquad \mu_0 \equiv g\sqrt{J(J+1)}\mu_{\text{B}}$$

(формула Хунда) (?????)

Магнитные свойства ионов переходных элементов и влияние поля кристаллической решётки

(4.4. Боровик)

Кроме редкоземельных (от La($Z = 57$) до Lu($Z = 71$)) в периодической системе есть ещё четыре группы элементов, у которых происходит достройка незаполненных внутренних слоёв:

1) группа железа - от Sc($Z = 21$) до Ni($Z = 28$) - заполняется $3d$-оболочка;

2) группа палладия - от Y($Z = 39$) до Pd($Z = 46$) - заполняется $4d$-оболочка;

3) группа платины - от Lu($Z = 71$) до Pt($Z = 78$) - заполняется $5d$-оболочка;

4) группа актинидов - от Ra($Z = 88$) до Md($Z = 101$); в этой группе сначала заполняется $6d$-оболочка, и у Th_{90} два $6d$-электрона; с Pa_{91} начинается заполнение $5f$-оболочки и получается типичная структура редкоземельных элементов (см. табл. П.1).

Однако, гораздо точнее результаты, если считать, что у нас orbital quenching, то есть в магнетизме участвует только спин, а орбитальный момент как-то заморожен:

$$\varkappa_p = \frac{N}{V}\frac{4S(S+1)\mu_{\text{B}}^2}{3kT} \qquad \mu_0 = 2\sqrt{S(S+1)}\mu_{\text{B}}.$$

Четыре основные причины отклонения от закона Кюри:

1. В случае мультиплетных интервалов, сравнимых с kT, постоянная Кюри сама меняется с температурой, а восприимчивость определяется формулой Ван Флека (например, Eu). (??)

2. Второй источник отклонений - Штарк-эффект внутри кристалла. Он был рассмотрен в настоящем параграфе. (????)

3. Третьим источником отклонения является электростатическое обменное взаимодействие электронов, приводящее в некоторых случаях к явлению ферромагнетизма (см. гл. 8). (??)

4. Четвёртый источник - магнитное взаимодействие магнитных моментов атома. Этот эффект существен лишь при низкой температуре. (??)

Некоторые парамагнитные молекулы (!!!!)

(4.5. Боровик)

(додумаю, тут крутые примеры!!! важный раздел для использования теории!!)

Формально на молекулы можно перенести все представления векторной модели сложных атомов, в частности способ обозначения термов. Только для них Боровик обозначает греческие буквы, т.е. вместо S - Σ; вместо P - Π; вместо D - Δ и т. д. Нпр. обозначение $^2\Pi_{3/2}$ будет отвечать состоянию $L = 1, J = 3/2$ и $S = 1/2$.

Молекула кислорода имеет восемь валентных электронов. Их суммарный орбитальный момент равен нулю. Спины шести электронов попарно направлены в противоположные стороны, а у двух спины параллельны. Таким образом, молекула кислорода обладает магнитным моментом, равным двум магнетонам Бора, и восприимчивость кислорода равна

$$\varkappa_{\mathrm{M}} = N\frac{\mu_{\mathrm{B}}^2 g^2 \Sigma(\Sigma-1)}{3kT} = \frac{8N\mu_{\mathrm{B}}^2}{3kT} = \frac{0,993}{T}.$$

Например, окись азота (NO): её основное состояние $-^2\Pi_{1/2}$, однако оно отделено от $^2\Pi_{3/2}$ лишь небольшим энергетическим интервалом ε. Разность энергий ε равна kT при $T = 175$ K. Поэтому при вычислении восприимчивости следует применять формулу Ван Флека (??? которую никто не знает, мб подумаю потом про это).

О деталях, особых случаях и экспериментах

4.6. Свойства парамагнетиков в сильных полях

(когда-то подумаю про это, пока не актуально)

Рассмотрим атомы, находящиеся в S-состоянии, но обладающие спиновым магнитным моментом (ион гадолиния). Терм его основного состояния $^{2S+1}S_S$, а магнитное квантовое число $m = -S, -S+1, \ldots\ldots, S-1, S$. Энергия в магнитном поле составляет $-2m\mu_{\mathrm{B}}H$, откуда

$$Z(T,H) = \sum_{m=-S}^{m=S} e^{2m\mu_{\mathrm{B}}H/(kT)} = \frac{e^{(2S+1)\mu_{\mathrm{B}}H/(kT)} - e^{-(2S+1)\mu_{\mathrm{B}}H/(kT)}}{e^{\mu_{\mathrm{B}}H/(kT)} - e^{-\mu_{\mathrm{B}}H/(kT)}} = \frac{\mathrm{sh}\left\{(2S+1)\mu_{\mathrm{B}}H/(kT)\right\}}{\mathrm{sh}\left\{\mu_{\mathrm{B}}H/(kT)\right\}}$$

откуда

$$\frac{\bar{\mu}}{\mu_{\mathrm{B}}} = \frac{kT}{\mu_{\mathrm{B}}}\frac{\partial}{\partial H}[\ln Z(T,H)] = (2S+1)\,\mathrm{cth}\left\{\frac{(2S+1)\mu_{\mathrm{B}}H}{kT}\right\} - \mathrm{cth}\,\frac{\mu_{\mathrm{B}}H}{kT}.$$

Рассмотрим последовательно область малых и больших полей. В случае малых полей (при $\mu_{\mathrm{B}}H/(kT) \ll 1$), воспользовавшись разложением $\mathrm{cth}(x) = 1/x + x/3\ldots$, получаем уже известный нам результат:

$$\frac{\bar{\mu}}{\mu_{\mathrm{B}}} = \frac{(2S+1)^2\mu_{\mathrm{B}}H}{kT} - \frac{\mu_{\mathrm{B}}H}{kT} = \frac{4S(S+1)\mu_{\mathrm{B}}^2}{3kT}.$$

В больших полях (при $\mu_B H/(kT) \gg 1$) с помощью формулы $\operatorname{cth}(x) \underset{x\to\infty}{\to} 1$ находим

$$\frac{\bar{\mu}}{\mu_B} = 2S.$$

$$\frac{\bar{\mu}}{\mu_0} = \frac{2S+1}{2S} \operatorname{cth}\left\{\frac{2S+1}{2S}\frac{\mu_0 H}{kT}\right\} - \frac{1}{2S}\operatorname{cth}\left(\frac{1}{2S}\frac{\mu_0 H}{kT}\right).$$

При $S \to \infty$, что отвечает переходу к классическому случаю, получаем

$$\frac{\bar{\mu}}{\mu_0} = \operatorname{cth}\frac{\mu_0 H}{kT} - \frac{kT}{\mu_0 H} = L\left(\frac{\mu_0 H}{kT}\right),$$

то есть классическую формулу Ланжевена.

Рассмотрим экспериментальные условия, в которых возможно достижение больших полей. Поскольку $\mu_B = eh/(4\pi mc) = 0,93 \cdot 10^{-20}$ эрг/Гс, при комнатной температуре даже выполнение условия

$$\frac{\mu_B H}{kT} = \frac{0,93 \cdot 10^{-20} H}{1,38 \cdot 10^{-16} \cdot 300} = 1$$

(!!! додумаю, укажу экспериментальные значения!!!)

4.7. Адиабатическое размагничивание и свойства парамагнетиков при температурах меньше 1 K

(это не так актуально, будет нужно - посмотрю. Пока не думаю, что это нужно знать для минимума. Потом мб 1/2 страницы сути добавлю.)

Paramagnetism

Quantum Theory of Paramagnetism

Rare Earth Ions

Iron Group

Ions Crystal

Field Splitting

Quenching of the Orbital Angular Momentum

Spectroscopic Splitting Factor

effect of crystal field on paramagnetism (!?!?!?!)

- magnetic atom in a lattice → Coulomb potential of the neighboring atoms
- extra term in the Hamiltonian: crystal field effect → the Brillouin description is no longer valid (only Zeeman term)
- in some cases small perturbation e.g. anisotropic paramagnetism (together with spin-orbit coupling)
- in some cases stronger:

changing the d-level occupation (filling order for m_s and m_l) (breaking the rules of lecture 1) and therefore the total magnetic moment

in the extreme: quenching of orbital momentum ($L = 0$)

- strong for most $3d$ transition metals; not for $4f$ rare earths ($4f$ orbitals are more shielded from surrounding atoms) 31

2.2.2 Magnetism of two-level systems and quantum gases

Магнетизм ДУС (????)

$$\tau \equiv k_B T$$
$$\frac{N_1}{N} = \frac{\exp(\mu B/\tau)}{\exp(\mu B/\tau) + \exp(-\mu B/\tau)};$$
$$\frac{N_2}{N} = \frac{\exp(-\mu B/\tau)}{\exp(\mu B/\tau) + \exp(-\mu B/\tau)};$$

here N_1, N_2 are the populations of the lower and upper levels, and $N = N_1 + N_2$ is the total number of atoms. The fractional populations are plotted in Fig. 3.

The projection of the magnetic moment of the upper state along the field direction is $-\mu$ and of the lower state is μ. The resultant magnetization for N atoms per unit volume is, with $x = \mu B/k_B T$,

$$M = (N_1 - N_2)\,\mu = N\mu \cdot \frac{e^x - e^{-x}}{e^x + e^{-x}} = N\mu \tanh x.$$

For $x \ll 1, \tanh x \cong x$, and we have

$$M \cong N\mu\,(\mu B/k_B T)\,.$$

Высокотемпературное приближение и закон Кюри (????)

О парамагнетизме электронного наза и законе Кюри в высокотемпературном приближении

Парамагнетизм электронного газа происходит за счет собственного момента электрона $\mu_B = e\hbar/2mc$. В высокотемпературном пределе $T \gg T_{\text{выр}}$ электронный газ больцмановский, его статистическая сумма факторизуется, поэтому вычисляем её для одного электрона в поле $\mathcal{B}$. Его энергия в поле $E = \pm\mu_B \mathcal{B}$,

$$z = e^{-\frac{\mu_B \mathcal{B}}{T}} + e^{\frac{\mu_B \mathcal{B}}{T}} = 2\,\mathrm{ch}\,x, \qquad x \equiv \mu_B \mathcal{B}/T$$
$$f = -T \ln z = -T \ln(2\,\mathrm{ch}\,x),$$
$$m = -\left(\frac{\partial f}{\partial \mathcal{B}}\right) = \mu_B \frac{\partial}{\partial x} \ln(2\,\mathrm{ch}\,x) = \mu_B\ \mathrm{th}\ x.$$

В нулевом поле ($x = 0$) температура «размешивает» моменты электронов по направлениям, и магнитный момент газа отсутствует. При небольших полях $\mu_B \mathcal{B} \ll T$ наведённый магнитный момент линеен по полю, и удобно ввести восприимчивость χ в расчете на одну частицу

$$\chi = \lim_{\mathcal{B}\to 0} m/\mathcal{B}. \qquad m = \chi \cdot \mathcal{B},$$
$$\chi = \mu_B^2/T.$$

При дальнейшем увеличении поля происходит «насыщение» магнитного момента, когда все магнитные моменты (и спины) электронов выстраиваются по (против) полю. Для этого необходимо, чтобы зеемановская энергия электронов была много больше температуры $\mu_B B \gg T$.

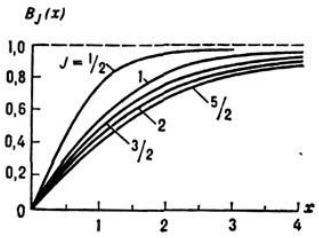
Для произвольного спина S: статистическая сумма содержит $2s+1$ слагаемых, вместо th x используется функция Бриллюэна:

$$m = \mu_B g_s s \mathbb{B}_s(x). \qquad x := \mu_B g_s s \mathcal{B}/T, \qquad g_s \text{ - фактор Ланде,}$$

$$\mathbb{B}_s(x) := \left(1+\frac{1}{2s}\right)\operatorname{cth}\left(1+\frac{1}{2s}\right)x - \frac{1}{2s}\operatorname{cth}\frac{x}{2s} \qquad \text{(функция Бриллюэна)}$$

$$\mathbb{B}_{B\ll T}(x) = \frac{s+1}{3s}x + \ldots \qquad x \ll 1,$$

$$\mathbb{B}_{B\gg T}(x) = \operatorname{cth} x - \frac{1}{x} \equiv \mathbb{L}(x) \qquad s \gg 1, \qquad \text{(функция Ланжевена)}$$



Тогда

$$\chi_{B\ll T} = \frac{\mu_B^2 g_s^2 s(s+1)}{3T}$$

$$\chi_{B\gg T} = \frac{\mu_{\text{кл}}^2}{3T}$$

где $\mu_{\text{кл}}^2 = \mu_B^2 g_s^2 s^2$ - магнитный момент атома в классическом пределе. Например, для невырожденных электронов

$$\chi_{\text{para}} = \left(\frac{\partial m}{\partial B}\right)_{B=0} = \frac{\mu_B^2}{T} \sim T^{-1} \qquad \text{(закон Кюри)}$$

В классическом пределе $T \gg T_{\text{выр}}$ магнетизм исчезает в полном соответствии с теоремой Бора-ван Леевен.

Но это не совсем верное описание. В реальности же числа заполнения вырожденного газа не малы, и статистическая сумма газа не факторизуется. С экспериментом это не согласуется: для низких температур получается слишком завышенное значение. (??? я пока не понял, как подправить эту модель? Можно через Омега потенциал считать и просто получить верные формулы и все, что ниже сделано.)

О диамагнетизме электронного газа в высокотемпературном приближении

Диамагнетизм возникает из-за дискретности уровней Ландау энергии электронов. Для простоты рассмотрим 2D газ:

$$\varepsilon_n = \hbar\omega(n+1/2)$$

$$g_L = \mathcal{B}S/\Phi_0 \qquad \text{- кратность вырождения}$$

Здесь $\hbar\omega = 2\mu_B\mathcal{B}$ - циклотронная частота, а $\Phi_0 = 2\pi\hbar c/e$ - квант потока, а g_L, кратность вырождения уровня Ландау, - это число квантов потока Φ/Φ_0, проходящего через поперечное к полю сечение образца.

$$z = g_L \sum_{n=0}^{\infty} e^{-\frac{\hbar\omega}{T}\left(n+\frac{1}{2}\right)} = \frac{g_L}{2\operatorname{sh} x}, \qquad x := \hbar\omega/2T = \mu_B B/T$$

$$f := -T \ln z = -T \ln \frac{x}{\operatorname{sh} x} + \dots,$$
$$m \equiv -\left(\frac{\partial f}{\partial \mathcal{B}}\right) = \mu_B \frac{\partial}{\partial x} \ln \frac{x}{\operatorname{sh} x} = -\mu_B \left(\operatorname{cth} x - \frac{1}{x}\right) \equiv -\mu_B \mathbb{L}(x).$$
$$\chi_{\text{dia}} = \left(\frac{\partial m}{\partial \mathcal{B}}\right)_{B=0} = -\frac{\mu_B^2}{3T}$$
$$\chi_{\text{dia}} = -\chi_{\text{пара}}/3.$$

Таким образом, в целом электронный газ парамагнитен: $\chi_{\text{dia}} + \chi_{\text{para}} > 0$.

(??? абзац про то, что делать в 3D????? не знаю, вопрос опускается, видимо, поправки какие-то должны также появиться.)

Как расчет магнитных свойств подтвердил справедливость квантовой механики?

(допишу про это потом, слышал, додумать предстоит, пока не до этого.)

Низкотемпературное приближение и слабые поля (?????)

(?? укажу в теории, как так обойти что не факторизуется статсумма через большое кан распределение????)

О низкотемпературном парамагнетизме Паули

(!?!?!? абзац про основную идею этой модели!)

Для типичных параметров магнитного поля и электронного газа в металле газ вырожден, а поле слабое: $\mu_B \mathcal{B} \ll T \ll \varepsilon_F$. Без поля то у половины электронов спин направлен по полю, а у других - против. После включения поля электроны приобретают дополнительную спиновую энергию $\pm\mu_B \mathcal{B}$, что эквивалентно замене μ на $\mu \pm \mu_B$. Поэтому, пока связанные с полем изменения химпотенциала малы,

$$\Omega(\mu) = \frac{1}{2}\Omega_{\mathcal{B}=0}\left(\mu + \mu_B(B)\right) + \frac{1}{2}\Omega_{\mathcal{B}=0}\left(\mu - \mu_B(B)\right) \approx \Omega_{\mathcal{B}=0}(\mu) + \frac{1}{2}\mu_B^2 \mathcal{B}^2 \frac{\partial^2 \Omega_{\mathcal{B}=0}}{\partial \mu^2},$$

$M = -\partial\Omega/\partial\mathcal{B}$, $\chi = M/\mathcal{B}$:

$$\chi_{\text{para}} = \mu_B^2 \cdot \left(\frac{\partial N}{\partial \mu}\right)_{T,V}. \qquad N \equiv -\partial\Omega/\partial\mu$$

Здесь χ_{para} относится ко всем N частицам. Полное вырождение электронного газа и слабость поля позволяет в этом же приближении $T \ll \varepsilon_F, \mu_B B \ll \varepsilon_F$

(?? тут важное рассуждение потеряно!)

дают $\mu(N) = \varepsilon_F = \left(3\pi^2\right)^{2/3} \frac{\hbar^2}{2m} \left(\frac{N}{V}\right)^{2/3}$,

$$\chi_{\text{пара}} = \mu_B^2 \frac{3N}{2\varepsilon_F} \equiv \mu_B^2 g\left(\varepsilon_F\right).$$

По сути включение слабого магнитного поля приводит увеличению числа электронов с магнитными спиновыми моментами «по» полю, чем электронов «против» поля на $g\left(\varepsilon_F\right)\mu_B\mathcal{B}$ штук. Это приводит к возникновению намагниченности $M = \mu_B^2 g\left(\varepsilon_F\right)\mathcal{B}$, отсюда и паулиевская восприимчивость.

О диамагнетизме Ландау при низких температурах и слабых полях

(!?!?!? абзац про основную идею этой модели!)

(???? что ниже с ур Ландау, чего такая формула для них???? в 2 раза больше? она скорее всего неверная, $\hbar$ из нее тоже должен выпадать)

Рассмотрим двумерный электронный газ в ящике, площадью S. Энергия электронов - уровни Ландау $\varepsilon_n = \mu_B \mathcal{B} \cdot (2n+1)$, кратность их вырождения с учетом спина $g_L = 2S\mathcal{B}/\Phi_0$, запишем омега-потенциал

$$\Omega = -T \sum_n \ln\left(1 + e^{\frac{\mu - \varepsilon_n}{T}}\right) \equiv 2\mu_B \mathcal{B} \cdot \sum_{n=0}^{\infty} f\left(\mu - \mu_B \mathcal{B} \cdot (2n+1)\right) \qquad f(X) := -\frac{TmS}{\hbar^2} \ln\left(1 + e^{\frac{X}{T}}\right)$$

Воспользуемся $\sum\limits_{n=0}^{\infty} f\left(n + \frac{1}{2}\right) = \int\limits_0^{\infty} f(x)dx + \frac{1}{24} \cdot f'(0) + \ldots$, получим:

$$\Omega = \Omega_0(\mu) - \frac{1}{6}\mu_B^2 \mathcal{B}^2 \frac{\partial^2 \Omega_0}{\partial \mu^2}.$$

Мы видим, что диамагнетизм Ландау в три раза слабее парамагнетизма Паули:

$$\chi_{\text{dia}} = -\chi_{\text{para}}/3, \qquad (s = 1/2),$$

(for other spins - different, for example, for s=3/2 $\chi_{\text{dia } s=3/2} = -\frac{8}{32}\chi_{\text{para } s=3/2}$)

так что, суммарная восприимчивость электронного газа при низких температурах $\chi = \chi_{\text{пара}} + \chi_{\text{dia}} = 2\chi_{\text{пара}}/3$ именно парамагнитна. Но некоторые металлы (а в них при комнатной температуре электронный газ как раз вырожден) могут быть диамагнитными, потому что в них спектр электронов $\varepsilon(\mathbf{p}) = \mathbf{p}^2/2m_*$ содержит эффективную массу m_*, она может быть существенно меньше m.

(???? а что с другими моделями магнетизма, они тоже есть, там магнетизм от атомов металла???? что с магнетизмом не для 2D газа, для 3D газа, например???)

Тогда,

$$\chi_{\text{para}} \sim \mu_B^2 \sim m^{-2} \qquad \chi_{\text{dia}} \sim \omega^2 \sim m_*^{-2} \qquad (\hbar\omega = 2\mu_B B)$$

В итоге, полная магнитная восприимчивость металла в принципе, может быть любого знака:

$$\chi = \chi_P \left(1 - m^2/3m_*^2\right),$$

Магнетизм в низкотемпературном приближении и сильных полях. Эффект де Гааза–ван Альфена (???)

(??? нет, это нужно заново читать оригинал, не понял)

Если магнитная энергия электронов больше температуры $T < \mu_B \mathcal{B} \ll \varepsilon_F$, то магнитное поле называют «квантующим». В этих условиях становится существенной дискретность уровней Ландау, что приводит к появлению у намагниченности электронного газа осциллирующей части. Амплитуда этих осцилляций не мала, а «шаг» осцилляций по обратному магнитному полю постоянен.

Рассмотрим самый простой случай: двумерный (D = 2) электронный газ при нулевой температуре ($T = 0$). Тогда в магнитном поле N электронов газа распределены по уровням Ландау следующим образом. На уровнях $0, 1, 2, \ldots, j$ «сидит» по g_L электронов,

а на последнем $j+1$-м уровне - оставшиеся $N-g_L(j+1)$ штук. Таким будет распределение электронов в интервале значений приложенного поля:

$$\frac{(j+1)}{\mathcal{B}_0} < \frac{1}{\mathcal{B}} < \frac{(j+2)}{\mathcal{B}_0},$$

где $\mathcal{B}_0 = N\Phi_0/2S$, S - площадь образца, $g_L = 2BS/\Phi_0$ - кратность вырождения уровня Ландау с учетом спина. Энергия основного состояния газа при $T=0$:

$$E = g_L \sum_{k=0}^{j} \hbar\omega \left(k+\frac{1}{2}\right) + \hbar\omega\left(j+\frac{3}{2}\right)(N - g_L(j+1)) =$$
$$= 2N\mu_B \left[\frac{\mathcal{B}}{\mathcal{B}_0}\left(j+\frac{3}{2}\right) - \left(\frac{\mathcal{B}}{\mathcal{B}_0}\right)^2 (j+1)\left(\frac{j}{2}+1\right)\right].$$

(?????? почему такая формула???? она тут точно не получается, там N^2 член получается точно, посмотрю по другой книге потом, не по Пухову)

$$M(\mathcal{B})|_{T=0} = -\partial F/\partial\mathcal{B} = -\partial E/\partial\mathcal{B} = N\cdot\frac{e\hbar}{mc}\cdot\left[(j+1)(j+2)\frac{\mathcal{B}}{\mathcal{B}_0} - j - \frac{3}{2}\right].$$

При изменении магнитного поля магнитный момент M осциллирует в интервале $\pm N\mu_B$ с постоянным по обратному полю шагом: последний уровень Ландау постепенно заполняется, пока число j скачком не увеличится на единицу, в следующем по $j \to j+1$ интервале обратных полей повторяется точно такая же зависимость $M(\mathcal{B})$.

$$\frac{1}{\mathcal{B}_0} = \frac{eS}{\pi\hbar cN} = \frac{\mu_B}{\varepsilon_F}.$$

Чтобы эти осцилляции были выражены и не «размывались» температурными эффектами, необходимо, чтобы поле было квантующим $T \ll \mu_B\mathcal{B}$, а электронный газ вырожденным $\mu_B\mathcal{B} \ll \varepsilon_F$.

Измеряя шаг осцилляций де Гааза-ван Альфена, можно судить о площади сечения поверхности Ферми. (??? почему???)

Шаг осцилляций де Гааза–ван Альфена очень мал, находится он так: осцилляции де Гааза-ван Альфена происходят при целочисленных отношениях числа электронов N к кратности вырождения g_L, т. е. при $1/B = 2Sn/N\Phi_0$, где $n = 1, 2, 3\ldots$. Это значит, что шаг осцилляций по обратному полю равен $\Delta(1/B) = 2S/N\Phi_0$. Выразим его через энергию Ферми. Она легко находится из 2D-условия: $\pi p_F^2 = N(2\pi\hbar)^2/2S$ для «клеток» одночастичных состояний. Отсюда:

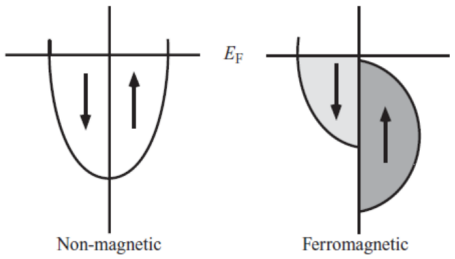
$$\varepsilon_F = \frac{\pi\hbar^2}{m}\frac{N}{S}.$$

(!! кст про ур Ландау в квантмехе допишу!!!)

2.2.3 Dia-, paramagnetism of conducting electrons, Pauli and Landau th.

Свойства электронов проводимости в металле

(это же в конд средах и статфизе (проверю, что написал!), типичная теория ферми сферы, т.п.!!!)



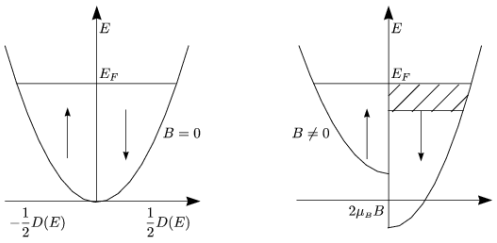
(?????? додумаю этот график????)

Pauli's paramagnetism (????????????)

(!!! нужно суть додумать!)
Magnetization induced by field B

$$M = \mu_B^2 D\left(E_F\right) \mathrm{B} = \mu_B^2 \frac{3n}{2\left(k_B T_F\right)^{3/2}} \sqrt{k_B T_F} B = \mu_B^2 \frac{3n}{2k_B T_F} B \qquad \mu_B := \frac{e\hbar}{2m_e}$$

Conduction e contribution to M is independent of Temperature
Quantitatively, need to add orbital effects, ionic cores, band effects, $e-e$ interactions...
Qualitatively, explains Paramagnetism of free spins, e.g. Na, K, Li (Alkali metals, with one free electron per atom)



For some alkali metals and noble metals, conduction electrons are weakly interacting and delocalized in space forming a Fermi gas. For a small magnetic field $\mathbf{H}$, the additional energy per electron from the interaction between an electron spin and the magnetic field is given by:

$$\Delta E = -\mu_0 \mathbf{H} \cdot \boldsymbol{\mu}_e = -\mu_0 \mathbf{H} \cdot \left(-g_e \frac{\mu_\mathrm{B}}{\hbar} \mathbf{S}\right) = \pm \mu_0 \mu_\mathrm{B} H$$

where μ_0 is the vacuum permeability, $\boldsymbol{\mu}_e$ is the electron magnetic moment, μ_B is the Bohr magneton, g-factor cancels with the spin $\mathbf{S} = \pm\hbar/2$. The $\pm$ indicates that the sign is positive (negative) when the electron spin component in the direction of $\mathbf{H}$ is parallel (antiparallel) to the magnetic field.

For low temperatures with respect to the Fermi temperature T_F ($\approx 10^4$ K for metals), the number density of electrons $n_\uparrow\left(n_\downarrow\right)$ pointing parallel (antiparallel) to the magnetic field can be written as:

$$n_\uparrow \approx \frac{n_e}{2} - \frac{\mu_0 \mu_\mathrm{B}}{2} g\left(E_\mathrm{F}\right) H; \qquad \left(n_\downarrow \approx \frac{n_e}{2} + \frac{\mu_0 \mu_\mathrm{B}}{2} g\left(E_\mathrm{F}\right) H\right),$$

with n_e the total free-electrons density and $g\left(E_\mathrm{F}\right)$ the electronic density of states (number of states per energy per volume) at the Fermi energy E_F.

In this approximation the magnetization is given as the magnetic moment of one electron times the difference in densities:

$$M = \mu_\mathrm{B}\left(n_\downarrow - n_\uparrow\right) = \mu_0 \mu_\mathrm{B}^2 g\left(E_\mathrm{F}\right) H,$$

which yields a positive paramagnetic susceptibility independent of temperature:

$$\chi_{\mathrm{P}} = \mu_0 \mu_{\mathrm{B}}^2 g\left(E_{\mathrm{F}}\right)$$

$$\boldsymbol{m} = \frac{\mu_{\mathrm{B}}}{2}\left[\int\limits_{-\mu_0\mu_{\mathrm{B}}H}^{E_{\mathrm{F}}} D\left(E+\mu_0\mu_{\mathrm{B}}H\right)dE - \int\limits_{+\mu_0\mu_{\mathrm{B}}H}^{E_{\mathrm{F}}} D\left(E-\mu_0\mu_{\mathrm{B}}H\right)dE\right]$$

$$\boldsymbol{m} = \mu_0\mu_{\mathrm{B}}^2 H D\left(E_{\mathrm{F}}\right) \quad D\left(E_{\mathrm{F}}\right) = \frac{3}{2}\frac{N}{E_{\mathrm{F}}}$$

Парамагнетизм свободных электронов

Диамагнетизм свободных электронов по Ландау

(спишу потом у Пухова, уже изучил и скорее всего в статфизе пропишсал)

Свойства переходных металлов

(см Боровика, потом додумаю)

2.2.4 Ferromagnetism, exchange interaction, Weiss theory (!)

Wiess theory

Main properties and ideas of Wiess theory

Ferromagnetism occurs when the energy cost of spin split is overcome by the exchange energy (with EF within the 3d band)

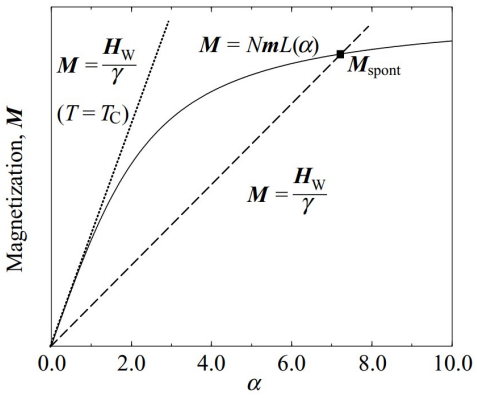
Weiss+Langevin (localized moment + molecular field) succeeds:

- correct temperature dependence (even better with Weiss+Brillouin)
- explains phase transition at T_C
- fails:
- moment per atom is equal in the paramagnetic and ferromagnetic phases
- moment per atom corresponds to an integer number of electrons (there is also delocalized moments = band magnetism)

(??? I'll write it better later)

Derivation of Wiess theory in a nutshell

$$\boldsymbol{M} = NmL(\alpha),$$
$$\alpha = \frac{mH_{\mathrm{W}}}{k_{\mathrm{B}}T} = \frac{m\gamma M}{k_{\mathrm{B}}T},$$
$$\boldsymbol{M} = \left(\frac{k_{\mathrm{B}}T}{\boldsymbol{m}_\gamma}\right)\alpha.$$



The Curie temperature can be determined by equating, at the origin, the slope of the magnetization described by the Langevin function (which is $\frac{1}{3}N\boldsymbol{m}$), with the slope of the straight line representing magnetization by the molecular field:

$$\frac{k_{\mathrm{B}}T_{\mathrm{C}}}{\boldsymbol{m}\gamma} = \frac{1}{3}N\boldsymbol{m},$$

$$T_{\mathrm{C}} = \frac{\gamma N m^2}{3k_{\mathrm{B}}}.$$

$$\gamma = \frac{3k_{\mathrm{B}}T_{\mathrm{C}}}{Nm^2}$$

$$\boldsymbol{H}_{\mathrm{W}} = \gamma\boldsymbol{M} = \gamma N\boldsymbol{m} = \frac{3k_{\mathrm{B}}T_{\mathrm{C}}}{\boldsymbol{m}}.$$

$$T_{\mathrm{C}} = \frac{\gamma N m_{\mathrm{eff}}^2}{3k_{\mathrm{B}}}$$

with $m_{\mathrm{eff}} = g\sqrt{J(J+1)}\mu_{\mathrm{B}}$. Multiplication by μ_0 produces the expression for T_{C} in SI units.

Теория Вейса по Пухову

(см. пока статфиз, хорошая теория)

Exchange interaction as a reason of ferromagnets

Types of exchange

Exchange interaction can be direct or indirect (RKKY, Super exchange), or others (double exchange, anisotropic, etc.(?)) more often it is indirect.

Electrons of neighbouring magnetic atoms interact via an exchange interaction. It don't need any intermediary.

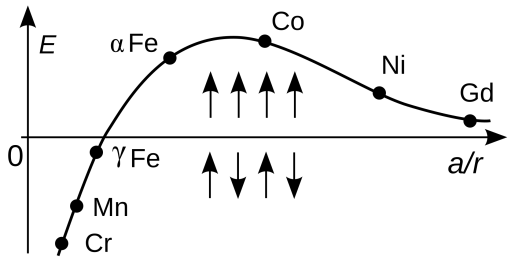
Direct interaction between neighbouring atoms is due to the spatial overlap of orbitals.

The simplest model for this kind of interactions is the Heisenberg model: $\mathbf{A} = -J\Sigma_{(ij)}\overrightarrow{s_i}\cdot\overrightarrow{s_j}$

Depending on J values some of the metals are FM and some are AFM (see, Bethe-Slater Curve).

The Bethe-Slater curve is a heuristic explanation for why certain metals are ferromagnetic and others are antiferromagnetic. It assumes a Heisenberg model of magnetism, and explains the differences in exchange energy of transition metals as due to the ratio of the interatomic

distance a to the radius r of the $3d$ electron shell. [1] When the magnetically important $3d$ electrons of adjacent atoms are relatively close to each other, the exchange interaction, J_{ex}, is negative, but when they are further away, the exchange interaction becomes positive, before slowly dropping off.



For occurrence of ferromagnetism the critical factor is

$$a/r_0 > 3$$

Element	Fe	Co	**Ni**	Cr	Mn	Gd
$\mathbf{a/r_0}$	3.26	3.64	3.94	2.6	2.94	3.12

(???)

Very often direct exchange cannot be an important mechanism in controlling the magnetic properties because there is insufficient direct overlap between neighbouring magnetic orbitals as in case of rare earths and transition metals. Then it becomes necessary to consider indirect exchange.

Молекула водорода и обменная энергия

(!!! основные идеи и формулы обменного взаимодействия!!!)

Let's make a rough classical estimate of the order of magnitude of the Coulomb repulsion between two electrons. If we assume that the average distance between electrons is around 1Å, then the Coulomb energy is

$$U = \frac{e^2}{4\pi\epsilon_0 r} \approx \frac{\left(1.6\cdot 10^{-19}\right)^2}{\left(1.1\cdot 10^{-10}\right)\left(1\cdot 10^{-10}\right)}\text{J} \approx 2.1\cdot 10^{-18}\ \text{J} = 1.4\cdot 10^{5}\ \text{K}.$$

This is about 10^5 times larger than the magnetic dipolar interaction which we calculated in Exercise 1.3(c). So if the electron distribution is changed even by a small amount, the effect on the total energy of an atom can be significant. This explains why the effective molecular field is so large!

$$\Psi\left(\mathbf{r}_1,\mathbf{r}_2\right) = \frac{1}{\sqrt{2}}\left[\phi_{1\ \mathrm{s}}\left(\mathbf{r}_1\right)\phi_{2\ \mathrm{s}}\left(\mathbf{r}_2\right) + \phi_{2\ \mathrm{s}}\left(\mathbf{r}_1\right)\phi_{1\ \mathrm{s}}\left(\mathbf{r}_2\right)\right]$$

$$\Psi\left(\mathbf{r}_1,\mathbf{r}_2\right) = \frac{1}{\sqrt{2}}\left[\phi_{1\ \mathrm{s}}\left(\mathbf{r}_1\right)\phi_{2\ \mathrm{s}}\left(\mathbf{r}_2\right) - \phi_{2\ \mathrm{s}}\left(\mathbf{r}_1\right)\phi_{1\ \mathrm{s}}\left(\mathbf{r}_2\right)\right]$$

(!! что ниже - перепишу, плохо написано, но пока нет времени, пока как в книге)

$$
\begin{aligned}
E &= \langle \Psi(\mathbf{r}_1, \mathbf{r}_2) | H | \Psi(\mathbf{r}_1, \mathbf{r}_2) \rangle \\
&= \frac{1}{2} \langle [\phi_{1\,\mathrm{s}}(\mathbf{r}_1)\phi_{2\,\mathrm{s}}(\mathbf{r}_2) \pm \phi_{2\,\mathrm{s}}(\mathbf{r}_1)\phi_{1\,\mathrm{s}}(\mathbf{r}_2)] \\
&\quad |(H_1 + H_2 + H_{12})| [\phi_{1\,\mathrm{s}}(\mathbf{r}_1)\phi_{2\,\mathrm{s}}(\mathbf{r}_2) \pm \phi_{2\,\mathrm{s}}(\mathbf{r}_1)\phi_{1\,\mathrm{s}}(\mathbf{r}_2)] \rangle \\
&= \frac{1}{2} [\langle \phi_{1\,\mathrm{s}}(\mathbf{r}_1) | H_1 | \phi_{1\,\mathrm{s}}(\mathbf{r}_1) \rangle + \langle \phi_{2\,\mathrm{s}}(\mathbf{r}_1) | H_1 | \phi_{2\,\mathrm{s}}(\mathbf{r}_1) \rangle \\
&\quad + \langle \phi_{1\,\mathrm{s}}(\mathbf{r}_2) | H_2 | \phi_{1\,\mathrm{s}}(\mathbf{r}_2) \rangle + \langle \phi_{2\,\mathrm{s}}(\mathbf{r}_2) | H_2 | \phi_{2\,\mathrm{s}}(\mathbf{r}_2) \rangle \\
&\quad + \langle \phi_{1\,\mathrm{s}}(\mathbf{r}_1)\phi_{2\,\mathrm{s}}(\mathbf{r}_2) | H_{12} | \phi_{1\,\mathrm{s}}(\mathbf{r}_1)\phi_{2\,\mathrm{s}}(\mathbf{r}_2) \rangle \\
&\quad + \langle \phi_{2\,\mathrm{s}}(\mathbf{r}_1)\phi_{1\,\mathrm{s}}(\mathbf{r}_2) | H_{12} | \phi_{2\,\mathrm{s}}(\mathbf{r}_1)\phi_{1\,\mathrm{s}}(\mathbf{r}_2) \rangle \\
&\quad \pm \langle \phi_{1\,\mathrm{s}}(\mathbf{r}_1)\phi_{2\,\mathrm{s}}(\mathbf{r}_2) | H_{12} | \phi_{2\,\mathrm{s}}(\mathbf{r}_1)\phi_{1\,\mathrm{s}}(\mathbf{r}_2) \rangle \\
&\quad \pm \langle \phi_{2\,\mathrm{s}}(\mathbf{r}_1)\phi_{1\,\mathrm{s}}(\mathbf{r}_2) | H_{12} | \phi_{1\,\mathrm{s}}(\mathbf{r}_1)\phi_{2\,\mathrm{s}}(\mathbf{r}_2) \rangle] \\
&= E_1 + E_2 + K \pm J,
\end{aligned}
$$

$$
\begin{aligned}
E_1 &= \langle \phi_{1\,\mathrm{s}}(\mathbf{r}_1) | H_1 | \phi_{1\,\mathrm{s}}(\mathbf{r}_1) \rangle = \langle \phi_{1\,\mathrm{s}}(\mathbf{r}_2) | H_2 | \phi_{1\,\mathrm{s}}(\mathbf{r}_2) \rangle \\
E_2 &= \langle \phi_{2\,\mathrm{s}}(\mathbf{r}_1) | H_1 | \phi_{2\,\mathrm{s}}(\mathbf{r}_1) \rangle = \langle \phi_{2\,\mathrm{s}}(\mathbf{r}_2) | H_2 | \phi_{2\,\mathrm{s}}(\mathbf{r}_2) \rangle \\
K &= \langle \phi_{1\,\mathrm{s}}(\mathbf{r}_1)\phi_{2\,\mathrm{s}}(\mathbf{r}_2) | H_{12} | \phi_{1\,\mathrm{s}}(\mathbf{r}_1)\phi_{2\,\mathrm{s}}(\mathbf{r}_2) \rangle \\
&= \langle \phi_{2\,\mathrm{s}}(\mathbf{r}_1)\phi_{1\,\mathrm{s}}(\mathbf{r}_2) | H_{12} | \phi_{2\,\mathrm{s}}(\mathbf{r}_1)\phi_{1\,\mathrm{s}}(\mathbf{r}_2) \rangle \\
J &= \langle \phi_{1\,\mathrm{s}}(\mathbf{r}_1)\phi_{2\,\mathrm{s}}(\mathbf{r}_2) | H_{12} | \phi_{2\,\mathrm{s}}(\mathbf{r}_1)\phi_{1\,\mathrm{s}}(\mathbf{r}_2) \rangle \\
&= \langle \phi_{2\,\mathrm{s}}(\mathbf{r}_1)\phi_{1\,\mathrm{s}}(\mathbf{r}_2) | H_{12} | \phi_{1\,\mathrm{s}}(\mathbf{r}_1)\phi_{2\,\mathrm{s}}(\mathbf{r}_2) \rangle .
\end{aligned}
$$

E_1 and E_2 represent the energies of the 1 s and 2 s orbitals, respectively, in the field of the helium nucleus; K is the Coulomb interaction between the electron densities $\phi^2_{1\,\mathrm{s}}$ and $\phi^2_{2\,\mathrm{s}}$; and J is the exchange interaction, which clearly has no classical analog.

Обменное взаимодействие и критерий ферромагнетизма

Conditions for FM

1. Existence of incomplete atomic shells
2. Positive exchange integral which causes parallel orientation of spins
3. $a/r_0 > 3$

Direct Exchange

Electrons of neighbouring magnetic atoms interact via an exchange interaction. It don't need any intermediary.

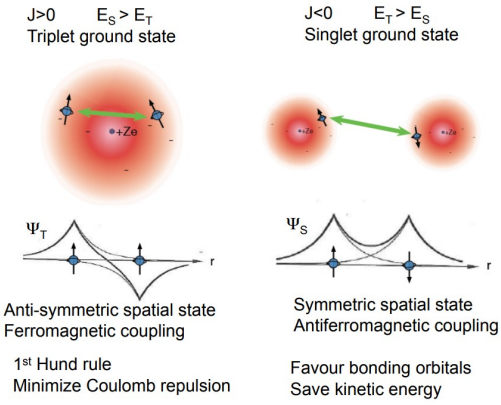
Direct interaction between neighbouring atoms is due to the spatial overlap of orbitals.

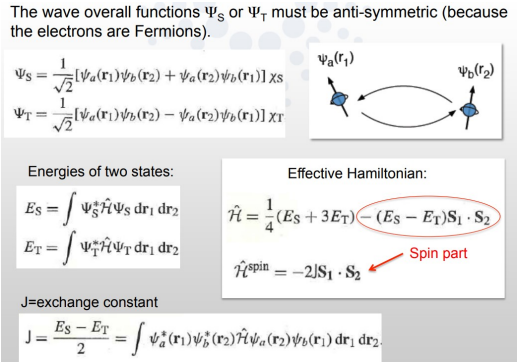
The simplest model for this kind of interactions is the

Depending on J values some of the metals are FM and some are AFM (see, Bethe-Slater Curve).

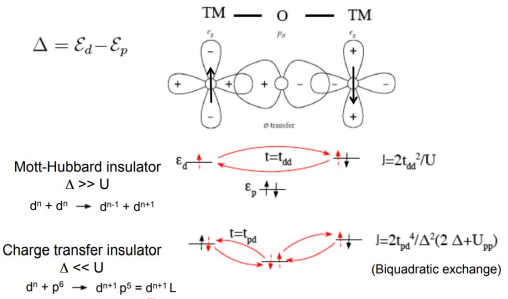
Very often direct exchange cannot be an important mechanism in controlling the magnetic properties because there is insufficient direct overlap between neighbouring magnetic orbitals as in case of rare earths and transition metals. Then it becomes necessary to consider indirect exchange.

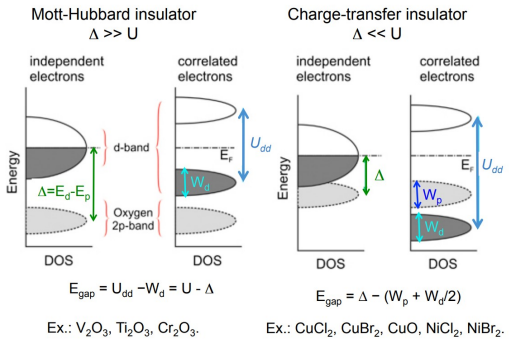
(???? still, why direct exchange is a rare mechanism? in other words, why there is typically insufficient direct overlap between neighboring magnetic orbitals????)





(see wonderful Paolsini lectures (use Google) about it!!!)





Зависимость свойств ферромагнетиков от температуры.

Атомные магнитные моменты ферромагнитных элементов

Кривая намагничивания ферромагнетика

(пока по Ландау)

$$\widetilde{\Phi} = \Phi_0(M) + \frac{\beta}{2}M_x^2 - \mathbf{HM} - \frac{H^2}{8\pi} =$$
$$= \Phi_0(M) + \frac{\beta M^2}{2}\sin^2\theta - M\left(H_x\sin\theta + H_z\cos\theta\right) - \frac{H^2}{8\pi}$$

Зависимость **M** от **H** определяется условием равновесия $\partial\widetilde{\Phi}/\partial\theta = 0$, откуда

$$\beta M\sin\theta\cos\theta = H_x\cos\theta - H_z\sin\theta.$$

По отношению к неизвестной $\xi = \sin\theta$ это есть алгебраическое уравнение четвертой степени

$$(\beta M\xi - H_x)^2\left(1-\xi^2\right) = H_z^2\xi^2$$

$$\frac{\partial}{\partial\theta} \quad \Rightarrow \quad \begin{gathered} \frac{H_x}{\sin\theta} - \frac{H_z}{\cos\theta} = \beta M \\ \frac{H_x}{\sin^3\theta} = -\frac{H_z}{\cos^3\theta}. \\ H_x^{2/3} + H_z^{2/3} = (\beta M)^{2/3} \end{gathered}$$

$$\widetilde{\Phi} = \Phi_0 - \frac{H^2}{8\pi} + \frac{\beta M^2}{2}\sin^2\theta - HM\sin\theta.$$

Если $H > \beta M$, то $\widetilde{\Phi}$ имеет лишь один минимум - при $\theta = \pi/2$, т. е. намагниченность направлена вдоль поля. Если же $H < \beta M$, то $\widetilde{\Phi}$ имеет минимум при

$$M_x = M\sin\theta = \frac{H}{\beta},$$

$$\bar{B}_x = H_x\left(1 + \frac{4\pi}{\beta}\right), \quad \bar{B}_z < 4\pi\sqrt{M^2 - \frac{H_x^2}{\beta^2}}.$$

Этими формулами определяется область значений средней индукции, соответствующая доменной структуре одноосного ферромагнетика.

(????? как) Определить область значений H, при которых тело будет обладать доменной структурой

Ферромагнетик вблизи точки Кюри по ЛЛ8.39

Энергия магнитной анизотропии по ЛЛ8.40

Кривая намагничения ферромагнетиков по ЛЛ8.41

Other exchange interactions (?)

(later maybe I'll create a subsubsection for each)

RKKY

Super exchange

double exchange

anisotropic

2.2.5 Antiferromagnetism, ferrimagnetism (!!)

(!!! короче, разбираться нужно, не малая тема)

Антиферромагнетизм fundamentals

У нас есть T_N, below which sublatices are antialigned under each other's molecular field.

under zero field H: net magnetization is always zero

under field: susceptibility depends on the angle between the direction of spontaneous magnetization of the sublattices and the direction of H

We have two sublattices, A and B, and their A-B interactions:

$$\boldsymbol{H}_{\mathrm{W}}^{A} = -\gamma \boldsymbol{M}_B \qquad \boldsymbol{H}_{\mathrm{W}}^{B} = -\gamma \boldsymbol{M}_A$$

Above the Néel temperature (T_{N}) there is a Curie law $\chi \sim 1/T$, so $\boldsymbol{M} \sim \boldsymbol{H}/\mathrm{T}$

$$\boldsymbol{M}_A = \frac{C'\left(\boldsymbol{H} - \gamma \boldsymbol{M}_B\right)}{T}; \qquad \boldsymbol{M}_B = \frac{C'\left(\boldsymbol{H} - \gamma \boldsymbol{M}_A\right)}{T}$$

$$\boldsymbol{M} = \boldsymbol{M}_A + \boldsymbol{M}_B;$$

$$\boldsymbol{M} = \frac{2C'\boldsymbol{H}}{T + C'\gamma}$$

$$\chi = \frac{\boldsymbol{M}}{\boldsymbol{H}} = \frac{2C'}{T + C'\gamma} = \frac{C}{T + \theta} \qquad \text{(Curie-Weiss law, } \theta < 0)$$

At T_{N} in field $H = 0$: $\boldsymbol{M}_A = \frac{-C'\gamma \boldsymbol{M}_B}{T_{\mathrm{N}}} = \frac{-\theta \boldsymbol{M}_B}{T_{\mathrm{N}}}$, in an antiferromagnet now $\boldsymbol{M} = 0$, so $\boldsymbol{M}_A = -\boldsymbol{M}_B$, so $\theta = T_{\mathrm{N}}$. But this is so only within the Weiss theory if we only consider $A - B$ interactions (nearest neighbors). In practice, if one accounts for A-A, B-B and other interactions, $T_{\mathrm{N}} < \theta$.

- below T_{N}: sublatices are antialigned under each other's molecular field

- under zero field H: net magnetization is always zero

- under field: susceptibility depends on the angle between the direction of spontaneous magnetization of the sublattices and the direction of H

- H parallel to spontaneous magnetizations: Langevin (or Brillouin) function dependence on H and T If H is applied parallel to $\boldsymbol{M}_A$: increase in $\boldsymbol{M}_A$ and decrease in $\boldsymbol{M}_B$

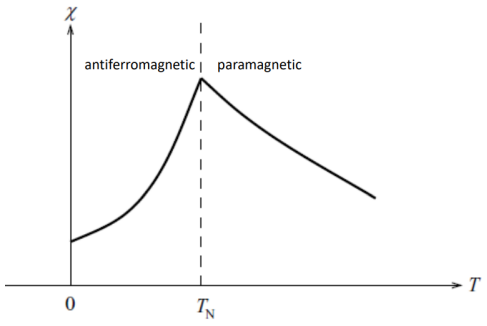
$$\boldsymbol{M} = \boldsymbol{M}_A - \boldsymbol{M}_B = |\delta \boldsymbol{M}_A| + |\delta \boldsymbol{M}_B|$$

$$\chi_{\parallel} = \frac{2Nm^2 B'(J,\alpha)}{2k_{\mathrm{B}}T + Nm^2\gamma B'(J,\alpha)}$$

$$\delta\alpha = \frac{m}{k_{\mathrm{B}}T}\delta H = \frac{m}{k_{\mathrm{B}}T}\left(H - \gamma\,|\delta M_B|\right)$$

$$\boldsymbol{H} = \boldsymbol{H}_{\mathrm{W}} = 2 \times \boldsymbol{H}_{\mathrm{W}}^{A} \sin\theta = 2\gamma \boldsymbol{M}_A \sin\theta = \gamma M$$

(since $\boldsymbol{M} = 2\boldsymbol{M}_A \sin\theta$)



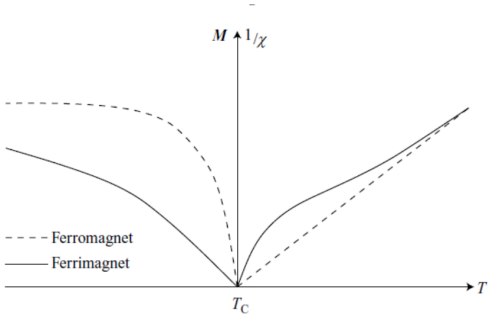
Properties of T_c of atiferromagnet if we have addings (?!?!)

(like how will it change if we account for something???)

If we add NNN interaction, then (???)

(просто подумаю - и пойму ответы, пока не до этого)

Ferrimagnetism fundamentals



(???? why such features of the plot???)

Вводим постоянные альфа и бета, показывающие, насколько велики спины, создающие намагниченность:

$$\boldsymbol{M}_A = \alpha n \mu_A,$$
$$\boldsymbol{M}_B = \beta n \mu_B.$$
$$\boldsymbol{M} = \boldsymbol{M}_A + \boldsymbol{M}_B = \alpha n \mu_A + \beta n \mu_B.$$

Again there are two Weiss molecular fields, one acting on each of the A and B sublattices, but they are no longer equal in magnitude:

$$\boldsymbol{H}^A_{\mathrm{W}} = -\gamma_{AB}\boldsymbol{M}_B + \gamma_{AA}\boldsymbol{M}_A; \qquad \boldsymbol{H}^B_{\mathrm{W}} = -\gamma_{AB}\boldsymbol{M}_A + \gamma_{BB}\boldsymbol{M}_B.$$
$$\boldsymbol{M}_A = \frac{C\left(\boldsymbol{H} + \boldsymbol{H}^A_{\mathrm{W}}\right)}{T}, \qquad \boldsymbol{M}_B = \frac{C\left(\boldsymbol{H} + \boldsymbol{H}^B_{\mathrm{W}}\right)}{T}.$$

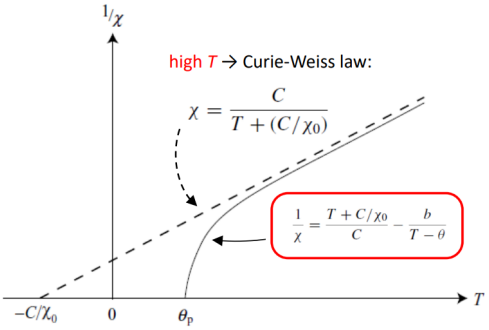
Solving for $\boldsymbol{M} = \boldsymbol{M}_A + \boldsymbol{M}_B$, we have for $\chi := \frac{\boldsymbol{M}}{\boldsymbol{H}}$

$$\frac{1}{\chi} = \frac{T + C/\chi_0}{C} - \frac{b}{T - \theta},$$

(???? this didn't work out for me, maybe later I'll solve it, need to read Spaldin)

$$\frac{1}{\chi_0} := \gamma_{AB}\left(2\alpha\beta - \frac{\gamma_{AA}}{\gamma_{AB}}\alpha^2 - \frac{\gamma_{BB}}{\gamma_{AB}}\beta^2\right); \quad b := \gamma_{AB}^2 C\alpha\beta\left[\alpha\left(1 + \frac{\gamma_{AA}}{\gamma_{AB}}\right) - \beta\left(1 + \frac{\gamma_{BB}}{\gamma_{AB}}\right)\right]^2; \quad \theta := \gamma_{AB}C\alpha\beta\left(2 + \frac{\gamma_{AA}}{\gamma_{AB}} + \frac{\gamma_{BB}}{\gamma_{AB}}\right)$$

Например, это (Fe_3O_4), из которого делают постоянные магниты, также это т.н. ферриты, имеющие общую формулу $MeOFe_2O_3$, где Me - ион двухвалентного металла, например Mg^{2+} Ni^{2+}, Mn^{2+} и т. д. (в случае магнетита это Fe^{2+}). Некоторые из ферритов, например $ZnOFe_2O_3$ не ферромагнитны.



Below T_C: - each sublattice is spontaneously magnetized: non-zero total magnetization (??? чего тогда на графике в ноль мы приходим???? и что такое там тета р???)

$$M = |M_A| - |M_B|$$

- each sublattice: Brillouin behavior under Weiss field

$$\boldsymbol{M}_A = N\boldsymbol{m}_A B\left(J, \frac{m_A H_{\mathrm{W}}^A}{k_{\mathrm{B}}T}\right)$$
$$\boldsymbol{M}_B = N\boldsymbol{m}_B B\left(J, \frac{m_B H_{\mathrm{W}}^B}{k_{\mathrm{B}}T}\right)$$

2.2.6 Typical Phase Transitions in Magnetism

Суть описания фазовых переходов

$$\Omega(T, \mathcal{M}) = \Omega_0(T) + \alpha(T)\cdot\mathcal{M}^2 + \frac{1}{2}b(T)\cdot\mathcal{M}^4 - \mathcal{M}\cdot\mathcal{B}.$$

Здесь $\Omega_0(T)$ часть Ω-потенциала, не связанная с магнетизмом, нечётных степеней $\mathcal{M}$ нет из-за изотропности ферромагнетика. Теперь $\mathcal{M}$ - независимая переменная, и равновесию соответствует минимум Ω-потенциала по $\mathcal{M}$. Пусть сначала поля нет: $\mathcal{B} = 0$, тогда

$$\frac{\partial\Omega}{\partial\mathcal{M}} = 2\alpha(T)\cdot\mathcal{M} + 2b(T)\cdot\mathcal{M}^3 = 0.$$

Для правильного описания перехода достаточно предположить, что $\alpha(T)$ меняет знак, а $b(T)$ постоянно в окрестности точки Кюри:

$$\alpha(T) = a\left(T - T_c\right), \qquad b(T) = b, \qquad a, b > 0.$$

Тогда

$$\mathcal{M}_0(T) = \begin{cases} \sqrt{\frac{a(T_c - T)}{b}}, & T < T_c, \\ 0, & T > T_c, \end{cases}$$

$$\Omega(T) = \begin{cases} \Omega_0(T) - \frac{a^2}{2b}\left(T_c - T\right)^2, & T < T_c, \\ \Omega_0(T), & T > T_c, \end{cases}$$

$$-T\partial^2\Omega/\partial T^2 \equiv C(T) = \begin{cases} C_0(T) + \frac{a^2}{b}T, & T < T_c, \\ C_0(T), & T > T_c. \end{cases}$$

В присутствии внешнего поля

$$\frac{\partial\Omega}{\partial\mathcal{M}} = 2\alpha(T)\cdot\mathcal{M} + 2b(T)\cdot\mathcal{M}^3 - \mathcal{B} = 0$$

появляется дополнительная намагниченность к спонтанной: $\mathcal{M} = \mathcal{M}_0 + \delta\mathcal{M}$. В небольших полях $\delta\mathcal{M} = \chi\cdot\mathcal{B}$ для восприимчивости χ получаем закон Кюри-Вейсса:

$$\chi(T) = \begin{cases} \frac{1}{2a(T_c - T)}, T < T_c, \\ \frac{1}{a(T - T_c)}, T > T_c. \end{cases}$$

(графики добавлю!!!!!)

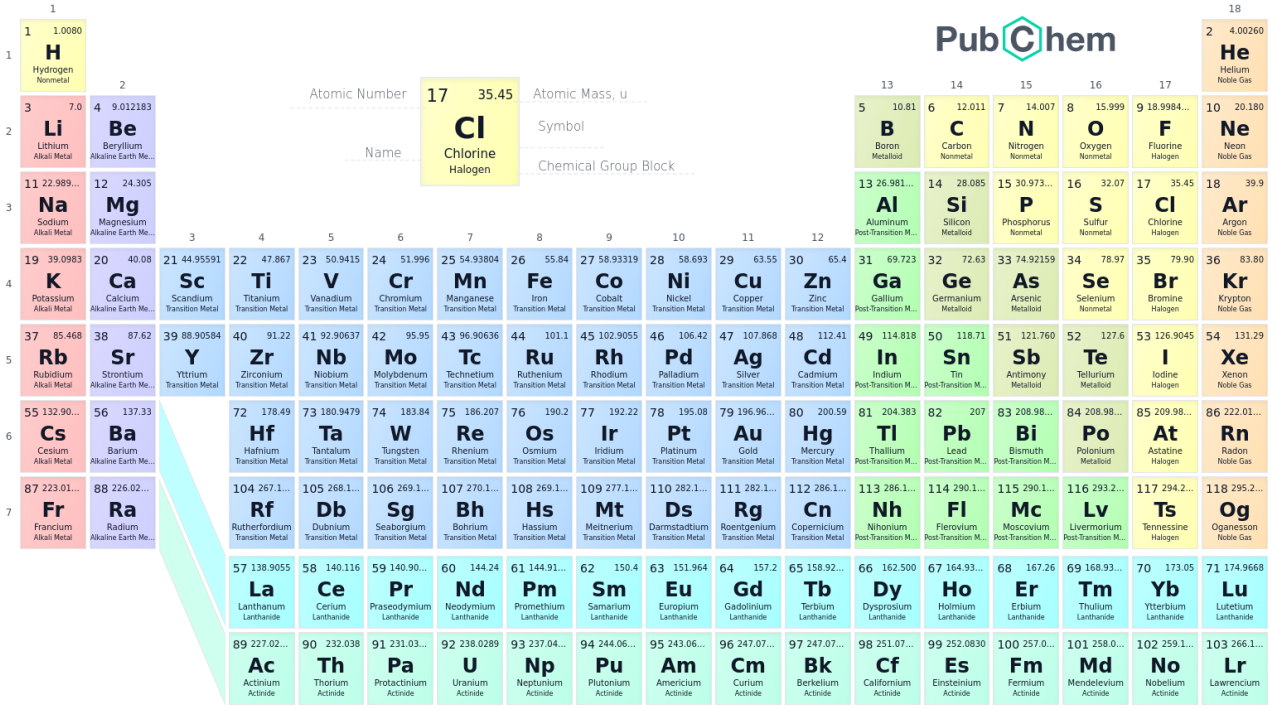
2.2.7 Magnetic Symmetries, Magnetic Classes, and Spatial Groups

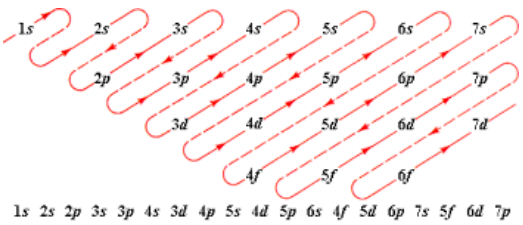
(ЛЛ8 пар 37, 38)
(потом поразбираюсь, подумать про это много нужно, пока не нужно это было)

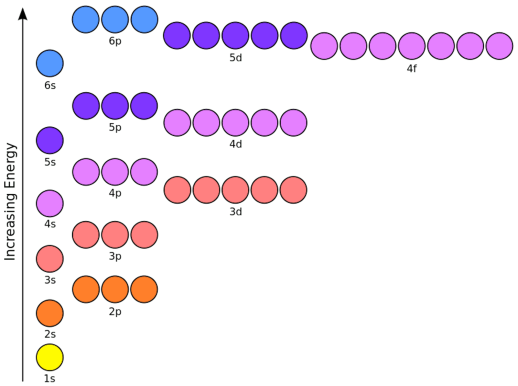
2.2.8 Atoms, matter, electrons, parameters, units background

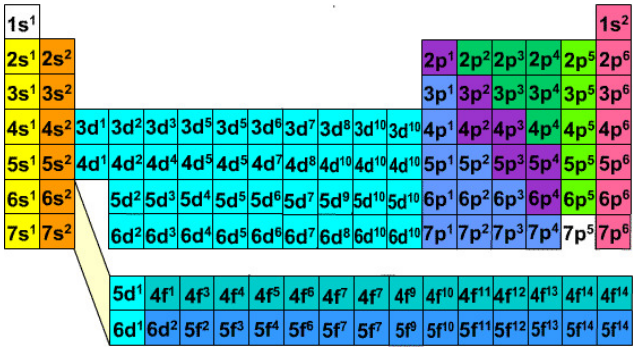
(суть общефиза и ЭД, в них это же, но подробнее!)

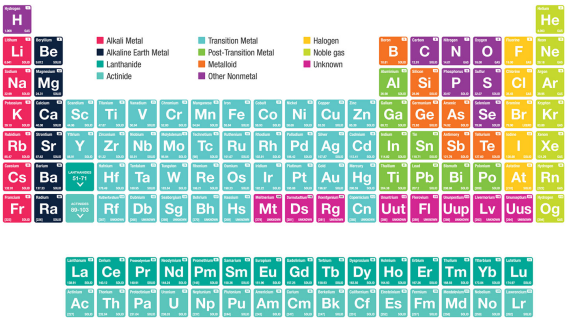
Многоэлектронные атомы

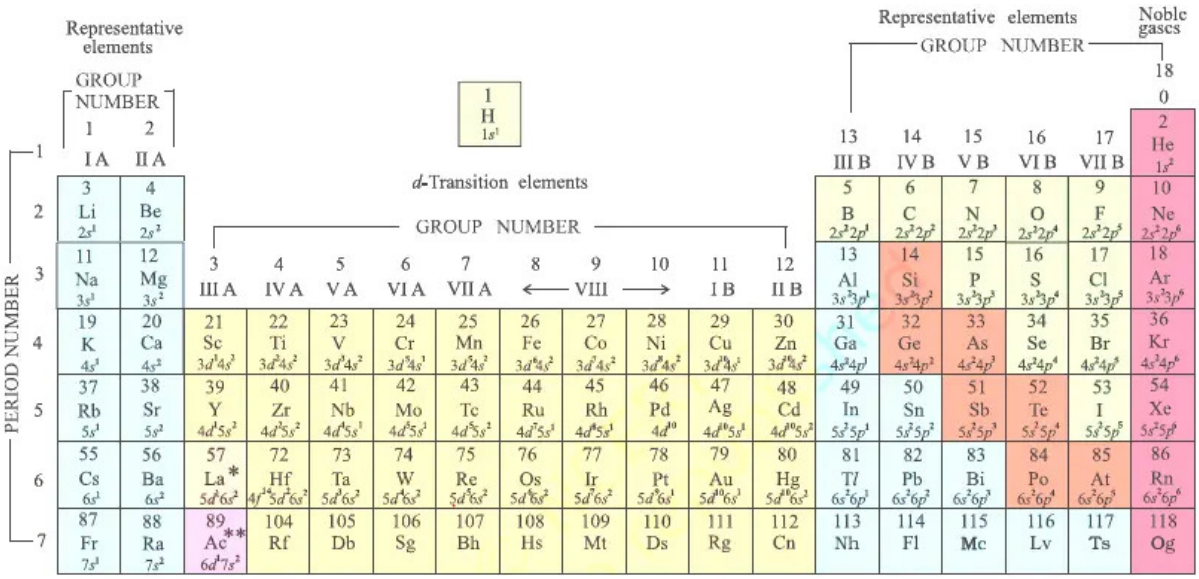


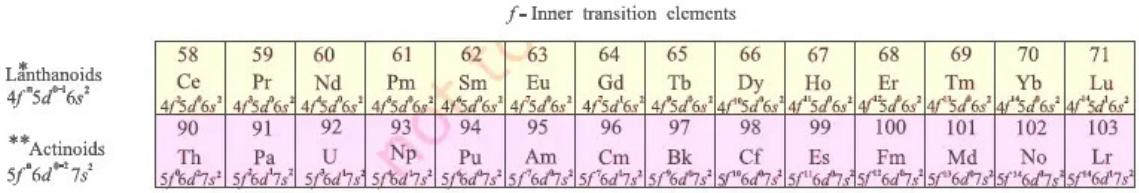


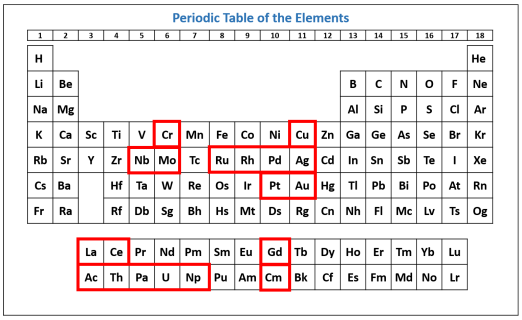


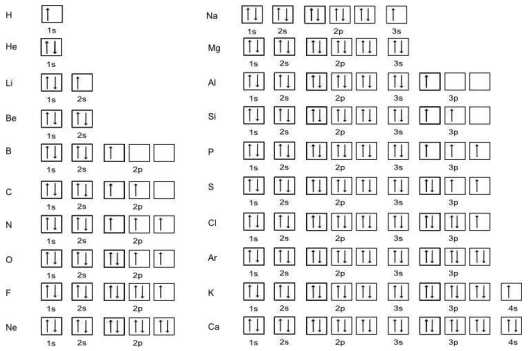

1s1 1s2
2s1 2s2 2p1 2p2 2p3 2p4 2p5 2p6
3s1 3s2 3p1 3p2 3p3 3p4 3p5 3p6
4s1 4s2 3d1 3d2 3d3 3d5 3d5 3d6 3d7 3d8 3d10 3d10 4p1 4p2 4p3 4p4 4p5 4p6
5s1 5s2 4d1 4d2 4d4 4d5 4d5 4d7 4d8 4d10 4d10 4d10 5p1 5p2 5p3 5p4 5p5 5p6
6s1 6s2 5d2 5d3 5d4 5d5 5d6 5d7 5d9 5d10 5d10 6p1 6p2 6p3 6p4 6p5 6p6
7s1 7s2 6d2 6d3 6d4 6d5 6d6 6d7 6d8 6d10 6d10 7p1 7p2 7p3 7p4 7p5 7p6
5d1 4f1 4f3 4f4 4f5 4f6 4f7 4f7 4f9 4f10 4f11 4f12 4f13 4f14 4f14
6d1 6d2 5f2 5f3 5f4 5f6 5f7 5f7 5f9 5f10 5f11 5f12 5f13 5f14 5f14


Alkali Metal
Alkaline Earth Metal
Lanthanide
Actinide
Transition Metal
Post-Transition Metal
Metalloid
Other Nonmetal
Halogen
Noble gas
Unknown


Representative elements
GROUP NUMBER
IA IIA
d-Transition elements
GROUP NUMBER
IIIA IVA VA VIA VIIA VIII IB IIB
Representative elements
GROUP NUMBER
IIIB IVB VB VIB VIIB
Noble gases
PERIOD NUMBER
1 H 1s1
2 He 1s2
3 Li 2s1 4 Be 2s2 5 B 2s22p1 6 C 2s22p2 7 N 2s22p3 8 O 2s22p4 9 F 2s22p5 10 Ne 2s22p6
11 Na 3s1 12 Mg 3s2 13 Al 3s23p1 14 Si 3s23p2 15 P 3s23p3 16 S 3s23p4 17 Cl 3s23p5 18 Ar 3s23p6
19 K 4s1 20 Ca 4s2 21 Sc 3d14s2 22 Ti 3d24s2 23 V 3d34s2 24 Cr 3d54s1 25 Mn 3d54s2 26 Fe 3d64s2 27 Co 3d74s2 28 Ni 3d84s2 29 Cu 3d104s1 30 Zn 3d104s2 31 Ga 4s24p1 32 Ge 4s24p2 33 As 4s24p3 34 Se 4s24p4 35 Br 4s24p5 36 Kr 4s24p6
37 Rb 5s1 38 Sr 5s2 39 Y 4d15s2 40 Zr 4d25s2 41 Nb 4d45s1 42 Mo 4d55s1 43 Tc 4d55s2 44 Ru 4d75s1 45 Rh 4d85s1 46 Pd 4d10 47 Ag 4d105s1 48 Cd 4d105s2 49 In 5s25p1 50 Sn 5s25p2 51 Sb 5s25p3 52 Te 5s25p4 53 I 5s25p5 54 Xe 5s25p6
55 Cs 6s1 56 Ba 6s2 57 La * 5d16s2 72 Hf 4f145d26s2 73 Ta 5d36s2 74 W 5d46s2 75 Re 5d56s2 76 Os 5d66s2 77 Ir 5d76s2 78 Pt 5d96s1 79 Au 5d106s1 80 Hg 5d106s2 81 Tl 6s26p1 82 Pb 6s26p2 83 Bi 6s26p3 84 Po 6s26p4 85 At 6s26p5 86 Rn 6s26p6
87 Fr 7s1 88 Ra 7s2 89 Ac ** 6d17s2 104 Rf 105 Db 106 Sg 107 Bh 108 Hs 109 Mt 110 Ds 111 Rg 112 Cn 113 Nh 114 Fl 115 Mc 116 Lv 117 Ts 118 Og


f-Inner transition elements
* Lanthanoids 4f n5d1-06s2
** Actinoids 5f n6d1-07s2
58 Ce 4f15d16s2 59 Pr 4f35d06s2 60 Nd 4f45d06s2 61 Pm 4f55d06s2 62 Sm 4f65d06s2 63 Eu 4f75d06s2 64 Gd 4f75d16s2 65 Tb 4f95d06s2 66 Dy 4f105d06s2 67 Ho 4f115d06s2 68 Er 4f125d06s2 69 Tm 4f135d06s2 70 Yb 4f145d06s2 71 Lu 4f145d16s2
90 Th 6d27s2 91 Pa 5f26d17s2 92 U 5f36d17s2 93 Np 5f46d17s2 94 Pu 5f67s2 95 Am 5f77s2 96 Cm 5f76d17s2 97 Bk 5f97s2 98 Cf 5f106d07s2 99 Es 5f116d07s2 100 Fm 5f126d07s2 101 Md 5f136d07s2 102 No 5f146d07s2 103 Lr 5f146d17s2


Periodic Table of the Elements
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
H He
Li Be B C N O F Ne
Na Mg Al Si P S Cl Ar
K Ca Sc Ti V Cr Mn Fe Co Ni Cu Zn Ga Ge As Se Br Kr
Rb Sr Y Zr Nb Mo Tc Ru Rh Pd Ag Cd In Sn Sb Te I Xe
Cs Ba Hf Ta W Re Os Ir Pt Au Hg Tl Pb Bi Po At Rn
Fr Ra Rf Db Sg Bh Hs Mt Ds Rg Cn Nh Fl Mc Lv Ts Og
La Ce Pr Nd Pm Sm Eu Gd Tb Dy Ho Er Tm Yb Lu
Ac Th Pa U Np Pu Am Cm Bk Cf Es Fm Md No Lr



Period	Element	Configuration						
1	He	$1s^2$						
2	Ne	$1s^2$	$2s^2\ 2p^6$					
3	Ar	$1s^2$	$2s^2\ 2p^6$	$3s^2\ 3p^6$				
4	Kr	$1s^2$	$2s^2\ 2p^6$	$3s^2\ 3p^6$	$4s^2\ 3d^{10}\ 4p^6$			
5	Xe	$1s^2$	$2s^2\ 2p^6$	$3s^2\ 3p^6$	$4s^2\ 3d^{10}\ 4p^6$	$5s^2\ 4d^{10}\ 5p^6$		
6	Rn	$1s^2$	$2s^2\ 2p^6$	$3s^2\ 3p^6$	$4s^2\ 3d^{10}\ 4p^6$	$5s^2\ 4d^{10}\ 5p^6$	$6s^2\ 4f^{14}\ 5d^{10}\ 6p^6$	
7	Og	$1s^2$	$2s^2\ 2p^6$	$3s^2\ 3p^6$	$4s^2\ 3d^{10}\ 4p^6$	$5s^2\ 4d^{10}\ 5p^6$	$6s^2\ 4f^{14}\ 5d^{10}\ 6p^6$	$7s^2\ 5f^{14}\ 6d^{10}\ 7p^6$

transition metals give up their 4 s electrons before their 3 d electrons on ionization. (??? это додумаю, проверю потом, как работает? пока просто считаю, что ионы если образуются, то именно от 4s электронов.)

1 Hund Rule for a ground state: в первом порядке ТВ энергия атома имеет минимальное значения в состояниях с максимальным значением полного спина S, а также с максимальным значением полного момента L.

2 Hund Rule for a ground state: (о влиянии тонкого расщепления) во втором порядке ТВ минимальную энергию имеет терм с $J = |L - S|$, если электронная оболочка заполнена меньше, чем наполовину, и с $J = L + S$, если она заполнена на половину или более.

(!!!! here is about why measured moment is not the same as we calculate using QM?!?!?!?!?)
-because of quenching (see Spalding sect 5.3)

(??? how in general can we measure total momentum or projection of a total momentum?????)

Планетарная модель, электрон как классическая частица

(суть того, что в общефизе в 1й части подробнее раза в 4)

Спин. Собственный магнитный момент электрона

(это вроде вообще не нужно)

$$p_s = \sqrt{s(s+1)}\frac{h}{2\pi} = \sqrt{\frac{1}{2}\left(\frac{1}{2}+1\right)}\frac{h}{2\pi},$$

$$\mu_{sH} = 2m_s\mu_\mathrm{B}.$$

Для спинового момента отношение магнитной составляющей к механической вдвое больше, чем для орбитального:

$$\frac{\mu_s}{p_s} = \frac{2\sqrt{s(s+1)}\mu_\mathrm{B}}{\sqrt{s(s+1)}h/(2\pi)} = \frac{e}{mc}$$

Магнитный момент атома. Фактор Ланде

$$\boldsymbol{\mu}_j = \mu_L \cos\left(\widehat{\mathbf{P}_L\mathbf{P}_J}\right) + \mu_S \cos\left(\widehat{\mathbf{P}_S\mathbf{P}_J}\right);$$

$$\mu_L = \sqrt{L(L+1)}\mu_\mathrm{B}$$

$$\mu_S = 2\sqrt{S(S+1)}\mu_\mathrm{B}$$

$$\begin{cases} \cos\left(\widehat{\mathbf{P}_L\mathbf{P}_J}\right) = \frac{L(L+1)+J(J+1)-S(S+1)}{2\sqrt{L(L+1)}\sqrt{J(J+1)}}, \\ \cos\left(\widehat{\mathbf{P}_S\mathbf{P}_J}\right) = \frac{S(S+1)+J(J+1)-L(L+1)}{2\sqrt{S(S+1)}\sqrt{J(J+1)}}. \end{cases}$$

$$g = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}$$

Эффект Зеемана

Поместим атом в магнитное поле, к его энергии добавится:

$$W_H = -\mu_H H = g m_J \mu_\text{B} H,$$

уровень, отвечающий заданному J, расщепляется на $2J+1$ подуровней. Смещение уровней в единицах $\mu_\text{B}H$ определяется gm_J.

Например, жёлтый дублет натрия - переходы из $3^2P_{1/2}$ в $3^2S_{1/2}$-состояние (линия D_1), и из $3^2P_{3/2}$ - в то же $3^2S_{1/2}$-состояние (линия D_2).

Терм	L	S	J	g	m_J	gm_J
$3^2S_{1/2}$	0	1/2	1/2	2	$1/2; -1/2$	$1; -1$
$3^2P_{1/2}$	1	1/2	1/2	2/3	$1/2; -1/2$	$1/3; -1/3$
$3^2P_{3/2}$	1	1/2	3/2	4/3	$3/2; 1/2; -1/2; -3/2$	$2; 2/3; -2/3; -2$

(??? видимо, график нужно добавить, это же в квантмехе!!!)
(!!! убежусь, что в квантмехе это как минимум тоже есть!!!)

Системы единиц в магнетизме (!!!)

(допишу потом, это же в ЭД)

2.3 Typical Constant Magnetic Fields and Their Interactions with Matter

2.3.1 Magnetic field of typical direct currents and current systems

(ЛЛ8 пар 29, 30, 33, а также общефиз про это тут)

2.3.2 Typical magnetostatics in a nutshell

(весь магнетизм общефиза тут указан будет)

2.3.3 Thermodynamic Relations and the Free Energy of a Magnet

(ЛЛ8 пар 31, 32)

Термодинамические соотношения в магнитном поле

Полная свободная энергия магнетика

О термодинамике магнитных явлений по Боровику

(У ЛЛ8 скорее всего еще про это много хорошего написано, нужно будет - разберусь)

2.1. Общие закономерности

$$dW_H = \frac{1}{4\pi}(\mathbf{H}\cdot d\mathbf{B}) = \frac{1}{4\pi}\mathbf{H}d\mathbf{H} + \mathbf{H}d\mathbf{I} \qquad (\mathbf{B} = \mathbf{H} + 4\pi\mathbf{I})$$

$$dQ = dU + dA = dU + pdV - HdM$$

2.2. Идеальные магнетики

$$dS = \frac{dQ}{T} = \frac{dU + dA}{T} = \frac{dU + pdV}{T} - \frac{H}{T}dM$$

является полным дифференциалом. Ланжевен предположил, что существуют вещества, у которых не только dS, но и входящие в формулу (2.5) слагаемые

$$dS' = \frac{dU + pdV}{T} \quad \text{и} \quad dS'' = -\frac{H}{T}dM$$

являются полными дифференциалами. Тогда из определения полного дифференциала следует, что коэффициент при dM должен быть функцией только M, то есть

$$\frac{H}{T} = f(M), \quad \text{или} \quad M = f\left(\frac{H}{T}\right).$$

Иначе говоря, намагниченность однозначно определяется отношением магнитного поля к температуре. При малых полях $M \sim H; I = \varkappa H$ и, следовательно,

$$\varkappa = \frac{C}{T},$$

2.3. Магнето-термические и магнето-калорические соотношения

$$dU = TdS - pdV + HdM.$$

Аналогом величины pV является величина HM. Поэтому (???):

$$\Phi := U - TS + pV - HM; \qquad d\Phi := -SdT + Vdp - MdH.$$

Поскольку Φ - функция состояния, а $d\Phi$ - полный дифференциал, то

$$S = -\left(\frac{\partial\Phi}{\partial T}\right)_{pH}; \quad V = \left(\frac{\partial\Phi}{\partial p}\right)_{TH}; \quad M = -\left(\frac{\partial\Phi}{\partial H}\right)_{Tp}.$$

$$\left(\frac{\partial V}{\partial H}\right)_{pT} = -\left(\frac{\partial M}{\partial p}\right)_{TH}.$$

(добавлю те, которые используются!!!)

2.4. Теплоёмкость

$$C_{pM} = \lim_{\Delta T \to 0} \left(\frac{\Delta Q}{\Delta T} \right)_{pM} = T \left(\frac{\partial S}{\partial T} \right)_{pM},$$

$$C_{pH} = \lim_{\Delta T \to 0} \left(\frac{\Delta Q}{\Delta T} \right)_{pH} = T \left(\frac{\partial S}{\partial T} \right)_{pH}$$

Пусть независимыми являются переменные p, T, H, и $S = S[T, p, M(p, T, H)]$. Тогда

$$\left(\frac{\partial S}{\partial T} \right)_{Hp} = \left(\frac{\partial S}{\partial T} \right)_{pM} + \left(\frac{\partial S}{\partial M} \right)_{pT} \left(\frac{\partial M}{\partial T} \right)_{pH},$$

$\left(\frac{\partial S}{\partial M}\right)_{pT} = \left(\frac{\partial S}{\partial H}\right)_{pT} \left(\frac{\partial H}{\partial M}\right)_{pT}$, а согласно (2.12) $\left(\frac{\partial S}{\partial H}\right)_{pT} = \left(\frac{\partial M}{\partial T}\right)_{pH}$. Отсюда

$$C_{pH} - C_{pM} = T \left(\frac{\partial S}{\partial T} \right)_{Hp} - T \left(\frac{\partial S}{\partial T} \right)_{pM} = T \left(\frac{\partial M}{\partial T} \right)^2_{pH} \left(\frac{\partial H}{\partial M} \right)_{pT}.$$

2.5. Особенности термодинамического поведения некоторых магнетиков

Приращение внутренней энергии при намагничивании тела составляет

$$dU' = HdM \equiv Hd(-aH) \equiv -aHdH = MdH = -dA''.$$

Отсюда, если $dV = 0$ и $dU = dU'$, имеем

$$dQ = dU + pdV - HdM = MdH - HdM = 0,$$

Рассмотрим адиабатический процесс $\left(dS = \frac{dQ}{T} = 0\right)$. Будем считать, что независимыми являются переменные T, p и H, а $S = S(p, T, H)$. Поскольку dS - полный дифференциал, то

$$dS = \left(\frac{\partial S}{\partial T} \right)_{pH} dT + \left(\frac{\partial S}{\partial p} \right)_{HT} dp + \left(\frac{\partial S}{\partial H} \right)_{pT} dH = 0.$$

Используя (2.12) и (2.14), получаем

$$C_{pH} \frac{\partial T}{T} - \left(\frac{\partial V}{\partial T} \right)_{pH} dp + \left(\frac{\partial M}{\partial T} \right)_{pH} dH = 0.$$

Для процесса при постоянном давлении $(dp = 0)$

$$\left(\frac{\partial T}{\partial H} \right)_{Sp} = -\frac{T(\partial M/\partial T)_{pH}}{C_{pH}}.$$

Для классических диамагнетиков $\left(\frac{\partial M}{\partial T}\right)_{pH} = 0$ и $\left(\frac{\partial T}{\partial H}\right)_{Sp} = 0$. Для парамагнитных тел $\left(\frac{\partial M}{\partial T}\right)_{pH} < 0$ и $\left(\frac{\partial T}{\partial H}\right)_{Sp} > 0$, т. е. у парамагнетиков, намагничиваемых без притока тепла, должно наблюдаться повышение температуры, а при размагничивании - её понижение.

2.6. Вычисление магнитного момента тела

$$\delta\varepsilon_j = \frac{\partial\varepsilon_j}{\partial H}dH = -\mu_{jH}dH,$$

где μ_{jH} - составляющая магнитного момента атома в направлении поля в состоянии с квантовым числом j (j условно есть совокупность квантовых чисел, характеризующих состояние). Среднее значение магнитного момента атома в направлении поля

$$\bar{\mu}_H = \sum_j \mu_{jH}W_j = -\sum_j \frac{\partial\varepsilon_j}{\partial H}W_j,$$

$$-\sum_j \frac{\partial\varepsilon_j}{\partial H}W_j = \frac{\sum\limits_j \frac{\partial\varepsilon_j}{\partial H}q_j e^{-\varepsilon_j/(kT)}}{\sum\limits_j q_j e^{-\varepsilon_j/(kT)}} = kT\frac{\partial}{\partial H}(\ln Z).$$

Окончательно получаем

$$M_H = -\left(\frac{\partial\Phi}{\partial H}\right)_{pT} = N\bar{\mu}_H = NkT\frac{\partial}{\partial H}(\ln Z).$$

Пользуясь следующей из (2.25) связью между Φ и Z : $\Phi = -NkT\ln Z$, можно из (2.11) получить значение энтропии:

$$S = -\left(\frac{\partial\Phi}{\partial T}\right)_{pH} = Nk\frac{\partial}{\partial T}(T\ln Z).$$

2.3.4 Self-Induction, Forces, and Gyromagnetic Phenomena

(ЛЛ8 пар 34, 35,36)

2.4 Typical Spin Models and Methods for Them

(мб отсюда перекину вдаль что-то, а мб и в раздел выше.)

2.4.1 Exchange Interaction, Heisenberg Model, Fundamentals of Spin Waves

On Exchange interaction and Heisenberg model

Модель: два электронных облака соседних атомов взаимодействуют кулоновским образом с энергией, зависящей от их спинового состояния.

Рассмотрим для примера обменное взаимодействие двух электронов 1 и 2 с суммарным спином $\hat{\mathbf{S}} = \hat{\mathbf{S}}_1 + \hat{\mathbf{S}}_2$. Поскольку в базисе собственных состояний $S(S+1) = S_1(S_1+1) + S_2(S_2+1) + 2\hat{\mathbf{S}}_1\hat{\mathbf{S}}_2$, а $S_1 = S_2 = 1/2$, то обменную разницу энергий синглетного ($S = 0$) и триплетного ($S = 1$) состояний $\approx J_{12}(r)$ удобно описывать слагаемым с обменным интегралом $-J_{12}\hat{\mathbf{S}}_1\hat{\mathbf{S}}_2$. Для системы взаимодействующих атомарных спинов ($g = 2$):

$$\hat{H} = -\mu_B g\sum_i \hat{\mathbf{S}}_i\cdot\boldsymbol{B} - \frac{1}{2}\sum_{i\neq j} J_{ij}\hat{\mathbf{S}}_i\hat{\mathbf{S}}_j \qquad \text{(гамильтониан Гейзенберга)}$$

В приближении самосогласованного среднего поля будем считать, что в образце существует единая для всей решётки средняя намагниченность $\boldsymbol{M} = \mu_B g n\langle\boldsymbol{S}\rangle$, а

отклонения спинов от среднего значения малы $\hat{\mathbf{S}}_i = \langle \mathbf{S} \rangle + \delta\hat{\mathbf{S}}_i, \delta S_i \ll \langle S \rangle$. (??? додумаю, почему такая подстановка сделана???) Тогда

$$\hat{H} = -\mu_B g \sum_i \hat{\mathbf{S}}_i \left(\boldsymbol{B} + \frac{1}{\mu_B g} \sum_{j\neq i}{}' J_{ij} \langle \mathbf{S} \rangle \right).$$

(!?!??!??! абзац про то, что же такое это обменное взаимодействие???? в чем его природа??? что с волновыми там функциями происходит??)

$$\mathcal{H} = \pm J \sum_{\langle \mathbf{r}\mathbf{r}' \rangle} \mathbf{S}_\mathbf{r} \cdot \mathbf{S}_{\mathbf{r}'} \equiv \pm J \sum_{\langle \mathbf{r}\mathbf{r}' \rangle} \left[S^z_\mathbf{r} S^z_{\mathbf{r}'} + \frac{1}{2} \left(S^+_\mathbf{r} S^-_{\mathbf{r}'} + S^-_\mathbf{r} S^+_{\mathbf{r}'} \right) \right] \quad \text{(Heisenberg Ham-an)}$$

(абзац почему вообще магнетизм связан с именно спинами??? в квантмехе это подробнее в 1й части тоже. типичный вопрос.)

$$\hat{S}^\pm_i = \hat{S}^x_i \pm i \cdot \hat{S}^y_i, \left[\hat{S}^z_i, \hat{S}^\pm_j \right] = \pm \delta_{ij} \hat{S}^\pm_i, \left[\hat{S}^+_i, \hat{S}^-_j \right] = 2\delta_{ij} \hat{S}^z_i.$$

$$\hat{H} = -\mu_B g \sum_i \mathcal{B} \cdot \hat{S}^z_i - \frac{1}{2} \sum_{i\neq j} J_{ij} \left(\hat{S}^z_i \hat{S}^z_j + \hat{S}^+_i \hat{S}^-_j \right).$$

$$i\hbar \frac{\partial \hat{S}^+_i}{\partial t} = \left[\hat{S}^-_i, \hat{H} \right]$$

$$i\hbar \frac{\partial \hat{S}^-_i}{\partial t} = \mu_B g \mathcal{B} \cdot \hat{S}^-_i + \sum_{j\neq i}{}' J_{ij} \left(\hat{S}^z_j \hat{S}^-_i - \hat{S}^-_j \hat{S}^z_i \right).$$

$$i\hbar \frac{\partial \hat{S}^-_i}{\partial t} = \mu_B g \mathcal{B} \cdot \hat{S}^-_i + \sum_{j\neq i}{}' \langle S \rangle J_{ij} \left(\hat{S}^-_i - \hat{S}^-_j \right).$$

Как решить такую систему? Ситуация здесь напоминает уравнения движения атомов в кристалле. Это дифференциальные уравнения с постоянными коэффициентами, распределённые в пространстве. Ситуация типичная для волнового уравнения, поэтому естественно поискать его решение в виде бегущей волны:

$$\hat{S}^-_i = \hat{S}^-_\mathbf{k} e^{-i\omega_\mathbf{k} t + \mathbf{i}\mathbf{k}\mathbf{k}_i}.$$

$$\hbar\omega_\mathbf{k} = \mu_B g \mathcal{B} + \langle S \rangle \sum_{j\neq i}{}' J_{ij} \left(1 - e^{i\mathbf{k}(\mathbf{r}_i - \mathbf{r}_j)} \right).$$

Считая, как и в (9.11), что кристалл имеет простую кубическую решётку и каждый спин взаимодействует одинаково $J_{ij} = J$ с $z = 6$ ближайшими соседями на расстоянии a, в длинноволновом пределе $ka \ll 1$ получаем

$$\hbar\omega_\mathbf{k} = \mu_B g \mathcal{B} + J\langle S \rangle a^2 k^2.$$

Таким образом, квант спиновых волн, который естественно назвать магноном, в отсутствие внешнего поля имеет такой же спектр как частица с массой $m*$:

$$\hbar\omega_\mathbf{k} = \frac{\hbar^2 k^2}{2m^*},$$
$$m^* \approx \frac{\hbar^2 S(S+1)}{\langle S \rangle a^2 T_c} \approx \frac{\hbar^2}{a^2 T_c}.$$

Если перейти к атомным единицам, которые связаны соотношением $\mathrm{Ry} \approx \hbar^2/m_e a_B^2$, то для отношения массы магнона к массе электрона получаем $m^*/m_e \approx (a_B/a)^2 \,\mathrm{Ry}/T_c$. При типичных значениях параметров ферромагнетика, масса магнона $m*$ раз в десять превышает массу электрона m_e.

Магноны и преобразование Холстейна-Примакова

При низких температурах возбужденные состояния системы близки к основному. При этом ферромагнетик содержит малое число магнонов, и можно считать, что их числа заполнения малы: $\langle n_j \rangle \ll S$. Это позволяет предложить приближенную процедуру квантования спиновых волн. Для этого нужно выразить операторы спина $\hat{\boldsymbol{S}}$ через операторы вторичного квантования для неких гармонических осцилляторов $\hat{a}_i^+$и $\hat{a}_i$, удовлетворяющих коммуникационным соотношениям $\left[a_i, a_i^+\right] = 1$. Мы приведем приближенные выражения для спиновых операторов через операторы $\hat{a}_i^+$и $\hat{a}_i$, в которых опущены слагаемые, содержащие отрицательные степени величины S. Это преобразование Холстейна-Примакова:

$$\hat{S}_i^+ = \sqrt{2S}\hat{a}_i, \hat{S}_i^- = \sqrt{2S}\hat{a}_i^+, \hat{S}_i^z = S - \hat{a}_i^+\hat{a}_i.$$

$$\hat{H} = -\frac{1}{2}\sum_{i\neq j} J_{ij}\left[S^2 + S\left(\hat{a}_i\hat{a}_j^+ + \hat{a}_i^+\hat{a}_j - \hat{a}_i^+\hat{a}_i - \hat{a}_j^+\hat{a}_j\right)\right] - \mu_B g\mathcal{B}\sum_i\left(S - \hat{a}_i^+\hat{a}_i\right).$$

$$\hat{a}_i = N^{-1/2}\sum_{\mathbf{k}} e^{i\mathbf{k}\mathbf{r}_i}\hat{a}_{\mathbf{k}}, \quad \hat{a}_i^+ = N^{-1/2}\sum_{\mathbf{k}} e^{-i\mathbf{k}\mathbf{r}_i}\hat{a}_{\mathbf{k}}^+,$$

$$J(\mathbf{k}) = \sum_{i\neq j} J_{ij}e^{-i\mathbf{k}(\mathbf{r}_i - \mathbf{r}_j)},$$

$$\sum_{i\neq j} e^{i(\mathbf{k}-\mathbf{k}')(\mathbf{r}_i - \mathbf{r}_j)} = N\delta_{\mathbf{k},\mathbf{k}'},$$

$$\sum_{i\neq j} J_{ij}\left(1 - \cos\mathbf{k}\mathbf{r}_{ij}\right) = J(0) - J(\mathbf{k}).$$

В итоге получаем

$$\hat{H} = E_0 + \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}}\hat{a}_{\mathbf{k}}^+\hat{a}_{\mathbf{k}},$$

где

$$E_0 = -NS^2J(0)/2 - \mu_B SgBN - \frac{S}{2}\sum_{\mathbf{k}} J(\mathbf{k})$$

- энергия основного состояния кристалла при $T = 0$,

$$\hbar\omega_{\mathbf{k}} = S[J(0) - J(\mathbf{k})] + \mu_B g\mathcal{B}$$

Полная намагниченность ферромагнетика равна

$$\mathcal{M}(T) = \frac{\mu_B}{V}\sum_i\left\langle\hat{S}_i^z\right\rangle = \frac{\mu_B g}{V}\left(NS - \sum_{\mathbf{k}}\left\langle\hat{a}_{\mathbf{k}}^+\hat{a}_{\mathbf{k}}\right\rangle\right) = \mathcal{M}(0) - \mu_B gV\int\frac{d^3\mathbf{k}}{(2\pi)^3}\left\langle\hat{a}_{\mathbf{k}}^+\hat{\mathbf{a}}_{\mathbf{k}}\right\rangle,$$

где $\mathcal{M}(0)$ - равновесная намагниченность при $T = 0$, а равновесное число магнонов в данной моде $\mathbf{k}$: $\left\langle\hat{a}_{\mathbf{k}}^+\hat{a}_{\mathbf{k}}\right\rangle = \left(\exp\left(\hbar\omega_{\mathbf{k}}/T\right) - 1\right)^{-1}$. При $T \ll T_c$ основной вклад в интеграл (9.28), который берется по первой зоне Бриллюэна, дают $k \leq a^{-1}\sqrt{T/JS} \ll k_B$, для которых $\hbar\omega_{\mathbf{k}} \leq T$. Поэтому интегрирование по k можно распространить до бесконечности:

$$\mathcal{M}(T) = \mathcal{M}(0) - \frac{\mu_B gV}{\pi^2}\int\limits_0^\infty\frac{k^2dk}{e^{\frac{\hbar^2k^2}{2m_sT}} - 1} = \mathcal{M}(0) - \frac{\mu_B gV\zeta(3/2)}{2\pi^{3/2}}\left(\frac{2m_sT}{\hbar^2}\right)^{3/2}.$$

Таким образом, при $T \ll T_c$ величина $\mathcal{M}(0) - \mathcal{M}(T) \propto T^{3/2}$ подчиняется закону Блоха.

Теплоёмкость и намагниченность магнонов

Обсудим физический смысл полученных в предыдущем пункте результатов. Итак, в низкотемпературном пределе теории среднего поля имеется магнонная ветвь возбуждений бозевского типа. Она даёт вклад в теплоёмкость и намагниченность ферромагнетика. Вклад магнонов в энергию кристалла равен

$$E = \sum_{\mathbf{k}} \frac{\hbar\omega_{\mathbf{k}}}{e^{\frac{\hbar\omega_{\mathbf{k}}}{T}} - 1} = \frac{V}{(2\pi\hbar)^3} \int\limits_0^\infty \frac{\hbar^2 k^2}{2m^*} \cdot \frac{k^2 dk}{e^{\frac{\hbar^2 k^2}{2m^* T}} - 1} \propto T^{5/2}.$$

Следовательно, теплоёмкость магнонов равна

$$C_V \propto T^{3/2}.$$

можно обосновать и чисто формально. Состояние $N^{-1/2} \sum\limits_i \hat{S}_i^+ |0\rangle$, соответствующее одному магнону (9.17), является собственным вектором оператора полного спина кристалла $\sum\limits_i \hat{S}_i^z$ с собственным значением -1. Число магнонов равно

$$N(T) = \sum_{\mathbf{k}} \frac{1}{e^{\frac{\hbar\omega_{\mathbf{k}}}{T}} - 1} = \frac{V}{(2\pi\hbar)^3} \int\limits_0^\infty \frac{4\pi k^2 dk}{e^{\frac{\hbar^2 k^2}{2m^* T}} - 1} \propto T^{3/2}.$$

Каждый магнон несет направленный против намагниченности кристалла $\mathcal{M}(0)$ магнитный момент $-g\mu_B$. Это значит, что $N(T)$ магнонов, возникших при температуре T, уменьшают намагниченность ферромагнетика на величину

$$\mathcal{M}(0) - \mathcal{M}(T) \propto T^{3/2}.$$

Это соотношение представляет собой закон трёх вторых Блоха

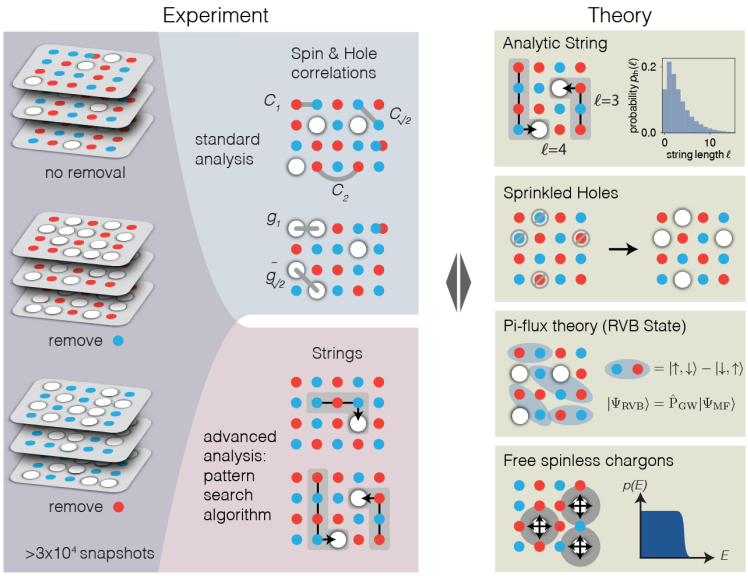
2.4.2 1D Ising Model (!)

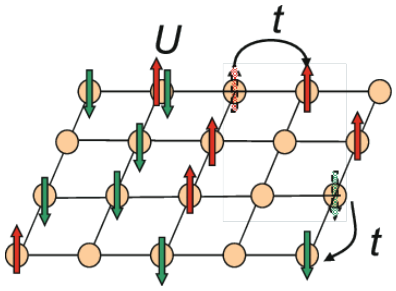
(всё о ней. в статфизе похожее, но с т зр статфиза)

2.4.3 2D Ising Model (!)

(всё о ней. в статфизе похожее, но с т зр статфиза)

2.4.4 Hubbard, SSH, tight-binding models





On Hubbard model

$$\hat{H}_{\text{hb}} := -t \sum_{i,\sigma} \left(\hat{c}^{\dagger}_{i,\sigma} \hat{c}_{i+1,\sigma} + \hat{c}^{\dagger}_{i+1,\sigma} \hat{c}_{i,\sigma} \right) + U \sum_{i} \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}; \qquad \overset{\hat{n}_i = \hat{n}_{i\uparrow} + \hat{n}_{i\downarrow}}{\hat{n}_{i\sigma} = \hat{c}^{\dagger}_{i\sigma} \hat{c}_{i\sigma}};$$

The Hubbard model states that each electron experiences competing forces: one pushes it to tunnel to neighboring atoms, while the other pushes it away from its neighbors (?? и что за вторая?). Its Hamiltonian thus has two terms: a kinetic term allowing for tunneling ("hopping") of particles between lattice sites and a potential term reflecting on-site interaction.

(!??!?! выпишу из конд сред про него!?!?!)

On SSH model

On tight-binding models

Using the atomic orbital as a basis state, the second quantization Hamiltonian operator in the tight binding framework can be written as:

$$H = -t \sum_{\langle i,j \rangle, \sigma} \left(c^{\dagger}_{i,\sigma} c_{j,\sigma} + \text{ h.c. } \right)$$

(я смотрел про это и вывод этого в кандидатской Кирилла, но не понял, как это с магнетизмом связано и вообще суть, как это работает. потом допишу скорее всего.)

2.4.5 Spin-Wave Theory and Magnons (!!!!)

(скоро выпишу, типичная тема вуза, уже продумывал частично это!)
Магноны - кванты спиновых волн, бозоны.

21. Динамика магнитной решётки. Спиновые волны

§

21.1. Общие представления о спиновых волнах

§

21.2. Полуклассическая теория спиновых волн в ферромагнетике

§

21.3. Спиновые волны в антиферромагнетике

§

21.4. Спектр спиновых волн в области малых квазиимпульсов

§

21.5. Экспериментальная проверка теории спиновых волн. Термодинамические свойства антиферромагнетиков

§

21.6. Неупругое рассеяние нейтронов в магнитоупорядоченных кристаллах

2.4.6 Jordan-Wigner, Holstein-Primakoff Transformation and Similar (??)

(важный раздел с мат указаниями!!! потом допишу, пока просто не нужны они были.)
(абзац про физические обоснования модели!!)

2.5 Typical Effects of Magnetism

2.5.1 de Haas-von Alphen effect in a Nutshell

(!! у Ароваса хорошо написано, вот я и впишу это!!!)

de Haas-von Alphen effect for anisotropic Fermi surfaces

2.5.2 Superparamagnetism, Magnetocrystalline Anisotropy

(выпишу из лекций, там 1 лекция про это, и хватит, просто пара интересных тем. лень про них как-то думать)

Superparamagnetism

(!!! пропишу это очень скоро уже!!!)

Origin of magnetocrystalline anisotropy

(!!! разберусь уже с тем, что такое easy axis, hard axis наконец-то!!!)

Symmetry of magnetocrystalline anisotropy

$$E = K_1\left(\alpha_1^2\alpha_2^2 + \alpha_2^2\alpha_3^2 + \alpha_3^2\alpha_1^2\right) + K_2\left(\alpha_1^2\alpha_2^2\alpha_3^2\right) + \cdots .$$

Here K_1, K_2, etc. are called the anisotropy constants. Typical values for iron at room temperature are $K_1 = 4.2 \times 10^5 \text{erg/cm}^3$ and $K_2 = 1.5 \times 10^5 \text{erg/cm}^3$. The energy, E, is that stored in the crystal when work is done against the anisotropy "force" to move the magnetization away from an easy direction. Note that the anisotropy energy is an even function of the direction cosines, and is invariant under interchange of the α_i s among themselves.

Cobalt is hexagonal, with the easy axis along the hexagonal (c) axis. The symmetry causes a uniaxial anisotropy energy, and its angular dependence is a function only of the angle θ between the magnetization vector and the hexagonal axis (see Fig. 11.4). In this case the anisotropy energy can be expanded as

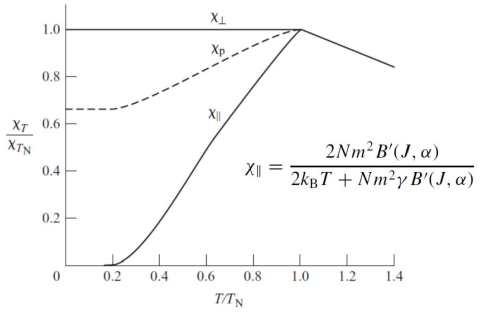
$$E = K_1 \sin^2\theta + K_2 \sin^4\theta + \cdots .$$

Typical values of the anisotropy constants for cobalt at room temperature are $K_1 = 4.1 \times 10^6 \text{erg/cm}^3$ and $K_2 = 1.0 \times 10^6 \text{erg/cm}^3$. Note that, in all materials,

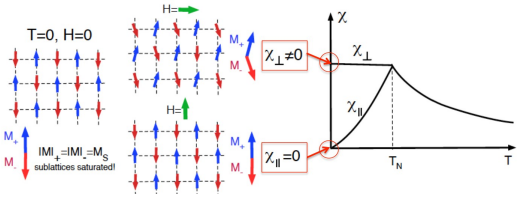
Shape anisotropy

$$\boldsymbol{H}_\text{d} = N_\text{d}\boldsymbol{M}$$
$$\boldsymbol{H}_\text{eff} = \boldsymbol{H}_\text{applied} - \boldsymbol{H}_\text{d}$$

N_d is smallest along long axis the $\boldsymbol{H}_\text{applied}$ required for saturation is smallest along long axis



(??? see theory for it!)



Energy of magnetic anisotropy from Landau

$$U_\text{a} = K\left(m_x^2 + m_y^2\right) = K\sin^2\theta \qquad \Leftrightarrow \qquad U_\text{a} = -Km_z^2 = -K\cos^2\theta,$$

$$\left(m_x^2 + m_y^2\right)^2 = \sin^4\theta$$

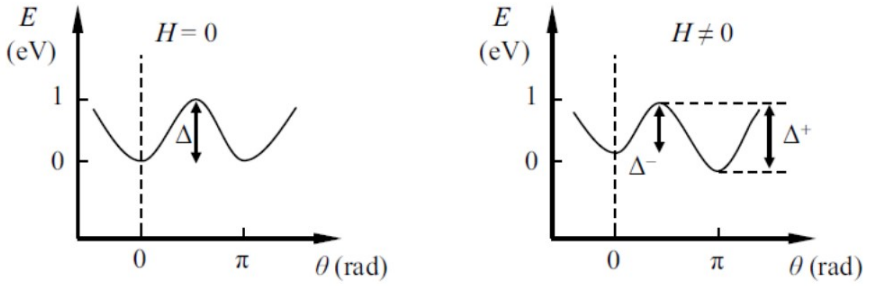
$$\frac{1}{2}m_z\left[\left(m_x + im_y\right)^3 + \left(m_x - im_y\right)^3\right] = \cos\theta\sin^3\theta\cos 3\varphi$$

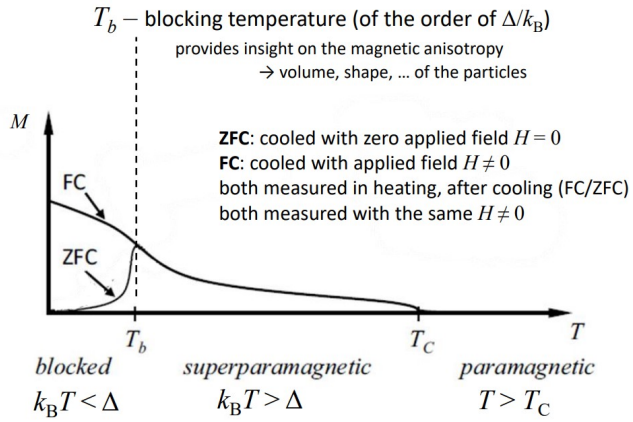
$$U_{\text{a}} = K\left(m_x^2 m_y^2 + m_x^2 m_z^2 + m_y^2 m_z^2\right) \qquad \Leftrightarrow \qquad U_{\text{aB}} = -\frac{1}{2}K\left(m_x^4 + m_y^4 + m_z^4\right);$$

$$AM_z^2 + \left(A + \frac{K}{M^2}\right)\left(M_x^2 + M_y^2\right)$$

Induced magnetic anisotropy

Superparamagnetism (??!!!)





(тут комментарии про смысл происходящего, которые я слышал на лекции!!!)

2.5.3 Energy of the Ferromagnetic State

(?????)

10.1. Характеристики технической кривой намагничивания127

10.2. Энергия естественной, или кристаллографической, магнитной анизотропии130

10.3. Энергия магнитострикционной деформации134

10.4. Экспериментальное определение констант магнитной анизотропии и их зависимость от температуры и состава ферромагнетика141

10.5. Энергия упруго деформированного ферромагнетика147

10.6. Энергия взаимодействия ферромагнетика с магнитным полем

2.5.4 Typical magnetoresistance

(см. пока кинетику, не особо этим занимался, но на 4 курсе задача про это была)

2.5.5 Magnetic Properties of Atomic Nuclei

§
16.1. Обнаружение магнитных свойств ядер

§
16.2. Энергия взаимодействия ядра с электронной оболочкой233

§
16.3. Эффект Зеемана для сверхтонкой структуры238

§
16.4. Правила перехода242

§
16.5. Методы определения величины магнитных моментов ядер....243

§
16.6. Измерение магнитного момента нейтронов246

§
16.7. Основные результаты измерения магнитных моментов ядер....249

§
16.8. Магнитные моменты нечётно-чётных ядер251

Part II
Special Magnetism in a Nutshell

3 Famous special properties and models of magnets

3.1 Ferromagnetics, anitferromagnetics special phenomena and properties

3.1.1 The Domain Structure of Ferromagnets

(см Боровик, ну и у Ландау тоже про это есть, доудмаю мб когда-то)

§
12.1. Причина образования доменов159

§
12.2. Переходные слои между доменами160

§
12.3. Доменная структура в одноосном ферромагнитном кристалле..163

§
12.4. Экспериментальное изучение ферромагнитных областей168

§
12.5. Доменная структура в магнитно-многоосном кристалле173

§
12.6. Тонкая структура доменной границы177

§
12.7. Цилиндрический магнитный домен180

§
12.8. Однодоменная структура184

§
12.9. Сверхпарамагнетизм186

3.1.2 Interaction of Magnons

(типичный статфиз и кинетика физтеха, потом напишу)

3.1.3 Phenomena with Ferromagnetism

(разделю потом по типам!!!)

51 Пьезомагнетизм и магнитоэлектрический эффект

$$\widetilde{\Phi}_{\text{пм}} = -\lambda_{i,kl} H_i \sigma_{kl},$$

где $\lambda_{i,kl}$ - тензор, симметричный по индексам kl (ср. (17.7)). Этот член приводит к появлению в индукции $B_i = -4\pi\partial\widetilde{\Phi}/\partial H_i$ дополнительного члена $4\pi\lambda_{i,kl}\sigma_{kl}$. Другими словами, при $\mathbf{H} = 0$ возникает линейная по деформации намагниченность:

$$M_i = \lambda_{i,kl}\sigma_{kl}.$$

Другим проявлением того же свойства является линейная магнитострикиия - возникновение деформации, линейной по наложенному на кристалл полю:

$$u_{kl} = -\frac{\partial\widetilde{\Phi}}{\partial\sigma_{kl}} = \lambda_{i,kl}H_i$$

(словами напишу, там еще на страницу умных мыслей у Ландау)

52 Геликоидальная магнитная структура

Особую категорию составляют магнитные структуры, в которых периоды «магнитной решетки» несоизмеримы с периодами основной, кристаллографической решетки. Возможны различные механизмы возникновения таких структур; мы рассмотрим здесь один из них (указанный И.Е. Дзялошинским, 1964), допускающий простую формулировку в макроскопических терминах. Сделаем это на конкретном примере: рассмотрим кристалл, относящийся к кубическому классу $\boldsymbol{T}$, причем обменное взаимодействие в нем само по себе устанавливало бы чисто ферромагнитное упорядочение магнитных моментов (В.Г. Баръяхтар, Е.П. Стефановский, 1969; P. Bak, M.H. Jensen, 1980)[1]).

$$U_{\text{неодн}} = \gamma\mathbf{M}\,\mathrm{rot}\,\mathbf{M} + \frac{\alpha}{2}\left[(\nabla M_x)^2 + (\nabla M_y)^2 + (\nabla M_z)^2\right].$$

$$\mathbf{M}(\mathbf{r}) = \frac{M}{\sqrt{2}}\left(\mathbf{m}e^{i\mathbf{kr}} + \mathbf{m}^* e^{-i\mathbf{kr}}\right)$$

$\mathbf{m}^2 = 0.$
$\mathbf{m} = (\mathbf{m}_1 + i\mathbf{m}_2)/\sqrt{2}$, $\mathbf{m}_1 \perp \mathbf{m}_2$

$$\mathbf{M}(\mathbf{r}) = M\left(\mathbf{m}_1\cos\mathbf{kr} - \mathbf{m}_2\sin\mathbf{kr}\right).$$

$$U_{\text{неодн}} = iM^2\gamma\mathbf{k}\left[\mathbf{m}\mathbf{m}^*\right] + \frac{1}{2}\alpha M^2k^2 = M^2\gamma\mathbf{k}\left[\mathbf{m}_1\mathbf{m}_2\right] + \frac{1}{2}\alpha M^2k^2.$$

$$k = \frac{\gamma}{\alpha} \ll \frac{1}{a}.$$

42 Магнитострикция ферромагнетиков (!!!)

$$U_{\text{M.y}} = -a_{iklm}\sigma_{ik}m_l m_m,$$

$$U_{\text{M.y}} = -a_1\left(\sigma_{xx}m_x^2 + \sigma_{yy}m_y^2 + \sigma_{zz}m_z^2\right) - 2a_2\left(\sigma_{xy}m_x m_y + \sigma_{xz}m_x m_z + \sigma_{yz}m_y m_z\right).$$

Тензор деформации:

$$u_{ik} = -\frac{\partial\widetilde{\Phi}}{\partial\sigma_{ik}},$$

причем в $\widetilde{\Phi}$ надо включить (с обратным знаком) также и обычную упругую энергию. У кубического кристалла последняя содержит три независимых упругих коэффициента и может быть представлена, например, в виде

$$U_{\text{упр}} = \frac{\mu_1}{2}\left(\sigma_{xx}^2 + \sigma_{yy}^2 + \sigma_{zz}^2\right) + \frac{\mu_2}{2}\left(\sigma_{xx} + \sigma_{yy} + \sigma_{zz}\right)^2 + \mu_3\left(\sigma_{xy}^2 + \sigma_{xz}^2 + \sigma_{yz}^2\right),$$

где μ_1, μ_2, μ_3 - положительные.

$$\widetilde{\Phi} = -\mathscr{M}\mathfrak{H} = -MV\mathfrak{H}$$

где $\mathscr{M} = MV$ - полный магнитный момент тела, однородно намагниченного в направлении, совпадающем с направлением поля;

$$\bar{u}_{ik} = -\frac{1}{V}\frac{\partial\Phi}{\partial\sigma_{ik}},$$

$$\bar{u}_{ik} = \frac{\mathfrak{H}}{V}\frac{\partial(VM)}{\partial\sigma_{ik}}.$$

(!!! нарешаю задачи, тут крутые задачи!!!)

11. Магнитострикция и эксперименты по ней по Боровику

§

11.1. Магнитострикция поликристаллических ферромагнетиков

§

11.2. Экспериментальные методы измерения магнитострикции151

§

11.3. Упругие напряжения и магнитострикция155

46 Ориентационные переходы

47 Флуктуации в ферромагнетике

3.1.4 Properties of Ferromagnetic Domains

43 Поверхностное натяжение доменной стенки

$$a_i(M)M_i\frac{\partial M_l}{\partial x_i} = \frac{1}{2}\mathbf{a}(M)\nabla M^2$$

$$U_{\text{неодн}} = \frac{1}{2}\alpha_{ik}\frac{\partial M_l}{\partial x_i}\frac{\partial M_l}{\partial x_k},$$

где α_{ik} - симметричный тензор;

$$U_{\text{неодн}} = \frac{\alpha}{2}\frac{\partial \mathbf{M}}{\partial x_i}\frac{\partial \mathbf{M}}{\partial x_i}.$$

В одноосном кристалле тензор α_{ik} имеет две независимые компоненты и энергия неоднородности имеет вид

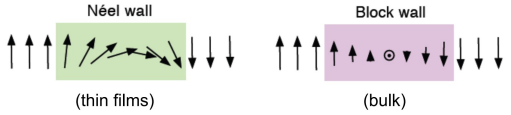
$$U_{\text{неодн}} = \frac{\alpha_1}{2}\left[\left(\frac{\partial \mathbf{M}}{\partial x}\right)^2 + \left(\frac{\partial \mathbf{M}}{\partial y}\right)^2\right] + \frac{\alpha_2}{2}\left(\frac{\partial \mathbf{M}}{\partial z}\right)^2.$$

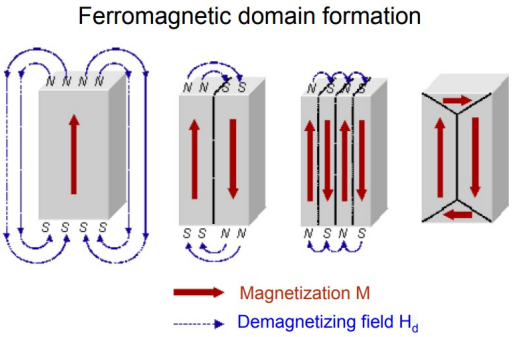
$$\int\limits_{-\infty}^{\infty}\left[\frac{\alpha_1}{2}\left(M_y'^2 + M_z'^2\right) + \frac{\beta}{2}M_y^2\right]dx = \frac{M^2}{2}\int\limits_{-\infty}^{\infty}\left(\alpha_1\theta'^2 + \beta\sin^2\theta\right)dx =$$

$$M^2\alpha_1\int\limits_{-\infty}^{\infty}\theta'^2 dx = M^2\sqrt{\alpha_1\beta}\int\limits_{0}^{\pi}\sin\theta d\theta = 2M^2\sqrt{\alpha_1\beta}.$$

Если рассматривать границу раздела между доменами как геометрическую поверхность, то эта величина есть поверхностное натяжение, которое должно быть приписано границе для учета энергии, необходимой для ее образования. Обозначим поверхностное натяжение доменной стенки как $M^2\Delta$, где Δ имеет размерность длины. Тогда

$$\Delta = 2\sqrt{\alpha_1\beta}.$$





44 Доменная структура ферромагнетиков

45 Однодоменные частицы

3.1.5 Antiferromagnets Under Special Conditions (?)

48 Антиферромагнетик вблизи точки Кюри

49 Бикритическая точка антиферромагнетика

50 Слабый ферромагнетизм

Существуют кристаллы, в которых обменное взаимодействие устанавливает антиферромагнитную структуру, но сравнительно слабые релятивистские взаимодействия приводят к небольшому искажению этой структуры, в результате чего появляется намагничение $\mathbf{M}$ - «аномально» малое в меру малости релятивистских взаимодействий по сравнению с обменными. Это явление называют слабым ферромагнетизмом.

(дальше умные рассуждения про кристаллы)

Члены разложения должны быть инвариантны по отношению ко всем преобразованиям группы $\boldsymbol{D}^6_{3d}$. Первые члены такого разложения:

$$\Phi = \Phi_0(L) + BM^2 + D(\mathbf{lM})^2 - \frac{\beta}{2}M_z^2 - \frac{\gamma L^2}{2}l_z^2 + \zeta L\left(M_x l_y - M_y l_x\right),$$

где $\Phi_0(L)$ - изотропная по $\mathbf{L}$ функция. Первые два (после Φ_0)

Будем считать, что постоянная $\gamma < 0$; тогда вектор 1 устанавливается в базисной плоскости ($l_z = 0$). Выбрав в качестве плоскости xz ту, в которой лежит l, и минимизируя Φ по $\mathbf{M}$ при заданном $\mathbf{L}$, найдем для ферромагнитного момента:

$$M_x = 0, \quad M_y = \frac{\zeta}{2B}L_x, \quad M_z = 0.$$

$$M_z = \chi_\parallel H_z, \quad M_\perp = \chi_\perp\left(H_\text{Д} + H_\perp\right),$$

где восприимчивости

$$\chi_\parallel = \frac{1}{2B-\beta}, \quad \chi_\perp = \frac{1}{2B}$$

и введено обозначение

$$H_\text{д} = |\zeta|L$$

для «эффективного поля», определяющего спонтанную намагниченность слабого ферромагнетика (его называют полем Дзялошинского). Поскольку $|\beta| \ll B$, то $\chi_\parallel \approx \chi_\perp$.

Точка Кюри определяется обращением в нуль коэффициента при L^2, так что вблизи нее

$$A - \frac{\zeta^2}{4B} = a\left(T - T_c\right)$$

($a > 0$): остальные коэффициенты полагаем равными их значениям при $T = T_c$ (при этом $C > 0$).

При наличии поля уравнения $\partial\widetilde{\Phi}/\partial L = 0, \partial\widetilde{\Phi}/\partial M = 0$ после исключения из них M дают следующее уравнение, определяющее L:

$$2CL^3 + a\left(T - T_c\right)L - \frac{\zeta H}{4B} = 0.$$

Уже отсюда видно, что в слабом ферромагнетике (как и в обычном) магнитное поле размывает фазовый переход). При этом оказывается отличным от нуля по обе стороны

от точки $T = T_c$ и антиферромагнитный вектор; магнитное поле вызывает в парамагнитной фазе антиферромагнитное упорядочение, устраняя тем самым различие в свойствах симметрии обеих фаз (А.С. Боровик-Романов, В.И. Ожогин, 1960). При $T > T_c$ на некотором расстоянии от этой точки L убывает по закону

$$L = \frac{\zeta H}{4aB\,(T - T_c)}.$$

3.1.6 Dynamics of Magnetic Domains

(see Borovik par. 26)

§

26.1. Предельная скорость доменной границы468

§

26.2. Вынужденное движение доменных границ474

§

26.3. Излучение звука при движении доменной границы479

§

26.4. Динамика цилиндрического магнитного домена481

3.1.7 Ferromagnets in Unsteady Magnetic Fields

§

15.1. Магнитная вязкость (магнитное последействие)215

§

15.2. Ферромагнетики в переменных магнитных полях217

§

15.3. Определение комплексной магнитной проницаемости и тангенса угла потерь222 Оглавление

§

15.4. Дисперсия магнитной проницаемости225

3.1.8 Ferromagnetic Resonance

(см. Боровик 18)

§

18.1. Введение294

§

18.2. Влияние формы образца на резонансную частоту295

§

18.3. Влияние магнитной анизотропии на резонансную частоту299

§

18.4. Влияние доменной структуры на резонансную частоту301

§

18.5. Фактор Ланде302

§

18.6. Ферромагнитный резонанс ферримагнетиков304

§

18.7. Ширина линий ферромагнитного резонансного поглощения....306

3.1.9 Nuclear Gamma Resonance (The Mossbauer Effect)

(см. Боровик 19)

§

19.1. Резонансная флюоресценция гамма-квантов ядрами свободных атомов. Аналогия с оптической флюоресценцией

§

19.2. Флюоресценция ядер в кристалле — эффект Мёссбауэра

§

19.3. Разрешающая способность, значение эффекта Мёссбауэра. Схема опыта

§

19.4. Применение эффекта Мёссбауэра к проблемам физики твёрдого тела

§

19.5. Экспериментальные методы исследования эффекта Мёссбауэра

§

19.6. Результаты экспериментальных исследований магнитоупорядоченных кристаллов методом ядерного гамма-резонанса

3.1.10 Stationary Magnetic Structures and the Influence of an External Field on Them

(Боровик 20.)

§

20.1. Нейтронографические исследования магнитоупорядоченных кристаллов и типы их магнитных структур

§

20.2. Магнитная симметрия

§

20.3. Теоретические методы отыскания магнитных структур

§

20.4. Влияние внешнего магнитного поля на магнитную структуру кристалла

§

20.5. Экспериментальное исследование опрокидывания магнитных подрешёток

3.2 Magnetism in Special Media

(все это пока не актуально вообще, нужно будет - каждую тему и доучу)

3.2.1 Special Models of Magnetism of Free Atoms

Van Vleck Temperature-Independent Paramagnetism

Cooling by Isentropic Demagnetization

Nuclear Demagnetization

3.2.2 Magnetism of Molecules and Chemical Compounds (????)

5.2. Свойства химических соединений

У химического соединения, состоящего из атомов сорта i с атомной восприимчивостью $\varkappa_{ai}$ молярная восприимчивость $\varkappa_м$ равна:

$$\varkappa_м = \sum_i n_i \varkappa_{ai} + \lambda, \qquad \text{(правило Паскаля)}$$

λ постоянная, зависящая от природы связей между атомами (???). Величины $\varkappa_{ai}$, носящие название постоянных Паскаля, не совпадают со значением атомных восприимчивостей простых веществ, но постоянны для широкого класса веществ. (?? и как найти их???)

Правило Паскаля основано на том, что при образовании химических соединений внутренние оболочки атома не изменяются, и на том, что данному характеру связи отвечает более или менее неизменная электронная конфигурация.

(не знаю тут ничего, нужно будет - изучу)

3.2.3 Magnetic Impurities

(???? see cond matter)

3.2.4 Special Statistical, Gas Models of Magnetism (???)

(закоулки статфиза, пока не актуально)

О других моделях магнетизма

О магнетизме вырожденного бозе-газа

(это бесполезная теория, так что чуть потом ее додумаю.)

Найдем парамагнитную и диамагнитную восприимчивость идеального нерелятивистского двухмерного газа заряженных бозе-частиц со спином $s = 1$ при низких температурах в слабом перпендикулярном магнитном поле $\mathcal{B}$. Химический потенциал без поля найдется из

$$N = \frac{g_s 2\pi m S}{(2\pi\hbar)^2} \int\limits_0^\infty \frac{d\varepsilon}{\left(e^{\frac{\varepsilon-\mu}{T}} - 1\right)}, \qquad g_s = 2s+1$$

интегрируем заменой, вводим $T_B := 2\pi\hbar^2 N/g_s m S$, получаем

$$\mu_{\beta=0} = T \ln\left(1 - e^{-\frac{T_B}{T}}\right).$$

(??? еще раз удостоверюсь в этой формуле!!???)

При $T \ll T_B$ это малая отрицательная величина $\mu_{\beta=0} = -T\exp\left(-T_B/T\right)$. Напомним, что для двумерного газа бозе-конденсация отсутствует. (??? почему???)

Так же как и для фермионов, при включении магнитного поля химический потенциал изменяется на малую относительную поправку $\left(\mu_B\mathcal{B}/\mu_{\beta=0}\right)^2$. Коэффициент пропорциональности у этой поправки складывается из диамагнитного $-1/6$ и парамагнитного $+1/3$ вкладов.

Для бозе-газа со спином единица:

$$\begin{aligned}
\Omega =& T\iint \frac{S dp_x dp_y}{(2\pi\hbar)^2} \sum_{m_s=-1}^{1} \ln\left\{1 - \exp\left(\frac{\mu - m_s\mu_B B - p^2/2m}{T}\right)\right\} \overset{\substack{dp_x dp_y = 2\pi p dp \\ \varepsilon = p^2/2m}}{=} \\
=& \frac{mST}{2\pi\hbar^2} \int\limits_0^\infty d\varepsilon \sum_{m_s=-1}^{1} \ln\left\{1 - \exp\left(\frac{\mu - m_s\mu_B B - \varepsilon}{T}\right)\right\} \overset{\text{b.parts}}{=} \\
=& \frac{mS}{2\pi\hbar^2} \int\limits_0^\infty \varepsilon d\varepsilon \sum_{m_s=-1}^{1} \frac{1}{\exp\left(\frac{m_s\mu_B\mathcal{B} + \varepsilon - \mu}{T}\right) - 1}.
\end{aligned}$$

Для вырожденного газа $T \ll T_B$ в слабых полях: $-\mu \ll T, \mu_B\mathcal{B} \ll T$ главный вклад в восприимчивость дает слагаемое с $m_s = 0$. После замены переменной $\varepsilon = Tx - m_s\mu_B B + \mu$ получаем

$$\Omega = \frac{3mST}{2\pi\hbar^2} \int\limits_{-\frac{\mu}{T}}^\infty \frac{(Tx+\mu)dx}{\exp(x) - 1}.$$

Выделяем в (8.41) слагаемое, зависящее от магнитного поля, и с учетом (8.36) получаем парамагнитную восприимчивость бозе-газа:

$$\chi_{\text{пара}} = \frac{2N\mu_B^2}{3\mu_{\mathcal{B}=0}}.$$

Далее вычислим Ω-потенциал диамагнитного бозе-газа:

$$\Omega = Tg_L \sum_{n=0}^{\infty} \ln\left\{1 - \exp\left(\frac{\mu - \varepsilon_n}{T}\right)\right\}.$$

Далее после какой-то математики и формулы Эйлера-Маклорана приходим к тому, что главная часть Ω-потенциала, связанная с магнитным полем, равна

$$\Omega(B) = -\frac{1}{6}\frac{(\mu_B\mathcal{B})^2}{\mu_{\beta=0}}.$$

Отсюда находим диамагнитную восприимчивость вырожденного бозе газа:

$$\chi_{\text{dia}} = -\frac{N\mu_B^2}{3\mu_{B=0}}.$$

Полученные результаты показывают, что при $T \to 0$ бозе-газ демонстрирует огромный магнитный отклик на внешнее поле. Точно такой же ответ получается и в трехмерном случае вблизи температуры начала бозе-конденсации T_B. Большая величина магнитной восприимчивости связана с тем, что значительное число атомов газа находятся в когерентном состоянии в конденсате или близком к конденсату.

(??? абзац про то, нужна ли где-то эта модель или нет???)

3.2.5 Electronic and Nuclear Paramagnetic Resonance

см. Боровик 17)

§

17.1. Открытие парамагнитного резонанса257

§

17.2. Теория ядерного парамагнитного резонанса258

§

17.3. Экспериментальные методы наблюдения ЯПР265

§

17.4. Применение ЯПР267

§

17.5. Теория электронного парамагнитного резонанса271

§

17.6. Методы наблюдения ЭПР276

§

17.7. Применение ЭПР в химии и биологии280

§

17.8. Влияние внутрикристаллического поля и анизотропия?-фактора284

§

17.9. Квантовые парамагнитные усилители289

3.2.6 Giant Magnetoresistance (!?!)

(если нобелевскую премию дали, значит стоит изучить суть)

Модель Валета - Ферта

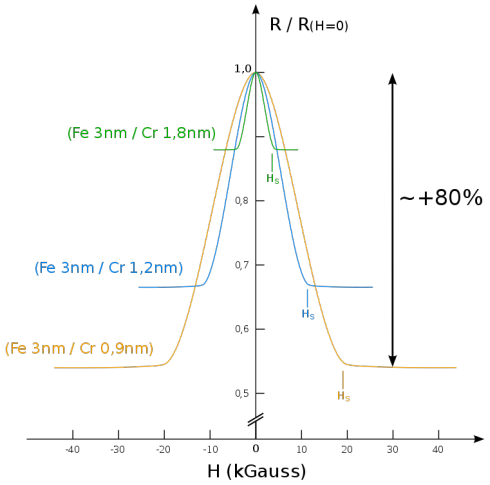
В 1993 году Тьери Валетом (англ. Thierry Valet) и Альбером Фертом была опубликована модель гигантского магнетосопротивления для CPP-геометрии, построенная на основе уравнений Больцмана. Суть теории заключается в рассмотрении расщепления химического потенциала на две функции внутри магнитного слоя, соответствующие электронам со спинами параллельными и антипараллельными намагниченности в нём. Если считать, что толщина немагнитного материала достаточно мала, то во внешнем поле E_0 поправки к электрохимическому потенциалу и полю внутри образца будут иметь вид

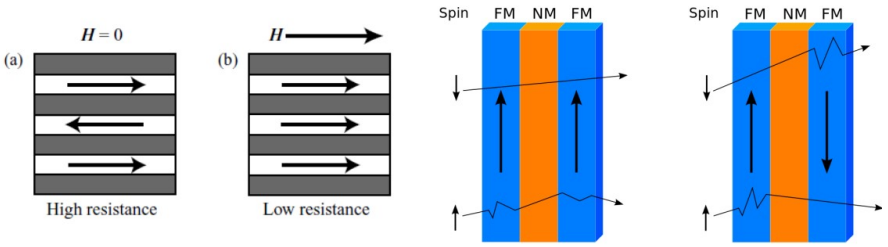
$$\Delta\mu = \frac{\beta}{1-\beta^2} eE_0 l_s e^{z/l_s},$$
$$\Delta E = \frac{\beta^2}{1-\beta^2} eE_0 l_s e^{z/l_s},$$

где I_s - средняя длина спиновой релаксации, а координата z отсчитывается от границы между магнитным и немагнитным слоями ($z < 0$ соответствует ферромагнетик). Отсюда следует, что на границе ферромагнетика будут накапливаться те электроны, для которых химический потенциал больше, что можно представить в виде потенциала спиновой аккумуляции V_AS, или так называемого интерфейсного сопротивления (присущего границе интерфейса ферромагнетик - немагнитный материал)

$$R_i = \frac{\beta\left(\mu_{\uparrow\downarrow} - \mu_{\uparrow\uparrow}\right)}{2ej} = \frac{\beta^2 l_{sN}\rho_N}{1 + \left(1-\beta^2\right) l_{sN}\rho_N / \left(l_{sF}\rho_F\right)},$$

где j - плотность тока в образце, I_SN и I_SF - длины спиновой релаксации в немагнитном и магнитном материалах соответственно [33].





3.3 Special General Theories in a Nutshell

3.3.1 Landau-Ginzburg Theory in Magnetism

(??? it is maybe famous)

3.3.2 Magneto-Optics

16.1 Magneto-optics basics 189 16.1.1 Kerr effect 189 16.1.2 Faraday effect 191 16.1.3 Physical origin of magneto-optic effects 191

3.3.3 Magnetic Hydrodynamics in a Nutshell

(просто дополнительный раздел, когда-то додумаю указания, вряд ли кому-то он нужен был на самом деле. пока н едо него.)

Уравнения движения жидкости в магнитном поле

329 66

Диссипативные процессы в магнитной гидродинамике

333 67

Магнитогидродинамическое течение между параллельными плоскостями

337 68

Равновесные конфигурации магнитных жидкостей

339 69

Магнитогидродинамические волны

344 70

Условия на разрывах

350 71

Тангенциальные и вращательные разрывы

351 72

Ударные волны

357 73

Условие зволюционности ударных волн

361 74

Турбулентное динамо

3.3.4 Magnetoresistance

(выпишу из кинетики, уже там думал про это! ничего не усвоено пока что)

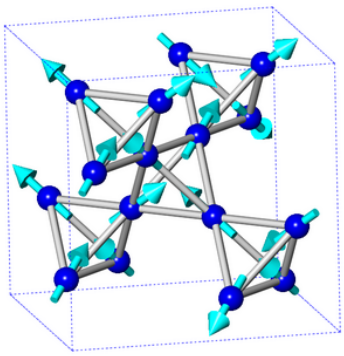
3.3.5 Spin liquid, glass, ice

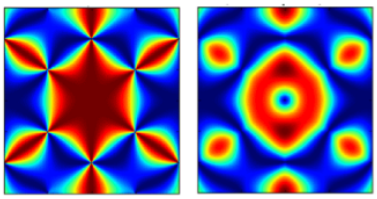
Spin glass in a nutshell (????)

(немного смотрел, но не до этого так или иначе)

Spin ice in a nutshell (????)

(немного смотрел, но не до этого так или иначе)





4 Famous device applications and novel materials in a nutshell

4.1 Typical device applications

4.1.1 Materials, alloys

(по Боровику п.25)

§

25.1. Характеристики постоянных магнитов

§

25.2. Магнитножесткие сплавы

§

25.3. Материалы из порошков и ферритов

4.1.2 Indirect exchanges: superexchange, Goodenough-Kanamori, Zener, RKKY, D-M

17.1.1 Direct exchange and superexchange 199

Indirect Exchange

* It is the coupling between magnetic moments over long distance and requires a mediator. & It can be of two types (a) Super Exchange (b) RKKY Exchange

17.1.2 Carrier-mediated exchange 199

17.1.3 Bound magnetic polarons (????)

The concept of bound magnetic polarons (BMPs) in connection with magnetic semiconductors was first introduced to explain the low-temperature metal–insulator transition in oxygen-deficient EuO [85]. In the BMP model, oxygen vacancies act both as electron donors and as electron traps, which can bind the electrons and maintain insulating behavior. Each trapped electron couples the local moments of the host lattice that lie within its orbit ferromagnetically, leading to a bound polaron with a large net magnetic moment. If neighboring polarons do not interact strongly, a paramagnetic, insulating phase results. However, for certain polaron–polaron distances and combinations of electron–electron and electron–local-moment exchange constants, the polarons couple ferromagnetically [86, 87]. The critical distance above which the exchange between two BMPs becomes ferromagnetic is typically of the order of a few Bohr radii [87].

The amplitude of the exchange interaction then drops off rapidly with distance. Above some critical electron density, the attractive potential of the vacancy is screened, the donor electrons become unbound, and the system becomes metallic [88].

Goodenough-Kanamori rules

Strong negative coupling when $\mathbf{M}_1 - \mathbf{O} - \mathbf{M}_2$ angle is equal to 180°: There is a strong antiferromagnetic exchange interaction if the half-filled orbitals of two cations overlap with the same empty or filled orbital of the intervening anion.

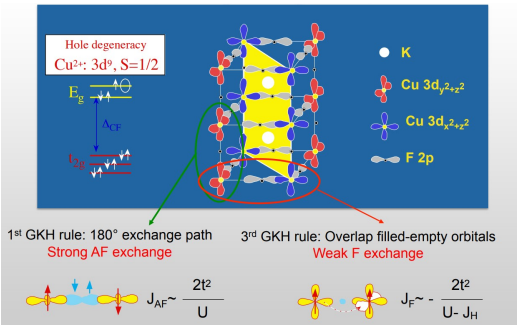
Weak positive coupling when $\mathbf{M}_1 - \mathbf{O} - \mathbf{M}_2$ angle is equal 90°: There is a weaker ferromagnetic exchange interaction if the half-filled orbitals of two cations overlap with orthogonal orbitals of the same intervening anion.

Goodenough-Kanamori-Anderson rules:

RULE 1: Half-filled orbitals: 180° superexchange

RULE 2: Half-filled 90° exchange

RULE 3: Overlap half-filled and empty 180° exchange



Superexchange.

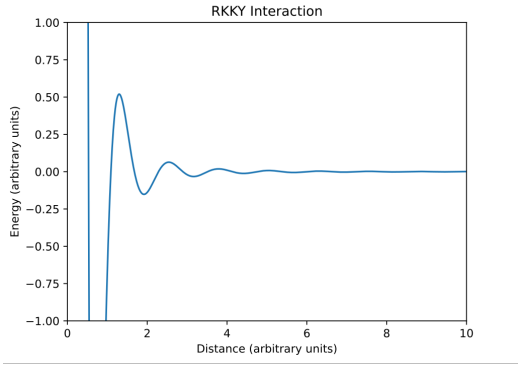
In many transition-metal oxides and related materials, magnetic interactions between the transition-metal ions are mediated by the intermediate anions. Such anion-mediated magnetic coupling is known as superexchange; we discussed this in Chapter 8 to explain the antiferromagnetism in MnO, in Chapter 9 in the context of the ferrites, and in Chapter 13 to understand the behavior of the colossal magnetoresistive manganites. Superexchange can also be described by a Heisenberg Hamiltonian, in which the sign of J_{ij} is determined by the metaloxygen-metal bond angle and the d electron configuration on the transition metal. These dependencies are articulated in the semi-empirical Goodenough-KanamoriAnderson rules (see, for example, [80]). Importantly for this chapter, 180° metaloxygen-metal angles between identical metal ions with both orbitals either filled or empty lead to antiferromagnetic interactions; 90° angles, on the other hand, result in ferromagnetism.

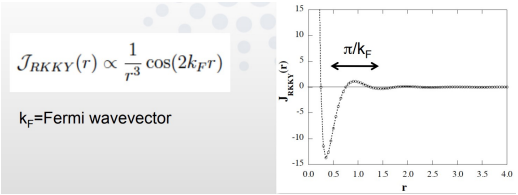
Indirect Exchange in Ionic Solid: Super Exchange

* In systems in which direct exchange cannot be realized due to insufficient overlap of magnetic orbitals, magnetic coupling may be mediated by orbitals of a nonmagnetic ligand in between them. It is the super exchange interaction, which is responsible for the magnetic properties of the most of magnetic materials, especially nonmetallic compounds, for example, oxides or fluorides. Generally found in metal oxides where the magnetic atoms are separated by non magnetic ions (O^{2-}).

RKKY

(оно тут используется, что это - не знаю, мб когда-то пойму)





In order for a band to spontaneously spin-split, the following condition must be satisfied: (??????)

$$Ug\left(E_F\right) \geq 1 \qquad \text{(Stoner criterion)}$$

where $U \propto \lambda$ and $g\left(E_F\right)$ is the density of states at the Fermi level. Materials in which the Stoner criterion is satisfied are spontaneously ferromagnetic.

Zener double-exchange model (???)

It explains the experimentally observed ferromagnetism in transition-metal oxides containing magnetic cations with two different valence states. In $La_{1-x}Ca_xMnO_3$, for example, with $0 < x < 1$, both Mn^{4+} (with three 3 d electrons) and Mn^{3+} (with four 3 d electrons) are present. The kinetic energy of the system is lowered if the magnetic moments align parallel, since parallel alignment allows electron transfer from Mn^{3+} to Mn^{4+}. This indirect coupling is again mediated by the oxygen atoms between neighboring Mn^{3+} and Mn^{4+} ions, but is distinguished from superexchange by the involvement of carriers.

Anisotropic Exchange Interaction (Dzyaloshinsky-Moriya (D-M) interaction)

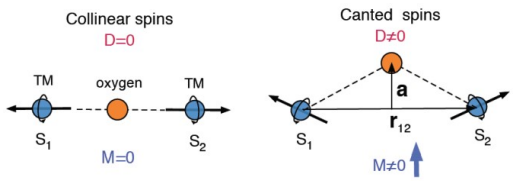
* This is also known as Dzyaloshinsky-Moriya (D-M) interaction. Here spin-orbit interaction plays the role as the oxygen atoms act in super exchange.

* The excited state is produced by the spin-orbit interaction in one of the magnetic ions and an exchange interaction occurs between excited state of one ion and ground state of other ion.

$\star$ This interaction includes a term in the Hamiltonian, $\widehat{\mathbf{H}}_{DM} = \overrightarrow{\mathbf{D}}.\overrightarrow{\mathbf{s}_1} \times \overrightarrow{\mathbf{s}_2}\ \overrightarrow{\mathbf{D}}$ will lie parallel or perpendicular to the line connecting two spins, depending on the symmetry.

$\therefore$ This interaction is such that it tries to force $\overrightarrow{\mathbf{s}_1}$ and $\overrightarrow{\mathbf{s}_2}$ to be at right angles in a plane perpendicular to $\overrightarrow{\mathbf{D}}$ in such an orientation as to ensure that energy is negative. Its effect is therefore very often cant the spins by small angle.

* For this spin canting, antiferromagnetic materials show some non-zero magnetic moment near absolute zero (Weak Ferrimagnetism).



4.1.3 Magnetic semiconductors, insulators (??)

(отдельно изучать нужно, куча информации)

Exchange interactions in magnetic semiconductors and insulators

17.2 II–VI diluted magnetic semiconductors – (Zn,Mn)Se 201
17.2.1 Enhanced Zeeman splitting 201
17.2.2 Persistent spin coherence 202
17.2.3 Spin-polarized transport 203
17.2.4 Other architectures 204
17.3 III–V diluted magnetic semiconductors – (Ga,Mn)As 204
17.3.1 Rare-earth–group-V compounds – ErAs 207
17.4 Oxide-based diluted magnetic semiconductors 208
17.5 Ferromagnetic insulators 210
17.5.1 Crystal-field and Jahn–Teller effects 210
17.5.2 YTiO3 and SeCuO3 211
17.5.3 BiMnO3 213
17.5.4 Europium oxide 214
17.5.5 Double perovskites 215

II–VI diluted magnetic semiconductors – (Zn,Mn)Se

Bulk Mn chalcogenides crystallize in either the hexagonal NiAs structure (α - MnTe) or in the cubic NaCl structure ($\alpha -$ MnSe and $\alpha -$ MnS). Only MnS has been grown in the cubic zincblende structure ($\beta -$ MnS) in bulk, but all Mn chalcogenides can be stabilized artificially in the zincblende structure by epitaxial growth, or by alloying with II-VI semiconductors [89]. The II-VI diluted magnetic semiconductors have been the subject of extensive experimental studies in the past and it has been established that the magnetic structure for sufficiently high manganese concentrations ($x_{\mathrm{Mn}} > 0.6$ for $(\mathrm{Zn,Mn})\mathrm{Se}$) is characterized by an antiferromagnetic correlation between the Mn^{2+} ($(3\ d)^5$) magnetic moments. For intermediate manganese concentrations, spin-glass structures are found, and for low concentrations ($x_{\mathrm{Mn}} < 0.3$ for $(\mathrm{Zn,Mn})\mathrm{Se}$), a paramagnetic behavior of the uncorrelated Mn spins becomes dominant [90]. The paramagnetic regime is particularly interesting, since the strong interaction between the Mn d states and the sp bands of the host semiconductor gives rise to a hundred-fold increase in the effective g-factors [91]. The resulting properties include enhanced Zeeman splitting, spin precession and persistent spin coherence, spin-polarized luminescence, and spin-polarized transport. We discuss these phenomena in more detail below.

(??? я вообще хз)

On III–V diluted magnetic semiconductors – (Ga,Mn)As

(ну есть такое, не актуально сейчас вообще)

$$H = -N_0\beta\vec{s}\cdot\vec{S}$$

where N_0 is the concentration of the cation sites and β is the p − d exchange integral. The product $N_0\beta$ is usually called the exchange constant. If one simply uses the mean-field approximation [100, 101], in which the magnetizations of both carriers and Mn ions are considered to be uniform in space, one obtains an expression for the Curie temperature:

$$T_{\mathrm{C}} = \frac{xN_0S(S+1)\beta^2\chi_s}{3k_{\mathrm{B}}\left(g^*\mu_{\mathrm{B}}\right)^2},$$

Oxide-based diluted magnetic semiconductors

Clearly, for device applications it is desirable to find a DMS material with a Curie temperature at or above room temperature. In addition to the current research aimed at understanding the origin of ferromagnetism in (Ga, Mn)As, there is also considerable exploration of novel DMS materials in the search for higher Curie temperatures. In particular, work on diluted magnetic semiconductors with oxide hosts was motivated by a prediction of above-room-temperature ferromagnetism within the Zener model for ZnO doped with 5%Mn and a large hole concentration of 3.5×10^{20} per cubic centimeter [101]. Oxides also have the advantage over selenides and arsenides of being earth-abundant and environmentally benign. Ferromagnetism with T_{C} above room temperature has indeed been reported in several oxide-based DMSs, including $Ti_{1-x}Co_xO_2$ [112], $Zn_{1-x}Mn_xO$ [113], and $Zn_{1-x}Co_xO$ [114]. Since many of the reports of ferromagnetism occur in insulating samples, an explanation based in bound magnetic polarons [115] is the most plausible. Results from different experiments, and indeed different computations, are widely conflicting, however, and it is still unclear whether the reports represent true ferromagnetic DMS behavior, or whether small concentrations of ferro- or ferrimagnetic impurity phases [116] or even residual magnetic ions in the magnetometers, are dominating the measurements. For a review of the conflicting experimental and theoretical results, and their possible origins, see [79]. This is clearly an area in which well-characterized

4.1.4 Magnetic data storage

(see Spaldin, interesting, but now no time for it)

Magnetic media 181 15.2.1 Materials used in magnetic media 181 15.2.2 The other components of magnetic hard disks 183 15.3 Write heads 183 Contents xi 15.4 Read heads 185 15.5 Future of magnetic data storage

5 Other Theoretical models in a nutshell

5.1 Overview of special interactions in magnets

(все это пока не актуально вообще, нужно будет - каждую тему и доучу)

5.1.1 RKKY Interaction (???)

Суть модели (????)

РККИ - обменное взаимодействие (Рудермана- Киттеля - Касуя - Иосиды) - косвенное обменное взаимодействие между магнитными ионами, осуществляемое через коллективизированные электроны проводимости.

Считаем, что электроны проводимости движутся в эффективном поле, создаваемом локализованным магнитным моментом одного узла.

$$E_{ij} = \frac{J^2}{\mu_0 \left(g\mu_B\right)^2} \mathbf{S}_i \mathbf{S}_j \chi_{ij}, \qquad \chi_{ij} := \frac{(x\cos x - \sin x)}{x^4}, \qquad x := 2\mathbf{k}_F \mathbf{R}_{ij} \gg 1$$

тут J - константа обмена, g - фактор Ланде, μ_B - магнетон Бора, μ_0 - магнитная постоянная, Здесь k_F - волновой вектор на уровне Ферми, R_{ij} - расстояние. Поэтому энергия взаимодействия имеет знакопеременный осциллирующий характер как функция от расстояния между взаимодействующими ионами. Знакопеременность энергии взаимодействия приводит к тому, что в зависимости от (???)

Это обменное взаимодействие возникает в металлах и полупроводниках, где коллективизированные электроны выступают посредниками обменного взаимодействия ионов, обладающих локализованными противоположно направленными спинами, частично заполненных d и f оболочек. Электроны проводимости взаимодействуют с эффективным магнитным полем i-го узла кристаллической решётки и приобретают некую спиновую поляризацию. При прохождении следующего узла решётки, релаксация магнитных моментов электрона и узла вызовет обоюдные изменения как спиновой поляризации, так и спина узла решётки.

(??? абзац, почему такая модель может что-то дать, что спиновые модели не могут???? пока не знаю.)

(??? додумаю потом)

5.1.2 Nagoaka's Model (???)

(когда-то почитаю про это, уже ниже есть в этой записи теория про это)

5.1.3 Other 1D Models by Tasaki (???)

(там много крутых тем, сперва в конд средах, потом сюда добавлю это. скорее всего пару разделов сделаю про них.)

5.1.4 Beta Ansatz

(see stat phys)

5.1.5 Anisotropic Magnetic Interactions

(???)

5.1.6 BKT Transition

(see cond matter or SC)

5.1.7 Frustrated Systems

(????)

5.1.8 Kitaev Model

(??)

5.1.9 Dzyaloshinskii-Moriya interaction

(???)

5.1.10 Condo Effect

(??? see cond matter or special mathem models)

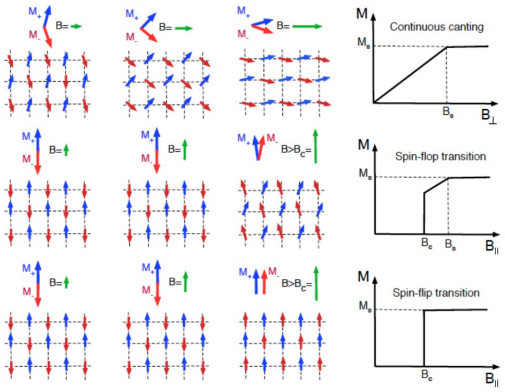
6 Other effects and applications in a nutshell

6.1 Other effects

(окажется важным - передвину выше)
(ниже - в основном другие темы по Боровику, Еременко, Мильнеру)

6.1.1 Spin-Flop Transition (??)

(что-то не понял все еще, что это? пока считаю не таким актуальным.)



6.1.2 Other Properties of Magnetization Curves

(Боровик 13)

§

13.1. Два типа процессов намагничивания188

§

13.2. Начальный участок кривой намагничивания189

§

13.3. Обратимые и необратимые процессы смещения границ191

§

13.4. Обратимые процессы вращения. Теория Акулова195

§

13.5. Влияние упругих напряжений на намагничивание ферромагне- тиков 199

§

13.6. Эффект Гопкинсона200

6.1.3 Magnetic Hysteresis

(Боровик 14)

§

14.1. Причины и типы гистерезиса2

§

14.2. Гистерезис, обусловленный задержкой смещения границ между доменами201

§

14.3. Гистерезис, обусловленный задержкой роста зародышей перемагничивания

§

14.4. Гистерезис, обусловленный необратимым процессом вращения

§

14.5. Магнитная анизотропия у антиферромагнетиков и смещённая петля гистерезиса

6.1.4 Magnetic Nanoparticles

(????????? maybe there is a lot of info about them)

6.1.5 Antiferromagnetic Resonance and Afm Interaction with Light

(см Боровик 22)

§

22.1. Антиферромагнитный резонанс — метод определения энергетической щели в спин-волновом спектре антиферромагнетика...391

§

22.2. Поглощение света в антиферромагнетиках. Исследование спин-волнового спектра вблизи границы зоны Бриллюэна406

§

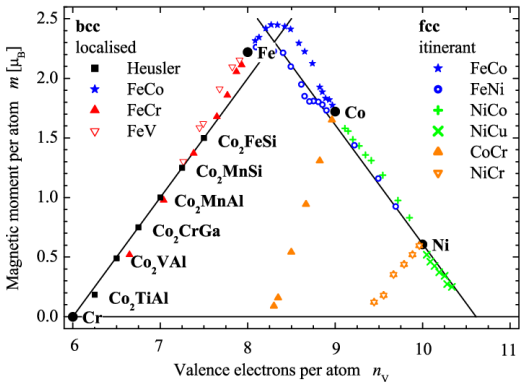
22.3. Рассеяние света в магнитоупорядоченных кристаллах418

6.1.6 Soft Magnetic Materials and Alloys

Sleter-Pauling curve for alloys (???)

the Slater–Pauling rule states that adding an element to a metal alloy will reduce the alloy's saturation magnetization by an amount proportional to the number of valence electrons outside of the added element's d shell. Conversely, elements with a partially filled d shell will increase the magnetic moment by an amount proportional to number of missing electrons.

Investigated by the physicists John C. Slater and Linus Pauling in the 1930s, the rule is a useful approximation for the magnetic properties of many transition metals.



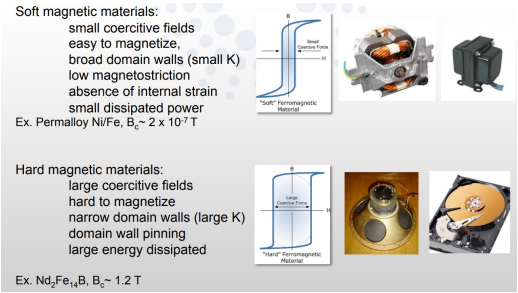
(?? I'll ask professionals about it if needed)

§

23.1. Введение
(см Боровик 23.)

§

23.2. Железо — мягкий магнитный материал426



§

23.3. Электротехническая железо-кремнистая сталь427

§

23.4. Железоникелевые сплавы431

§

23.5. Мягкие магнитные материалы с особыми магнитными свойствами436

6.1.7 High-Frequency Magnetic Materials and Their Applications

(см Боровик 24.)

§

24.1. Ферриты442

§

24.2. Магнитные свойства ферритов443

§

24.3. Некоторые магнито-оптические явления в ферритах447

§

24.4. Некоторые применения высокочастотных ферромагнетиков

§

24.5. Основные типы ферритов, применяемых в технике высоких частот450

6.2 Other applications and magnetic devises

6.2.1 Magnetoresistive RAM, magnetic memory materials

(есть такие, что-то про них когда-то мб напишу, изучать пока нет планова)

§

27.1. Запись информации в магнитофоне485

§

27.2. Запись информации на ЦМД (Bubble memory)

§

27.3. Технологические основы ЦМД-устройств (Bubble memory)

§

27.4. Перспективы развития магнитной памяти491

6.2.2 Magneto-optic recording

16.2.1 Other types of optical storage, and the future of magneto-optic recording

6.3 Special magnetic field in space

6.3.1 On special magnetostatics

(весь сложный магнетизм будет указан тут)

6.3.2 Magnetism of the Planets and Between Them

(????)

О магнитосфере Земли (????)

(она там еще смещается, короче, это загадка для меня, и не актуально пока это)

6.4 Experiments on Magnetism in a Nutshell

6.4.1 X-ray magnetic scattering

See wonderful lecture by L. Paolasini
https://www.esrf.fr/files/live/sites/www/files/events/Paolasini_magnetism%20lecture9.pdf
(not my field of work, but here is a lot of information)

6.4.2 Experimental Measurement of Magnetic Susceptibility

(это я потом когда-то досмотрю!)

3.1. Источники магнитного поля

(интересно, но не до этого)

Типичный источник поля - соленоид:

$$H = 0,4\pi\frac{N}{l}i,$$

тут l - длина соленоида в см; N - число витков; i - ток в Амперах. Или, с учетом конечной длины и толщины обмотки, $H = 0,4\pi\frac{N}{l}i\left[\frac{l/2+x}{2(b-a)}\ln\frac{b+s_2}{a+s_1} + \frac{l/2-x}{2(b-a)}\ln\frac{b+r_2}{a+r_1}\right]$. Здесь x - расстояние по оси от середины соленоида; $2a$ и $2b$ - внутренний и наружный диаметры обмотки соленоида соответственно (рис. 3.1);

$$s_1^2 = a^2 + \left(\frac{l}{2}+x\right)^2; \quad s_2^2 = b^2 + \left(\frac{l}{2}+x\right)^2; \quad r_1^2 = a^2 + \left(\frac{l}{2}-x\right)^2 \quad r_2^2 = b^2 + \left(\frac{l}{2}-x\right)^2$$

$$\oint H_l dl = \frac{4\pi}{c}\int_s j_n ds = \frac{4\pi}{c}Ni$$

Однако из-за наличия железа поле на большей части l_{Fe} контура интегрирования сильно ослаблено и

$$\oint H_l dl \cong H_{\text{B}} l_{\text{B}}.$$

Если $l_{\text{в}}$ меньше размера катушек ($2l_{\text{к}}$), то мы получаем усиление поля по сравнению с полем $H_{\text{к}}$, которое можно создать в тех же катушках без железа.

Если i выражать в амперах, то поле, созданное катушками в воздушном зазоре электромагнита, даётся формулой

$$H_{\text{B}} \sim \frac{4\pi}{10}\frac{N}{l_{\text{B}}}i.$$

Поле в тех же катушках без железа

$$H_{\text{K}} = \frac{4\pi}{10}k\frac{N}{2l_{\text{K}}}i,$$

3.2. Макроскопические характеристики магнитных свойств вещества и основные методы их измерения

В настоящее время международным соглашением принята единая система единиц СИ. В ней, наряду с массой M, длиной L и временем t, в число основных величин входит сила тока i. В СИ индукция B измеряется в единицах, носящих название тесла (1 Т $= 10^4$Гс), а магнитное поле и намагниченность - в амперах на метр (1 А/м $= 1/79,5775$ Э). Таким образом, в этой системе единиц размерности магнитной индукции и магнитного поля не одинаковы:

$$[B] = \left[Mt^{-2}i^{-1}\right]; \quad [H] = \left[L^{-1}i\right].$$

В результате приходится ввести не имеющее физического смысла понятие магнитной проницаемости вакуума $-\mu_0$. Величина μ_0 имеет размерность

$$[\mu_0] = \left[LMt^{-2}i^{-2}\right]$$

$$\mu_0 = 4\pi \cdot 10^{-7}\frac{\text{Гн}}{\text{м}} = 1,256637 \cdot 10^{-6}\frac{\text{Гн}}{\text{м}}.$$

Уравнения, связывающие индукцию и поле, в этом случае приобретают вид

$$B = \mu_0\mu H = 1,256637 \cdot 10^{-6}\mu H = 1,256637 \cdot 10^{-6}(H + 4\pi I).$$

Поскольку определить с большой точностью массу тела легче, чем объём, результаты измерений часто выражают не в виде объёмной ($\varkappa$), а в виде массовой ($\varkappa_\rho$) восприимчивости:

$$\varkappa_\rho = \frac{1}{\rho}\varkappa,$$

где ρ - плотность вещества.

При сравнении магнитных свойств разных тел целесообразно сравнивать магнитные моменты, приходящиеся на 1 г.моль вещества. Соответственно вводят молярную восприимчивость:

$$\varkappa_\text{М} = m\varkappa_\rho = \frac{m}{\rho}\varkappa = V_m\varkappa,$$

где m - молекулярный вес; V_m - молярный объём. По значениям $\varkappa$ вещества можно разделить на три группы:

I	II	III
$\varkappa < 0$	$\varkappa > 0$	$\varkappa > 0$
$\lvert\varkappa\rvert \ll 1$	$\lvert\varkappa\rvert \ll 1$	$\lvert\varkappa\rvert > 1$
(диамагнетики)	(парамагнетики)	(ферромагнетики)

3.3. Методы измерения восприимчивости слабомагнитных тел

(вообще не актуально сейчас)

Как уже говорилось, большинство методов исследования свойств слабомагнитных тел основано на измерении сил, действующих на них в неоднородном поле. Энергия тела в магнитном поле равна

$$U = -\int HdM = -\Delta V\int HdI = -\Delta V\int_0^H H\varkappa dH = -\frac{1}{2}\varkappa H^2\Delta V,$$

откуда

$$f_x = -\frac{\partial U}{\partial x} = \frac{1}{2}\Delta V\varkappa\frac{\partial}{\partial x}\left(H^2\right) = \Delta V\varkappa H\frac{\partial H}{\partial x}$$

(Метод Фарадея какой-то есть???)

Гуи разработал метод, в котором применяются образцы значитель- ных размеров, так что уже нельзя, как в методе Фарадея, считать поле на протяжении образца неизменным.

3.4. Методы исследования тел с большой восприимчивостью

(вообще не актуально сейчас)

При исследовании свойств тел с большой восприимчивостью наибольшее распространение получил так называемый баллистический метод. В этом методе для измерения используется явление электромагнитной индукции. На образец надевается катушка, соединённая с гальванометром. При изменении намагниченности образца изменяется поток Φ, пронизывающий катушку, и в ней возникает ЭДС электромагнитной индукции E. Если время, в течение которого изменяется намагниченность, мало по сравнению с периодом гальванометра, то отброс последнего будет пропорционален

$$\int E dt = k \int_{t_1}^{t_2} \frac{\partial \Phi}{\partial t} dt = k \left(\Phi_2 - \Phi_1 \right).$$

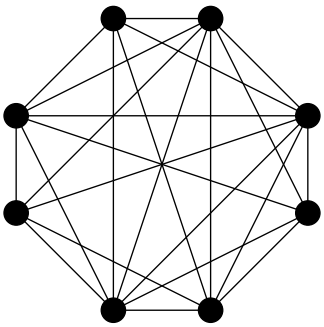
(?? как-то это делается, ну и ладно, не актуально никогда было)

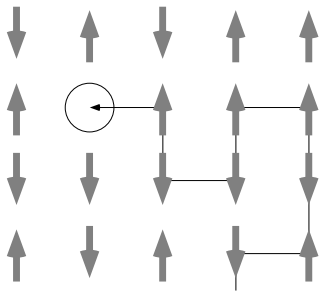
6.4.3 Experimental Methods for Studying Magnetic Materials

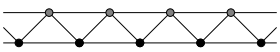
7.3. Гиромагнитные опыты

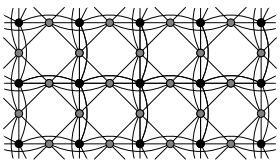
(не уверен, что это актуально)

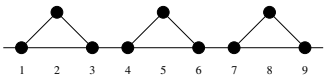
7 Other Pictures



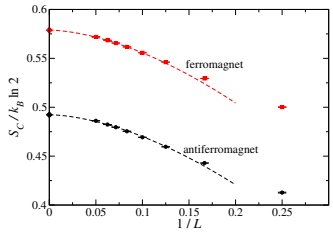


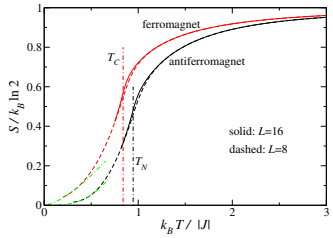


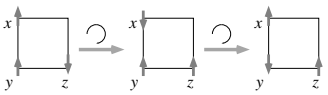


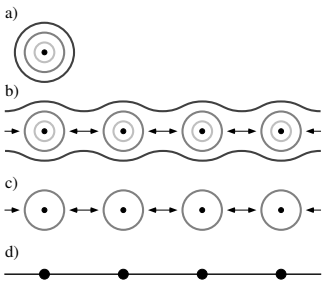

1 2 3 4 5 6 7 8 9

1

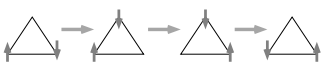

ferromagnet
antiferromagnet
$S_C / k_B \ln 2$
1 / L

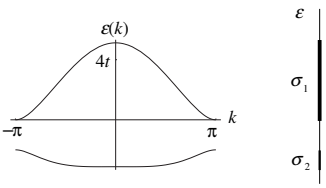

ferromagnet
antiferromagnet
T_C
T_N
solid: L=16
dashed: L=8
$S / k_B \ln 2$
$k_B T / |J|$

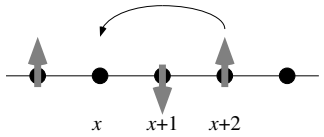

x
y
z

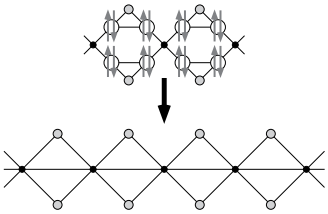

a)
b)
c)
d)

2

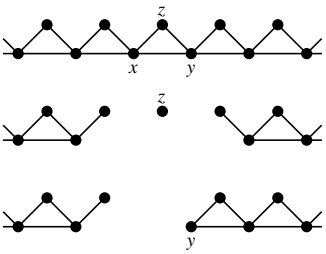


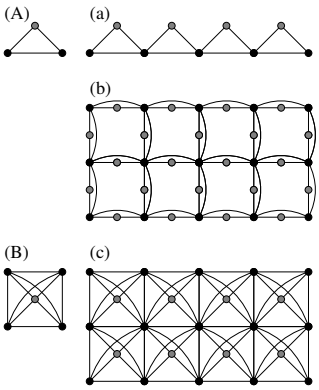

ε(k)
4t
−π
π
k
ε
σ1
σ2

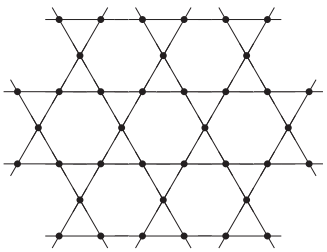


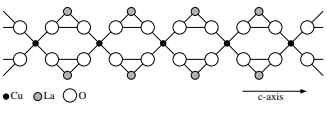


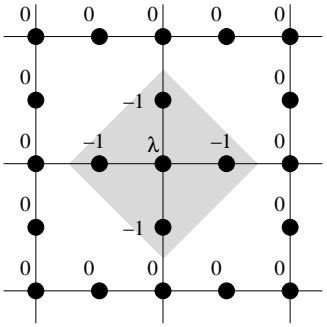
3










0 0 0 0 0
0 −1 0
0 −1 λ −1 0
0 −1 0
0 0 0 0 0

Part III
Theory

8 Fundamentals of Magnetism (???)

8.0.1 On Concept of Spin

(спишу из квантмеха, там об этом все уже писал.)

9 Magnetism in Condensed Matter by Blundell

9.0.1 Preface

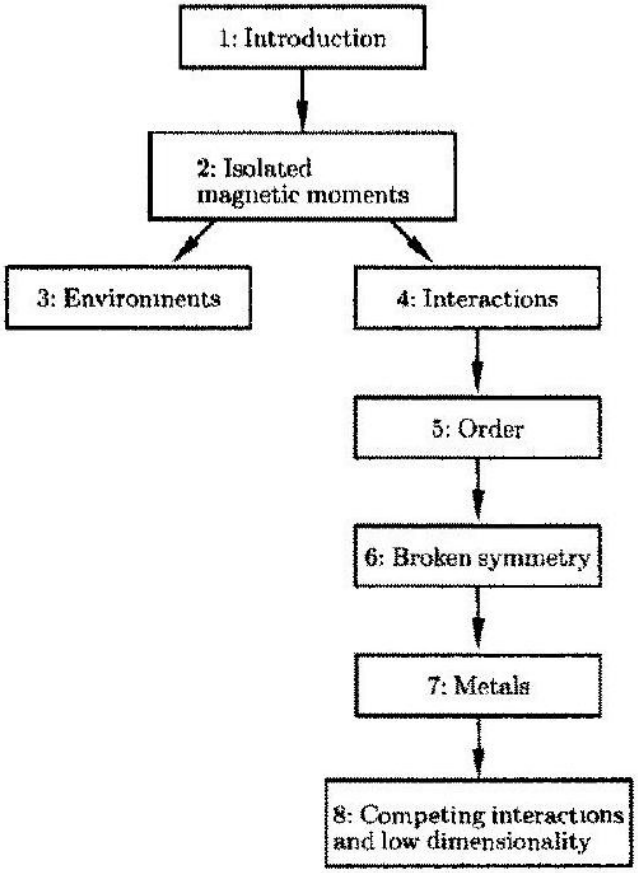
'... in Him all things hold together.'
(Colossians 1^{17})

Magnetism is a subject which has been studied for nearly three thousand years. Lodestone, an iron ore, first attracted the attention of Greek scholars and philosophers, and the navigational magnetic compass was the first technological product resulting from this study. Although the compass was certainly known in Western Europe by the twelfth century AD, it was not until around 1600 that anything resembling a modern account of the working of the compass was proposed. Progress in the last two centuries has been more rapid and two major results have emerged which connect magnetism with other physical phenomena. First, magnetism and electricity are inextricably linked and are the two components that make up light, which is called an electromagnetic wave. Second, this link originates from the theory of relativity, and therefore magnetism can be described as a purely relativistic effect, due to the relative motion of an observer and charges moving in a wire, or in the atoms of iron. However it is the magnetism in condensed matter systems including ferromagnets, spin glasses and low-dimensional systems, which is still of great interest today. Macroscopic systems exhibit magnetic properties which are fundamentally different from those of atoms and molecules, despite the fact that they are composed of the same basic constituents. This arises because magnetism is a collective phenomenon, involving the mutual cooperation of enormous numbers of particles, and is in this sense similar to superconductivity, superfluidity and even to the phenomenon of the solid state itself. The interest in answering fundamental questions runs in parallel with the technological drive to find new materials for use as permanent magnets, sensors, or in recording applications.

This book has grown out of a course of lectures given to third and fourth year undergraduates at Oxford University who have chosen a condensed matter physics option. There was an obvious need for a text which treated the fundamentals but also provided background material and additional topics which could not be covered in the lectures. The aim was to produce a book which presented the subject as a coherent whole, provided useful and interesting source material, and might be fun to read. The book also forms part of the Oxford Master Series in Condensed Matter Physics; the other volumes of the series cover electronic properties, optical properties, superconductivity, structure and soft condensed matter.

The prerequisites for this book are a knowledge of basic quantum mechanics and electromagnetism and a familiarity with some results from atomic physics. These are summarized in appendices for easy access for the reader and to present a standardized notation.

Structure of the book:



Some possible course structures:
(1) Short course (assuming Chapter 1 is known):

- Chapter 2 (omit 2.6-2.8)
- Chapter 3 (omit 3.2)
- Chapter 4 (omit 4.2.5, 4.2.6)
- Chapter 5 (omit 5.4-5.7)
- Chapter 6 (omit 6.4-6.5)
- Chapter 7 (omit 7.1-7.2)
  (2) Longer course:
- Chapters 1-4
- Chapter 5 (5.6 and 5.7 as background reading)
- Chapter 6
- Chapter 7 (7.5-7.9 as background reading)
- Chapter 8, selected topics

The interesting magnetic effects found in condensed matter systems have two crucial ingredients: first, that atoms should possess magnetic moments and second, that these moments should somehow interact. These two subjects are discussed in Chapters 2 and 4 respectively. Chapter 2 answers the question 'why do atoms have magnetic moments?' and shows how they behave and can be studied if they do not interact. Chapter 3 describes how these magnetic moments can be affected by their local environment inside a crystal and the techniques which can be used to study this. Chapter 4 then answers the question 'how do the magnetic moments on different atoms interact with each other?' With these ingredients in place, magnetic order can occur, and this is the subject of Chapters 5 and 6. Chapter 5 contains a description of the different types of magnetic order which can be found in the solid state. Chapter 6 considers order again, but starts from basic ideas of broken symmetry and describes phase transitions, excitations and domains. A strong emphasis is the link between magnetic order and other types of broken-symmetry ground states like superconductivity. Chapter 7 is devoted to the magnetic properties of metals, in which magnetism can often be associated with delocalized conduction electrons. Chapter 8 describes some of the subtle and complex effects which can occur when competing magnetic interactions are present and/or the system has a reduced dimensionality. These topics are the subject of intense research activity and there are many outstanding questions which remain to be resolved. Throughout the text, I discuss properties and applications to demonstrate the implications of all these ideas for real materials, including ferrites, permanent magnets and also the physics behind various magneto-optical and magnetoresistance effects which have become of enormous technological importance in recent years. This is a book for physicists and therefore the emphasis is on the clear physical principles of quantum mechanics, symmetry, and electromagnetism which underlie the whole field. However this is not just a 'theory book' but attempts to relate the subject to real measurements and experimental techniques which are currently used by experimental physicists and to bridge the gulf between the principles of elementary undergraduate physics and the topics of current research interest.

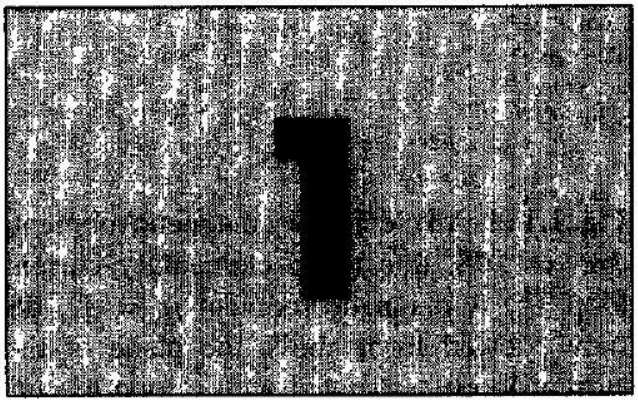
Chapters 1-7 conclude with some further reading and problems. The problems are of varying degrees of difficulty but serve to amplify issues addressed in the text. Chapter 8 contains no problems (the subjects described in this chapter are all topics of current research) but has extensive further reading.

It is a great pleasure to thank those who have helped during the course of writing this book. I am grateful for the support of Sönke Adlung and his team at Oxford University Press, and also to the other authors of this Masters series, Mansfield College, Oxford and the Oxford University Department of Physics have provided a stimulating environment in which to work. I wish to record my gratitude to my students who have sometimes made me think very hard about things I thought I understood. In preparing various aspects of this book, I have benefitted greatly from discussions with Hideo Aoki, Arzhang Ardavan, Deepto Chakrabarty, Amalia Coldea, Radu Coldea, Roger Cowley, Steve Cox, Gillian Gehring, Matthias Gester, John Gregg, Martin Greven, Mohamedally Kurmoo, Steve Lee, Wilson Poon, Francis Pratt, John Singleton and Candadi Sukumar. I owe a special debt of thanks to the friends and colleagues who have read the manuscript in various drafts and whose
exacting criticisms and insightful questions have immensely improved the final result: Katherine Blundell, Richard Blundell, Andrew Boothroyd, Geoffrey Brooker, Bill Hayes, Brendon Lovett, Lesley Parry-Jones and Peter Riedi. Any errors in this book which I discover after going to press will be posted on the web-site for this book which may be found at:

`http://users.ox.ac.uk/~sjb/magnetism/`

Most of all, I want to thank Katherine, dear wife and soulmate, who more than anyone has provided inspiration, counsel, friendship and love. This work is dedicated to her.

9.1 Introduction



This book is about the manifestation of magnetism in condensed matter. Solids contain magnetic moments which can act together in a cooperative way and lead to behaviour that is quite different from what would be observed if all the magnetic moments were isolated from one another. This, coupled with the diversity of types of magnetic interactions that can be found, leads to a surprisingly rich variety of magnetic properties in real systems. The plan of this book is to build up this picture rather slowly, piece by piece. In this introductory chapter we shall recall some facts about magnetic moments from elementary classical and quantum physics. Then, in the following chapter, we will discuss how magnetic moments behave when large numbers of them are placed in a solid but are isolated from each other and from their surroundings. Chapter 3 considers the effect of their immediate environment, and following this in Chapter 4, the set of possible magnetic interactions between magnetic moments is discussed. In Chapter 5 we will be in a position to discuss the occurrence of long range order, and in Chapter 6 how that is connected with the concept of broken symmetry. The final chapters follow through the implications of this concept in a variety of different situations. SI units are used throughout the book (a description of cgs units and a conversion table may be found in Appendix A).

9.1.1 1.1 Magnetic moments

The fundamental object in magnetism is the magnetic moment. In classical electromagnetism we can equate this with a current loop. If there is a current I around an elementary (i.e. vanishingly small) oriented loop of area |dS| (see Fig. 1.1(a)) then the magnetic moment $\mathbf{d}\mu$ is given by

$$\mathbf{d}\mu = I\mathbf{dS}, \tag{1.1}$$

and the magnetic moment has the units of Am^2. The length of the vector **dS** is equal to the area of the loop. The direction of the vector is normal to the loop and in a sense determined by the direction of the current around the elementary loop.

This object is also equivalent to a magnetic dipole, so called because it behaves analogously to an electric dipole (two electric charges, one positive and one negative, separated by a small distance). It is therefore possible to imagine a magnetic dipole as an object which consists of two magnetic monopoles of opposite magnetic charge separated by a small distance in the same direction as the vector **dS** (see Appendix B for background information concerning electromagnetism).

1.1 Magnetic moments
1.2 Classical mechanics and magnetic moments 6
1.3 Quantum mechanics of spin

Alber Einstein (1879-1955)
Wander Johannes de Haas (1878-1960)

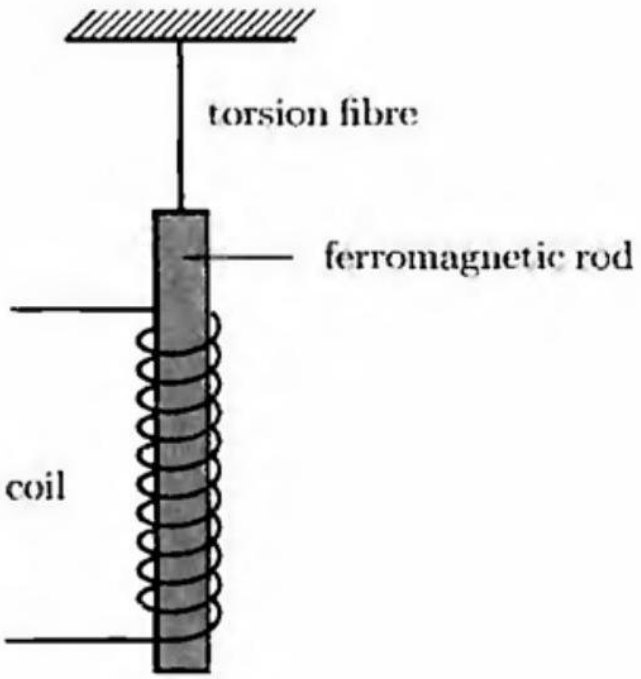
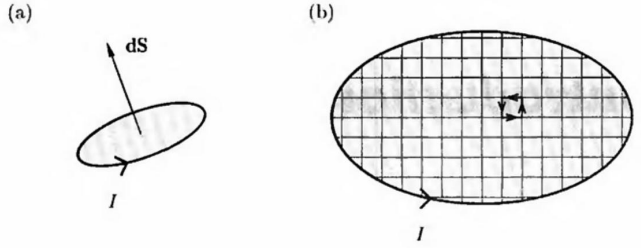


Fig. 1.2 The Einstein-de Haas effect. A ferromagnetic rod is suspended from a thin fibre A coil is used to provide a magnetic tield which magnetires the fertomagnet and produces a rotation. The experiment can be done resonantly, by periodically reversing the current in the coil, and hence the magnetization in the ferromagnet, and observing the angular response as a function of frequency.

Samuel Jackson Barnell (1873-1956)

Fig. 1.1 (a) An elementary magnetic moment. $\mathbf{d

\mu=I \mathrm{dS}$, due to an elementary current loop. (b) A magnetic moment $\mu = I \int \mathrm{dS}$ (now viewed from above the plane of the current loop) associated with a loop of current I can be considered by summing up the magnetic moments of lots of infinitesimal current loops.}



The magnetic moment $\mathbf{d}\boldsymbol{\mu}$ points normal to the plane of the loop of current and therefore can be either parallel or antiparallel to the angular momentum vector associated with the charge which is going around the loop. For a loop of finite size, we can calculate the magnetic moment μ by summing up the magnetic moments of lots of equal infinitesimal current loops distributed throughout the area of the loop (see Fig. 1.1(b)). All the currents from neighbouring

infinitesimal loops cancel, leaving only a current running round the perimeter of the loop. Hence,

$$\mu = \int \mathrm{d}\mu = I \int \mathrm{dS}. \tag{1.2}$$

1.1.1 Magnetic moments and angular momentum

A current loop occurs because of the motion of one or more electrical charges. All the charges which we will be considering are associated with particles that have mass. Therefore there is also orbital motion of mass as well as charge in all the current loops in this book and hence a magnetic moment is always connected with angular momentum.
In atoms the magnetic moment μ associated with an orbiting electron lies along the same direction as the angular momentum $\mathbf{L}$ of that electron and is proportional to it. Thus we write

$$\boldsymbol{\mu} = \gamma \mathbf{I}. \tag{1.3}$$

where γ is a constant known as the gyromagnetic ratio. This relation between the magnetic moment and the angular momentum is demonstrated by the Einstein-de Haas effect, discovered in 1915, in which a ferromagnetic rod is suspended vertically. along its axis, by a thin fibre (see Fig. 1.2). It is initially at rest and unmagnetized, and is subsequently magnetized along its length by the application of a vertical magnetic field. This vertical magnetization is due to the alignment of the atomic magnetic moments and corresponds to a net angular momentum. To conserve total angular momentum, the rod begins turning about its axis in the opposite sense. If the angular momentum of the rod is measured, the angular momentum associated with the atomic magnetic moments, and hence the gyromagnetic ratio, can be deduced. The Einstein-de Haas effect is a rotation induced by magnetization, but there is also the reverse effect. known as the Barnett effect in which magnetization is
induced by rotation. Both phenomena demonstrate that magnetic moments are associated with angular momentum.

1.1.2 Precession

We now consider a magnetic moment μ in a magnetic field $\mathbf{B}$ as shown in Fig. 1.3. The energy E of the magnetic moment is given by

$$E = -\boldsymbol{\mu} \cdot \mathbf{B} \tag{1.4}$$

(see Appendix B) so that the energy is minimized when the magnetic moment lies along the magnetic field. There will be a torque $\mathbf{G}$ on the magnetic moment given by

$$\mathbf{G} = \mu \times \mathbf{B}, \tag{1.5}$$

(see Appendix B) which, if the magnetic moment were not associated with any angular momentum, would tend to turn the magnetic moment towards the magnetic field. [1]

However, since the magnetic moment is associated with the angular momentum $\mathbf{L}$ by eqn 1.3, and because torque is equal to rate of change of angular momentum, eqn 1.5 can be rewritten as

$$\frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}t} = \gamma \boldsymbol{\mu} \times \mathbf{B}. \tag{1.6}$$

This means that the change in μ is perpendicular to both μ and to $\mathbf{B}$. Rather than turning $\boldsymbol{\mu}$ towards $\mathbf{B}$, the magnetic field causes the direction of $\boldsymbol{\mu}$ to precess around $\mathbf{B}$. Equation 1.6

also implies that $|\boldsymbol{\mu}|$ is time-independent. Note that this situation is exactly analogous to the spinning of a gyroscope or a spinning top. [2]

In the following example, eqn 1.6 will be solved in detail for a particular case.

Example 1.1

Consider the case in which **B** is along the z direction and μ is initially at an angle of θ to **B** and in the xz plane (see Fig. 1.4). Then

$$\dot{\mu_x} = \gamma B \mu_y \tag{1.7}$$
$$\dot{\mu_y} = -\gamma B \mu_x \tag{1.8}$$
$$\dot{\mu_z} = 0, \tag{1.9}$$

so that μ_z is constant with time and μ_x and μ_y both oscillate. Solving these differential equations leads to

$$\mu_x(t) = |\boldsymbol{\mu}| \sin\theta \cos(\omega_{\mathrm{L}} t) \tag{1.10}$$
$$\mu_y(t) = |\boldsymbol{\mu}| \sin\theta \sin(\omega_{\mathrm{L}} t) \tag{1.11}$$
$$\mu_z(t) = |\boldsymbol{\mu}| \cos\theta \tag{1.12}$$

where

$$\omega_{\mathrm{L}} = \gamma B \tag{1.13}$$

is called the Larmor precession frequency.

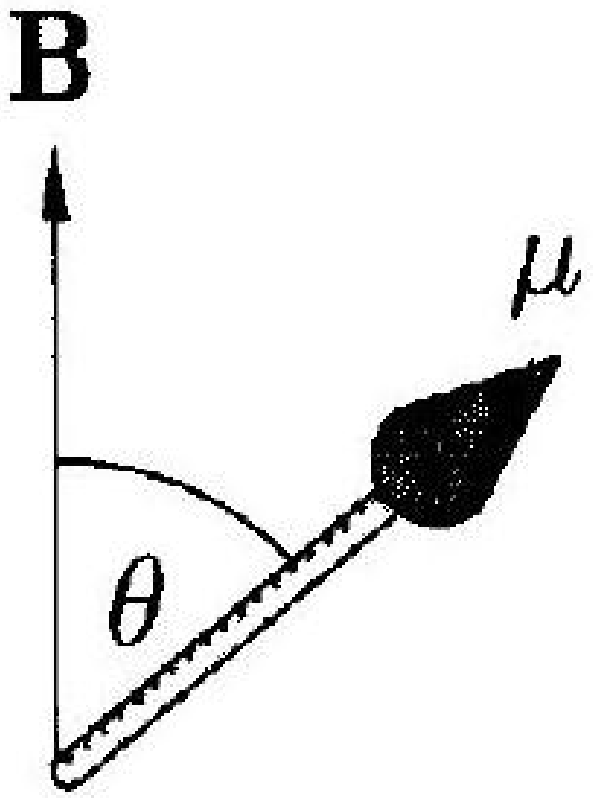


Fig. 1.3 A magnetic moment μ in a magnetic field $\mathbf{B}$ has an energy equal to $-\boldsymbol{\mu} \cdot \mathbf{B} = -\mu B \cos\theta$.}

[1] For an electric dipole $\mathbf{p}$, in an electric field $\boldsymbol{\varepsilon}$, the energy is $\boldsymbol{E} = -\mathbf{p} \cdot \boldsymbol{\varepsilon}$ and the torque is $\mathbf{G} = \mathbf{p} \times \boldsymbol{\varepsilon}$. A stationary electric dipole moment is just two separated stationary electric charges; it is not associated with any angular momentum, so if $\boldsymbol{\varepsilon}$ is not aligned with $\mathbf{p}$, the torque $\mathbf{G}$ will tend to turn $\mathbf{p}$ towards $\mathcal{E}$. A stationary magnetic moment is associated with angular momentum and so behaves differently.

[2] Imagine a top spinning with its axis inclined to the vertical. The weight of the top, acting downwards, exerts a (horizontal) torque on the top. If it were not spinning it would just fall over. But because it is spinning, it has angular momentum parallel to its spinning axis, and the torque causes the axis of the spinning top to move parallel to the torque, in a horizontal plane. The spinning top precesses.

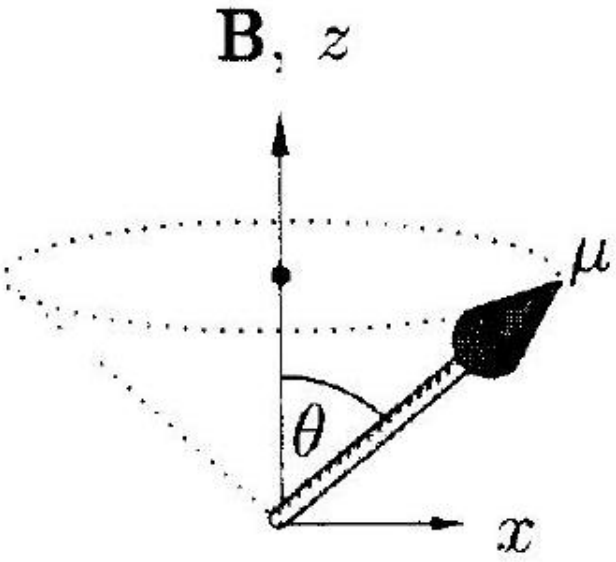


Fig. 1.4 A magnetic moment μ in a magnetic field $\mathbf{B
$ precesses around the magnetic field at the Larmor precession frequency, γB, where γ is the gyromagnetic ratio. The magnetic field **B** lies along the z-axis and the magnetic moment is initially in the xz-plane at an angle θ to B. The magnetic moment precesses around a cone of semi-angle θ.}

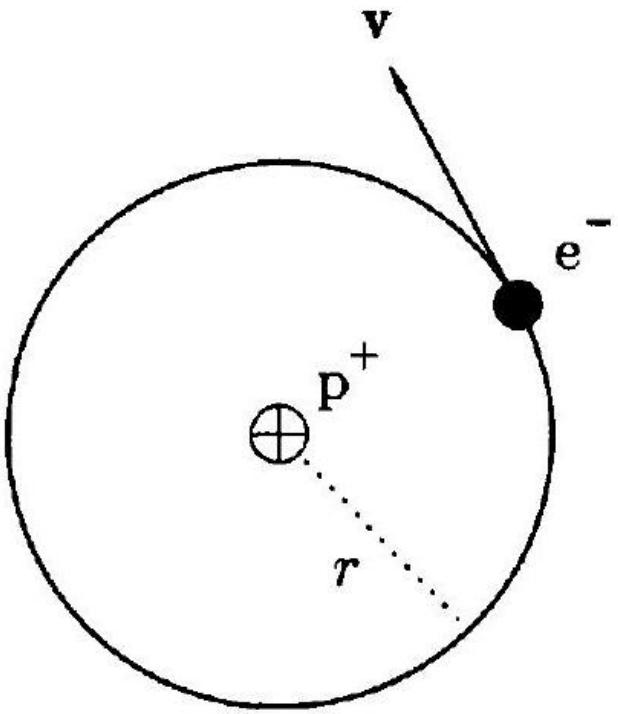


Fig. 1.5 An electron in a hydrogen atom orbiting with velocity $\mathbf{v
$ around the nucleus which consists of a single proton.}

Niels Bohr (1885-1962)

Note that the gyromagnetic ratio γ is the constant of proportionality which connects both the angular momentum with the magnetic moment (through eqn 1.3) and the precession frequency with the magnetic field (eqn 1.13). The phenomenon of precession hints at the subtlety

of what lies ahead: magnetic fields don't only cause moments to line up, but can induce a variety of dynamical effects.

1.1.3 The Bohr magneton

Before proceeding further, it is worth performing a quick calculation to estimate the size of atomic magnetic moments and thus deduce the size of the gyromagnetic ratio. Consider an electron (charge $-e$, mass $m_{\rm e}$) performing a circular orbit around the nucleus of a hydrogen atom, as shown in Fig. 1.5. The current I around the atom is $I = -e/\tau$ where $\tau = 2\pi r/v$ is the orbital period, $v = |\mathbf{v}|$ is the speed and r is the radius of the circular orbit. The magnitude of the angular momentum of the electron, $m_{\rm e}vr$, must equal $\hbar$ in the ground state so that the magnetic moment of the electron is

$$\mu = \pi r^2 I = -\frac{e\hbar}{2m_{\rm e}} \equiv -\mu_{\rm B} \tag{1.14}$$

where μ_B is the Bohr magneton, defined by

$$\mu_{\rm B} = \frac{e\hbar}{2m_{\rm e}}. \tag{1.15}$$

This is a convenient unit for describing the size of atomic magnetic moments and takes the value $9.274 \times 10^{-24}\mathrm{Am}^2$. Note that sign of the magnetic moment in eqn 1.14 is negative. Because of the negative charge of the electron, its magnetic moment is antiparallel to its angular momentum. The gyromagnetic ratio for the electron is $\gamma = -e/2m_{\rm e}$. The Larmor frequency is then $\omega_{\rm L} = |\gamma|B = eB/2m_{\rm e}$.

1.1.4 Magnetization and field

A magnetic solid consists of a large number of atoms with magnetic moments. The magnetization $\mathbf{M}$ is defined as the magnetic moment per unit volume. Usually this vector quantity is considered in the 'continuum approximation', i.e. on a lengthscale large enough so that one does not see the graininess due to the individual atomic magnetic moments. Hence $\mathbf{M}$ can be considered to be a smooth vector field, continuous everywhere except at the edges of the magnetic solid.

In free space (vacuum) there is no magnetization. The magnetic field can be described by the vector fields $\mathbf{B}$ and $\mathbf{H}$ which are linearly related by

$$\mathbf{B} = \mu_0 \mathbf{H}, \tag{1.16}$$

where $\mu_0 = 4\pi \times 10^{-7}\mathrm{Hm}^{-1}$ is the permeability of free space. The two magnetic fields $\mathbf{B}$ and $\mathbf{H}$ are just scaled versions of each other, the former measured in Tesla (abbreviated to T) and the latter measured in A m $^{-1}$.

In a magnetic solid the relation between $\mathbf{B}$ and $\mathbf{H}$ is more complicated and the two vector fields may be very different in magnitude and direction. The general vector relationship is

$$\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M}). \tag{1.17}$$

In the special case that the magnetization $\mathbf{M}$ is linearly related to the magnetic field $\mathbf{H}$, the solid is called a linear material, and we write

$$\mathbf{M} = \chi \mathbf{H}, \tag{1.18}$$

where χ is a dimensionless quantity called the magnetic susceptibility. In this special case there is still a linear relationship between $\mathbf{B}$ and $\mathbf{H}$, namely

$$\mathbf{B} = \mu_0(1+\chi)\mathbf{H} = \mu_0\mu_\mathrm{r}\mathbf{H} \tag{1.19}$$

where $\mu_\mathrm{r} = 1 + \chi$ is the relative permeability of the material.
A cautionary tale now follows. This arises because we have to be very careful in defining fields in magnetizable media. Consider a region of free space with an applied magnetic field given by fields $\mathbf{B}_a$ and $\mathbf{H}_a$, connected by $\mathbf{B}_a = \mu_0\mathbf{H}_a$. So far, everything is simple. Now insert a magnetic solid into that region of free space. The internal fields inside the solid, given by $\mathbf{B}_\mathrm{i}$ and $\mathbf{H}_\mathrm{i}$ can be very different from $\mathbf{B}_a$ and $\mathbf{H}_a$ respectively. This difference is because of the magnetic field produced by all magnetic moments in the solid. In fact $\mathbf{B}_\mathrm{i}$ and $\mathbf{H}_i$ can both depend on the position inside the magnetic solid at which you measure them. [3] This is true except in the special case of an ellipsoidal shaped sample (see Fig. 1.6). If the magnetic field is applied along one of the principal axes of the ellipsoid, then throughout the sample

$$\mathbf{H_i} = \mathbf{H}_\mathrm{a} - N\mathbf{M}, \tag{1.20}$$

where N is the appropriate demagnetizing factor (see Appendix D). The 'correction term' $\mathbf{H}_\mathrm{d} = -N\mathbf{M}$, which you need to add to $\mathbf{H}_\mathrm{a}$ to get $\mathbf{H}_\mathrm{i}$, is called the demagnetizing field. Similarly

$$\mathbf{B_i} = \mu_0\left(\mathbf{H_i} + \mathbf{M}\right) = \mathbf{B}_\mathrm{a} + \mu_0(1-N)\mathbf{M} \tag{1.21}$$

Example 1.2

For a spherically shaped sample, $N = \frac{1}{3}$ and so the internal fields inside the sphere are

$$\mathbf{H_i} = \mathbf{H}_a - \frac{\mathbf{M}}{3}, \tag{1.22}$$

$$\mathbf{B_i} = \mathbf{B}_a + \frac{2\mu_0\mathbf{M}}{3}. \tag{1.23}$$

For historical reasons, standard convention dictates that $\mathbf{B}$ is called the magnetic induction or mognetic flux density and $\mathbf{H}$ is called the magnetic field strength. However, such terms are cumbersome and can be misleading. Following common usage, we refer to both simply as the magnetic field. The letters ' B ' and ' H ' will show which one is meant.
[3] A magnetized sample will also influence the magnetic field outside it, as well as inside it (considered here), as you may know from playing with a bar magnet and iron filings.

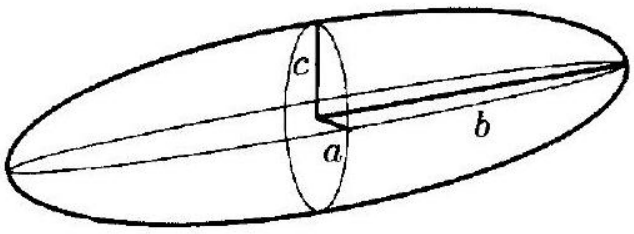


Fig. 1.6 An ellipsoidal shaped sample of a magnetized solid with principal axes a, b and c. This includes the special cases of a sphere ($a = b = c$) and a flat plate ($a, b \to \infty$, $c = 0$).

[4] In accurate experimental work on even these materials, demagnetizing fields must still be considered.
[5] See Section 6.7 for more on magnetic domains.

When the magnetization is large compared to the applied field $|\mathbf{H}_\mathrm{a}| = |\mathbf{B}_\mathrm{a}|/\mu_0$ (measured before the sample was inserted) these demagnetizing corrections need to be taken seriously. However, it is possible to sweep these complications under the carpet for the special case of weak magnetism. For a linear material with $\chi \ll 1$, we have that $M \ll H, H_\mathrm{i} \approx H_\mathrm{a}$ and $B_\mathrm{i} \approx \mu_0 H_\mathrm{i}$. We can then get away with imagining that the magnetic field in the material is the same as the magnetic field that we apply. This approximation will be used in Chapters 2 and 3 concerning the relatively weak effects of diamagnetism. [4] In ferromagnets, demagnetizing effects are always significant.

Example 1.3

The intrinsic magnetic susceptibility of a material is

$$\chi_\mathrm{intrinsic} = \frac{M}{H_\mathrm{i}} \tag{1.24}$$

This intrinsic material property is not what you measure experimentally. This is because you measure the magnetization M in response to an applied field H_a. You therefore measure

$$\chi_\mathrm{experimental} = \frac{M}{H_\mathrm{a}}. \tag{1.25}$$

The two quantities can be related by

$$\chi_\mathrm{experimental} = \frac{M}{H_\mathrm{i} + NM} = \frac{M/H_\mathrm{i}}{1 + NM/H_\mathrm{i}} = \frac{\chi_\mathrm{intrinsic}}{1 + N\chi_\mathrm{intrinsic}}. \tag{1.26}$$

When $\chi_\mathrm{intrinsic} \ll 1$, the distinction between $\chi_\mathrm{intrinsic}$ and $\chi_\mathrm{experimental}$ is academic. When $\chi_\mathrm{intrinsic}$ is closer or above 1 , the distinction can be very important. For example, in a ferromagnet approaching the Curie temperature from above (see Chapter 4), $\chi_\mathrm{intrinsic} \to \infty$, but $\chi_\mathrm{experimental} \to 1/N$.

One last word of warning at this stage: a ferromagnetic material may have no net magnetic moment because it consists of magnetic domains. [5] In each domain there is a uniform magnetization, but the magnetization of each domain points in a different direction from its neighbours. Therefore a sample may appear not to be magnetized, even though on a small enough scale, all the magnetic moments are locally aligned.

In the rest of this chapter we will consider some further aspects of magnetic moments that relate to classical mechanics (in Section 1.2) and quantum mechanics (in Section 1.3).

9.1.2 1.2 Classical mechanics and magnetic moments

In this section, we describe the effect of an applied magnetic field on a system of charges using purely classical arguments. First, we consider the effect on a single charge and then use this result to evaluate the magnetization of a system of charges. A summary of some important results in electromagnetism may be found in Appendix B.

1.2.1 Canonical momentum

In classical mechanics the force $\mathbf{F}$ on a particle with charge q moving with velocity $\mathbf{v}$ in an electric field $\boldsymbol{\varepsilon}$ and magnetic field $\mathbf{B}$ is

$$\mathbf{F} = q(\boldsymbol{\varepsilon} + \mathbf{v} \times \mathbf{B}) \tag{1.27}$$

and is called the Lorentz force. With this familiar equation, one can show how the momentum of a charged particle in a magnetic field is modified. Using $\mathbf{F} = m\mathrm{d}\mathrm{v}/\mathrm{d}t, \mathbf{B} = \nabla \times \mathbf{A}$ and $\boldsymbol{\varepsilon} = -\nabla V - \partial\mathbf{A}/\partial t$, where V is the electric potential, $\mathbf{A}$ is the magnetic vector potential and m is the mass of the particle, eqn 1.27 may be rewritten as

$$m\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -q\nabla V - q\frac{\partial \mathbf{A}}{\partial t} + q\mathbf{v} \times (\nabla \times \mathbf{A}) \tag{1.28}$$

The vector identity

$$\mathbf{v} \times (\nabla \times \mathbf{A}) = \nabla(\mathbf{v} \cdot \mathbf{A}) - (\mathbf{v} \cdot \nabla)\mathbf{A} \tag{1.29}$$

can be used to simplify eqn 1.28 leading to

$$m\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} + q\left(\frac{\partial \mathbf{A}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{A}\right) = -q\nabla(V - \mathbf{v} \cdot \mathbf{A}) \tag{1.30}$$

Note that $m\mathrm{d}\mathrm{v}/\mathrm{d}t$ is the force on a charged particle measured in a coordinate system that moves with the particle. The partial derivative $\partial\mathbf{A}/\partial t$ measures the rate of change of $\mathbf{A}$ at a fixed point in space. We can rewrite eqn 1.30 as

$$\frac{\mathbf{d}}{\mathrm{d}t}(m\mathbf{v} + q\mathbf{A}) = -q\nabla(V - \mathbf{v} \cdot \mathbf{A}) \tag{1.31}$$

where $\mathrm{d}\mathbf{A}/\mathrm{d}t$ is the convective derivative of $\mathbf{A}$, written as

$$\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} = \frac{\partial \mathbf{A}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{A}, \tag{1.32}$$

which measures the rate of change of $\mathbf{A}$ at the location of the moving particle. Equation 1.31 takes the form of Newton's second law (i.e. it reads 'the rate of change of a quantity that looks like momentum is equal to the gradient of a quantity that looks like potential energy') and therefore motivates the definition of the canonical momentum

$$\mathbf{p} = m\mathbf{v} + q\mathbf{A} \tag{1.33}$$

and an effective potential energy experienced by the charged particle, $q(V - \mathbf{v} \cdot \mathbf{A})$, which is velocity-dependent. The canonical momentum reverts to the familiar momentum $m\mathbf{v}$ in the case of no magnetic field, $\mathbf{A} = 0$. The kinetic energy remains equal to $\frac{1}{2}mv^2$ and this can therefore be written in terms of the canonical momentum as $(\mathbf{p} - q\mathbf{A})^2/2m$. This result will be used below, and also later in the book where the quantum mechanical operator associated with kinetic energy in a magnetic field is written $(-\mathrm{i}\hbar\nabla - q\ \mathrm{A})^2/2m$.

Hendrik Lorentz (1853-1928)

See Appendix G for a list of vector identities. Note also that $\mathbf{v}$ does not vary with position. [6] See also Appendix E.

Niels Bohr (1885-1962)

Hendreka J. van Leeuwen (1887-1974)

Ludwig Boltzmann (1844-1906)

1.2.2 The Bohr-van Leeuwen theorem

The next step is to calculate the net magnetic moment of a system of electrons in a solid. Thus we want to find the magnetization, the magnetic moment per unit volume, that is induced by the magnetic field. From eqn 1.4, the magnetization is proportional to the rate of change of energy of the system with applied magnetic field. [6] Now, eqn 1.27 shows that the effect of a magnetic field is always to produce forces on charged particles which are perpendicular to their

velocities. Thus no work is done and therefore the energy of a system cannot depend on the applied magnetic field. If the energy of the system does not depend on the applied magnetic field, then there can be no magnetization.
This idea is enshrined in the Bohr-van Leeuwen theorem which states that in a classical system there is no thermal equilibrium magnetization. We can prove this in outline as follows: in classical statistical mechanics the partition function Z for N particles, each with charge q, is proportional to

$$\iint \cdots \int \exp\left(-\beta E\left(\{\mathbf{r}_i, \mathbf{p}_i\}\right) \mathrm{d}\mathbf{r}_1 \cdots \mathrm{d}\mathbf{r}_N \, \mathrm{d}\mathbf{p}_1 \cdots \mathrm{d}\mathbf{p}_N \right. \tag{1.34}$$

where $\beta = 1/k_{\mathrm{B}}T$, k_{B} is the Boltzmann factor, T is the temperature, and $i = 1, \ldots, N$. Here $E\left(\{\mathbf{r}_i, \mathbf{p}_i\}\right)$ is the energy associated with the N charged particles having positions $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N$, and momenta $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N$. The integral is therefore over a $6N$-dimensional phase space ($3N$ position coordinates, $3N$ momentum coordinates). The effect of a magnetic field, as shown in the preceding section, is to shift the momentum of each particle by an amount $q\mathbf{A}$. We must therefore replace $\mathbf{p}_i$ by $\mathbf{p}_i - q\mathbf{A}$. The limits of the momentum integrals go from $-\infty$ to ∞ so this shift can be absorbed by shifting the origin of the momentum integrations. Hence the partition function is not a function of magnetic field, and so neither is the free energy $F = -k_{\mathrm{B}}T \log Z$ (see Appendix E). Thus the magnetization must be zero in a classical system.

This result seems rather surprising at first sight. When there is no applied magnetic field, electrons go in straight lines, but with an applied magnetic field their paths are curved (actually helical) and perform cyclotron orbits. One is tempted to argue that the curved cyclotron orbits, which are all curved in the same sense, must contribute to a net magnetic moment and hence there should be an effect on the energy due to an applied magnetic field. But the fallacy of this argument can be understood with reference to Fig. 1.7, which shows the orbits of electrons in a classical system due to the applied magnetic field. Electrons do indeed perform cyclotron orbits which must correspond to a net magnetic moment. Summing up these orbits leads to a net anticlockwise circulation of current around the edge of the system (as in Fig. 1.1). However, electrons near the surface cannot perform complete loops and instead make repeated elastic collisions with the surface, and perform so-called skipping orbits around the sample perimeter. The anticlockwise current due to the bulk electrons precisely cancels out with the clockwise current associated with the skipping orbits of electrons that reflect or scatter at the surface.

The Bohr-van Leeuwen theorem therefore appears to be correct, but it is at odds with experiment: lots of real systems containing electrons do have a net magnetization. Therefore the assumptions that went into the theorem must be in doubt. The assumptions are classical mechanics! Hence we conclude that classical mechanics is insufficient to explain this most basic property of

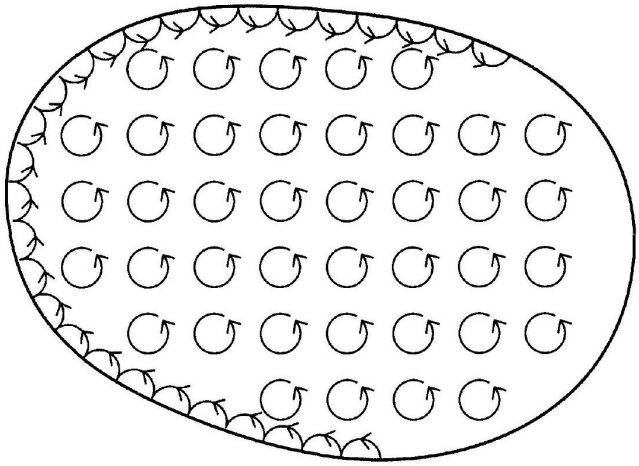


Fig. 1.7 Electrons in a classical system with an applied magnetic field undergo cyclotron orbits in the bulk of the system. These orbits here precess in an anti-clockwise sense. They contribute a net orbital current in an anticlockwise sense (see Fig. 1.1). This net current precisely cancels out with the current due to the skipping orbits associated with electrons which scatter at the surface and precess in a clockwise sense around the sample.

magnetic materials, and we cannot avoid using quantum theory to account for the magnetic properties of real materials. In the next section we will consider the quantum mechanics of electrons in some detail.

9.1.3 1.3 Quantum mechanics of spin

In this section I will briefly review some results concerning the quantum mechanics of electron spin. A fuller account of the quantum mechanics of angular momentum may be found in the further reading given at the end of the chapter. Some results connected with quantum and atomic physics are also given in Appendix C.

1.3.1 Orbital and spin angular momentum

The electronic angular momentum discussed in Section 1.1 is associated with the orbital motion of an electron around the nucleus and is known as the orbital angular momentum. In a real atom it depends on the electronic state occupied by the electron. With quantum numbers l and m_l defined in the usual way (see Appendix C) the component of orbital angular momentum along a fixed axis (in this case the z axis) is $m_l\hbar$ and the magnitude [7] of the orbital angular momentum is $\sqrt{l(l+1)}\hbar$. Hence the component of magnetic moment along the z axis is $-m_l\mu_{\rm B}$ and the magnitude of the total magnetic dipole moment is $\sqrt{l(l+1)}\mu_{\rm B}$.

The situation is further complicated by the fact that an electron possesses an intrinsic magnetic moment which is associated with an intrinsic angular momentum. The intrinsic angular momentum of an electron is called spin. It is so termed because electrons were once thought to precess about their own axes, but since an electron is a point particle this is rather hard to imagine.

[7] Strictly, it is the square of the angular momentum and the square of the magnetic dipole moment which are well defined quantities. The operator $\hat{\mathbf{L}}^2$ has eigenvalue $l(l+1)\,\hbar^2$ and $\hat{L}_z$ has eigenvalue $m_l\hbar$. Similarly, the operator $\hat{\mu}^2$ has eigenvalue $l(l+1)\mu_{\mathbf{B}}^2$ and $\hat{\mu}_z$ has eigenvalue $-m_l\mu_{\rm B}$.

[8] Strictly, the eigenvalue of the operator $\hat{\mathbf{S}}^2$ is $s(s+1)h^2$.

The g-factor is discussed in more detail in Appendix C.6.

[9] The $\mp$ sign is this way up because the magnetic moment is antiparallel to the angular momentum. This arises because of the negative charge of the electron. When $m_s = +\frac{1}{2}$ the moment is $-\mu_{\rm B}$. When $m_s = -\frac{1}{2}$ the moment is $+\mu_{\rm B}$.

Pieter Zeeman (1865-1943)

Wolfgang Pauli (1900-1958)

The concept has changed but the name has stuck. This is not such a bad thing because electron spin behaves so counterintuitively that it would be hard to find any word that could do it full justice!

The spin of an electron is characterized by a spin quantum number s, which for an electron takes the value of $\frac{1}{2}$. The value of any component of the angular momentum can only take one of $2s+1$ possible values, namely: $s\hbar, (s-1)\hbar, \ldots, -s\hbar$. The component of spin angular momentum is written $m_s\hbar$. For an electron, with $s = \frac{1}{2}$, this means only two possible values so that $m_s = \pm\frac{1}{2}$. The component of angular momentum along a particular axis is then $\hbar/2$ or $-\hbar/2$. These alternatives will be referred to as 'up' and 'down' respectively. The magnitude [8] of the spin angular momentum for an electron is $\sqrt{s(s+1)}\hbar = \sqrt{3}\hbar/2$.

The spin angular momentum is then associated with a magnetic moment which can have a component along a particular axis equal to $-g\mu_{\rm B}m_s$ and a magnitude equal to $\sqrt{s(s+1)}g\mu_{\rm B} = \sqrt{3}g\mu_{\rm B}/2$. In these expressions, g is a constant known as the g -factor. The g -factor takes a value of approximately 2 , so that the component of the intrinsic magnetic moment of the electron along the z axis is [9] $\approx \mp\mu_{\mathbf{B}}$, even though the spin is half-integral. The energy of the electron in a magnetic field B is therefore

$$E = g\mu_{\rm B}m_sB. \tag{1.35}$$

The energy levels of an electron therefore split in a magnetic field by an amount $g\mu_{\rm B}B$. This is called Zeeman splitting.

In general for electrons in atoms there may be both orbital and spin angular momenta which combine. The g-factor can therefore take different values in real atoms depending on the relative contributions of spin and orbital angular momenta. We will return to this point in the next chapter.

The angular momentum of an electron is always an integral or half-integral multiple of $\hbar$. Therefore it is convenient to drop the factor of $\hbar$ in expressions for angular momentum operators, which amounts to saying that these operators measure the angular momentum in units of $\hbar$. In the rest of this book we will define angular momentum operators, like $\hat{\mathbf{L}}$, such that the angular momentum is $\hbar\hat{\mathbf{L}}$. This simplifies expressions which appear later in the book.

1.3.2 Pauli spin matrices and spinors

The behaviour of the electron spin turns out to be connected to a rather strange algebra, based on the three Pauli spin matrices, which are defined as

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.36}$$

It will be convenient to think of these as a vector of matrices,

$$\boldsymbol{\sigma} = (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)\,. \tag{1.37}$$

Before proceeding, we recall a few results which can be proved straightforwardly by direct substitution. Let

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \tag{1.38}$$

be a three-component vector. Then $\sigma \cdot \mathbf{a}$ is a matrix given by

$$\boldsymbol{\sigma} \cdot \mathbf{a} = \begin{pmatrix} a_3 & a_1 - \mathrm{i}a_2 \\ a_1 + \mathrm{i}a_2 & -a_3 \end{pmatrix}. \tag{1.39}$$

Such matrices can be multiplied together, leading to results such as

$$(\boldsymbol{\sigma} \cdot \mathbf{a})(\boldsymbol{\sigma} \cdot \mathbf{b}) = \mathbf{a} \cdot \mathbf{b} + \mathrm{i}\boldsymbol{\sigma} \cdot (\mathbf{a} \times \mathbf{b}) \tag{1.40}$$

and

$$(\sigma \cdot \mathbf{a})^2 = |\mathbf{a}|^2. \tag{1.41}$$

We now define the spin angular momentum operator by

$$\hat{\mathbf{S}} = \frac{1}{2}\hat{\sigma} \tag{1.42}$$

so that

$$\hat{S}_x = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{S}_y = \frac{1}{2}\begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \hat{S}_z = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.43}$$

Notice again that we are using the convention that angular momentum is measured in units of $\hbar$, so that the angular momentum associated with an electron is actually $\hbar\mathbf{S}$. (Note that some books choose to define $\hat{\mathbf{S}}$ such that $\hat{\mathbf{S}} = \hbar\hat{\boldsymbol{\sigma}}/2$.)

It is only the operator $\hat{S}_z$ which is diagonal and therefore if the electron spin points along the z-direction the representation is particularly simple. The eigenvalues of $\hat{S}_z$, which we will give the symbol m_s, take values $m_s = \pm\frac{1}{2}$ and the corresponding eigenstates are $|\uparrow_z\rangle$ and $|\downarrow_z\rangle$ where

$$|\uparrow_z\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{1.44}$$

$$|\downarrow_z\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{1.45}$$

and correspond to the spin pointing parallel or antiparallel to the z axis respectively. (The 'bra and ket' notation, i.e. writing states in the form $|\psi\rangle$ is reviewed in Appendix C.) Hence

$$\hat{S}_z\,|\uparrow_z\rangle = \frac{1}{2}\,|\uparrow_z\rangle\,, \quad \hat{S}_z\,|\downarrow_z\rangle = -\frac{1}{2}\,|\downarrow_z\rangle\,. \tag{1.46}$$

The eigenstates corresponding to the spin pointing parallel or antiparallel to the x - and y-axes are

$$\begin{array}{ll} |\uparrow_x\rangle = \frac{1}{\sqrt{2}}\binom{1}{1}, & |\downarrow_x\rangle = \frac{1}{\sqrt{2}}\binom{1}{-1} \\ |\uparrow_y\rangle = \frac{1}{\sqrt{2}}\binom{1}{i}, & |\downarrow_y\rangle = \frac{1}{\sqrt{2}}\binom{1}{-i}. \end{array} \tag{1.48}$$

This two-component representation of the spin wave functions is known as a spinor representation and the states are referred to as spinors. A general state can be written

$$|\psi\rangle = \begin{pmatrix} a \\ b \end{pmatrix} = a\,|\uparrow_z\rangle + b\,|\downarrow_z\rangle \tag{1.49}$$

where a and b are complex numbers [10] and it is conventional to normalize the state so that

$$|a|^2 + |b|^2 = 1 \tag{1.50}$$

Note that all the terms in eqns 1.40 and 1.41 are matrices. The term $\mathbf{a}\cdot\mathbf{b}$ is shorthand for $\mathbf{a}\cdot$ b $\mathbf{I}$ where $\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is the identity matrix. Similarly $|\mathbf{a}|^2$ is shorthand for $|\mathbf{a}|^2\mathbf{1}$.

[10] They could of course be functions of position in a general case.

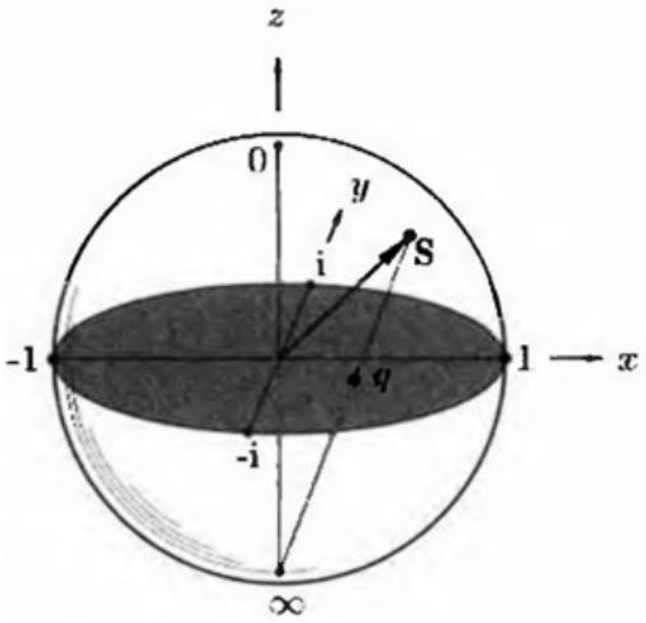


Fig. 1.8 The Riemann sphere represents the spin states of a spin- $\frac{1}{2}$ particle. The spin vector S lies on a unit sphere. A line from the south pole of the sphere to S cuts the horizontal equitorial plane (shaded) at $y = x + \mathrm{i}y$ where the horizontal plane is considered as an Argand diagram. The numerical value of the complex number q is shown for six cases. namely S parallel or antiparallel to the x, y and i axes.}

George F. B. Riemann (1826-1866)

The total spin angular momentum operator $\hat{\mathbf{S}}$ is defined by

$$\hat{\mathbf{S}} = \begin{pmatrix} \hat{S}_x \\ \hat{S}_y \\ \hat{S}_z \end{pmatrix} = \mathbf{i}\hat{S}_x + \mathbf{j}\hat{S}_y + \mathbf{k}\hat{S}_z. \tag{1.51}$$

where $\mathbf{i}, \mathbf{j}$ and $\mathbf{k}$ are the unit cartesian vectors. The operator $\hat{\mathbf{S}}^2$ is then given by

$$\hat{\mathbf{S}}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2. \tag{1.52}$$

Since the eigenvalues of $\hat{S}_x^2, \hat{S}_y^2$ or $\hat{S}_z^2$ are always $\frac{1}{4} = \left(\pm\frac{1}{2}\right)^2$, we have the result that for any spin state $|\psi\rangle$

$$\hat{\mathbf{S}}^2|\psi\rangle = \left(\hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2\right)|\psi\rangle = \left(\frac{1}{4} + \frac{1}{4} + \frac{1}{4}\right)|\psi\rangle = \frac{3}{4}|\psi\rangle. \tag{1.53}$$

Many of these results can be generalized to the case of particles with spin quantum number $s > \frac{1}{2}$. The most important result is that the eigenvalue of $\hat{\mathbf{S}}^2$ becomes $s(s+1)$. In the case of $s = \frac{1}{2}$ which we are considering in this chapter, $s(s+1) = \frac{3}{4}$, in agreement with eqn 1.53 . The commutation relation hetween the spin operators is

$$\left[\hat{S}_x, \hat{S}_y\right] = \mathrm{i}\hat{S}_z \tag{1.54}$$

and cyclic permutations thereof. This can be proved very simply using eqns 1.40 and 1.42 . Each of these operators commutes with $\hat{\mathbf{S}}^2$ so that

$$\left[\hat{\mathbf{S}}^2, \hat{S}_z\right] = 0. \tag{1.55}$$

Thus it is possible simultaneously to know the total spin and one of its components, but it is not possible to know more than one of the components simultaneously.

A useful geometric construction that can aid thinking about spin is shown in Fig. 1.8. The spin vector $\mathbf{S}$ points in three-dimensional space. Because the quantum states are normalized, S lies on the unit sphere. Draw a line from the end of the vector $\mathbf{S}$ to the south pole of the sphere and observe the point, q, at which this line intersects the horizontal plane (shown shaded in Fig. 1.8). Treat this horizontal plane as an Argand diagram, with the x axis as the real axis and the y axis as the imaginary axis. Hence $q = x + iy$ is a complex number. Then the spinor representation of $\mathbf{S}$ is $\binom{1}{q}$, which when normalized is

$$\mathbf{S} = \frac{1}{\sqrt{1+|q|^2}}\binom{1}{q}. \tag{1.56}$$

In this representation the sphere is known as the Riemann sphere.

1.3.3 Raising and lowering operators

The raising and lowering operators $\hat{S}_+$and $\hat{S}_-$are defined by

$$\begin{aligned} \hat{S}_+ &= \hat{S}_x + i\hat{S}_y \\ \hat{S}_- &= \hat{S}_x - i\hat{S}_y. \end{aligned} \tag{1.57}$$

For an operator $\hat{A}$ to be Hermitian, one must have that $\hat{A}^\dagger = \hat{A}$ where † implies an adjoint operation (for matrices this means 'take the transpose and then complex conjugate each element'). The raising and lowering operators are not Hermitian (because $\hat{S}_+^\dagger = \hat{S}_-$and $\hat{S}_-^\dagger = \hat{S}_+$) and therefore they do not correspond to observable quantities. They are nevertheless very useful. Straightforward application of eqns 1.54 and 1.57 yields the following commutation relations :

$$\left[\hat{S}_+, \hat{S}_-\right] = 2\hat{S}_z \tag{1.58}$$

$$\left[\hat{S}_z, \hat{S}_\pm\right] = \pm\hat{S}_\pm \tag{1.59}$$

and

$$\left[\hat{\mathbf{S}}^2, \hat{S}_\pm\right] = 0 \tag{1.60}$$

Another useful relation, proven by direct substitution is

$$\hat{S}_+\hat{S}_- + \hat{S}_-\hat{S}_+ = 2\left(\hat{S}_x^2 + \hat{S}_y^2\right) \tag{1.61}$$

and this provides a convenient representation for $\hat{\mathbf{S}}^2$, namely

$$\hat{\mathbf{S}}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2 \tag{1.62}$$

$$= \frac{1}{2}\left(\hat{S}_+\hat{S}_- + \hat{S}_-\hat{S}_+\right) + \hat{S}_z^2 \tag{1.63}$$

Expressed as matrices the raising and lowering operators are

$$\hat{S}_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{1.64}$$

$$\hat{S}_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{1.65}$$

and using eqns $1.43, 1.63, 1.64$ and 1.65 this then yields

$$\hat{\mathbf{S}}^2 = \frac{3}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{1.66}$$

in agreement with eqn 1.53 .

1.3.4 The coupling of two spins

Now consider two spin- $\frac{1}{2}$ particles coupled by an interaction described by a Hamiltonian $\hat{\mathcal{H}}$ given by [11]

$$\hat{\mathcal{H}} = A\hat{\mathbf{S}}^a \cdot \hat{\mathbf{S}}^b, \tag{1.67}$$

where $\hat{\mathbf{S}}^a$ and $\hat{\mathbf{S}}^b$ are the operators for the spins for the two particles. Considered as a joint entity, the total spin can also be represented by an operator:

$$\hat{\mathbf{S}}^{\text{tot}} = \hat{\mathbf{S}}^a + \hat{\mathbf{S}}^b \tag{1.68}$$

so that

$$\left(\hat{\mathbf{S}}^{\text{tot}}\right)^2 = \left(\hat{\mathbf{S}}^a\right)^2 + \left(\hat{\mathbf{S}}^b\right)^2 + 2\hat{\mathbf{S}}^a \cdot \hat{\mathbf{S}}^b \tag{1.69}$$

Combining two spin $-\frac{1}{2}$ particles results in a joint entity with spin quantum number $s = 0$ or 1 . The eigenvalue of $\left(\hat{\mathbf{S}}^{\text{tot}}\right)^2$ is $s(s+1)$ which is therefore

The raising and lowering operators get their name from their effect on spin states. You can show directly that

$$\hat{S}_+ \left|\uparrow_z\right\rangle = 0$$
$$\hat{S}_+ \left|\downarrow_z\right\rangle = \left|\uparrow_z\right\rangle$$
$$\hat{S}_- \left|\uparrow_z\right\rangle = 0$$
$$\hat{S}_- \left|\downarrow_z\right\rangle = \left|\downarrow_z\right\rangle$$

So a raising operator will raise the z component of the spin angular momentum by $\hbar$, a lowering operator will lower the z component of the spin angular momentum by h. If the z

component of the spin angular momentum is already at its maximum (minimum) level, $\hat{S}_+$ ($\hat{S}_-$) will just annihilate the state.

[11] The type of interaction in eqn 1.67 will turn out to be very important in this book. The hyperfine interaction (see Chapter 2) and the Heisenberg exchange interaction (see Chapter 4) both take this form.

Table 1.1 The eigenstates of $\hat{\mathbf{S}}^a \cdot \hat{\mathbf{S}}^b$ and the corresponding values of m_s, s and the eigenvalue of $\hat{\mathbf{S}}^a \cdot \hat{\mathbf{S}}^b$.

Eigenstate	m_s	s	$\hat{\mathbf{S}}^a \cdot \hat{\mathbf{S}}^b$
$\vert\uparrow\uparrow\rangle$	1	1	$\frac{1}{4}$
$\frac{\vert\uparrow\downarrow\rangle+\vert\downarrow\uparrow\rangle}{\sqrt{2}}$	0	1	$\frac{1}{4}$
$\vert\downarrow\downarrow\rangle$	-1	1	$\frac{1}{4}$
$\frac{\vert\uparrow\downarrow\rangle-\vert\downarrow\uparrow\rangle}{\sqrt{2}}$	0	0	$-\frac{3}{4}$

either 0 or 2 for the cases of $s = 0$ or 1 respectively. The eigenvalues of both $\left(\hat{\mathbf{S}}^a\right)^2$ and $\left(\hat{\mathbf{S}}^b\right)^2$ are $\frac{3}{4}$ from eqn 1.53. Hence from eqn 1.69

$$\hat{\mathbf{S}}^a \cdot \hat{\mathbf{S}}^b = \begin{cases} \frac{1}{4} & \text{if} \quad s = 1 \\ -\frac{3}{4} & \text{if} \quad s = 0. \end{cases} \tag{1.70}$$

Because the Hamiltonian is $\hat{\mathcal{H}} = A\hat{\mathbf{S}}^a \cdot \hat{\mathbf{S}}^b$, the system therefore has two energy levels for $s = 0$ and 1 with energies given by

$$E = \begin{cases} \frac{A}{4} & \text{if} \quad s = 1 \\ -\frac{3A}{4} & \text{if} \quad s = 0. \end{cases} \tag{1.71}$$

The degeneracy of each state is given by $2s + 1$, hence the $s = 0$ state is a singlet and the $s = 1$ state is a triplet. The z component of the spin of this state, m_s, takes the value 0 for the singlet, and one of the three values $-1, 0, 1$ for the triplet.

Equation 1.70 has listed the eigenvalues of $\hat{\mathbf{S}}^a \cdot \hat{\mathbf{S}}^b$, but it is also useful to describe the eigenstates. Let us first consider the following basis:

$$\vert\uparrow\uparrow\rangle, \quad \vert\uparrow\downarrow\rangle, \quad \vert\downarrow\uparrow\rangle, \quad \vert\downarrow\downarrow\rangle. \tag{1.72}$$

In this representation the first arrow refers to the z component of the spin labelled a and the second arrow refers to the z component of the spin labelled b. The eigenstates of $\hat{\mathbf{S}}^a \cdot \hat{\mathbf{S}}^b$ are linear combinations of these basis states and are listed in Table 1.1. The calculation of these eigenstates is treated in Exercise 1.9. Notice that m_s is equal to the sum of the z components of the individual spins. Also, because the eigenstates are a mixture of states in the original basis, it is not possible in general to know both the z components of the original spins and the total spin of the resultant entity. This is a general feature which will become more important in more complicated situations.

Our basis in eqn 1.72 was unsatisfactory from another point of view: the wave function must be antisymmetric with respect to exchange of the two electrons. Now the wave function is a product of a spatial function $\psi_{\text{space}}(\mathbf{r}_1, \mathbf{r}_2)$ and the spin function χ, where χ is a linear combination of the states listed in eqn 1.72. The spatial wave function can be either symmetric or antisymmetric with respect to exchange of electrons. For example, the spatial wave function

$$\psi_{\text{space}}(\mathbf{r}_1, \mathbf{r}_2) = \frac{\phi(\mathbf{r}_1)\,\xi(\mathbf{r}_2) \pm \phi(\mathbf{r}_2)\,\xi(\mathbf{r}_1)}{\sqrt{2}} \tag{1.73}$$

is symmetric (+) or antisymmetric (−) with respect to exchange of electrons depending on the ±. This type of symmetry is known as exchange symmetry. In eqn 1.73, $\phi(\mathbf{r}_i)$ and $\xi(\mathbf{r}_i)$ are single-particle wave functions for the i^{th} electron. Whatever the exchange symmetry of the spatial wave function, the spin wave function χ must have the opposite exchange symmetry. Hence χ must be antisymmetric when the spatial wave function is symmetric and vice versa. This is in order that the product $\psi_{\text{space}}(\mathbf{r}_1, \mathbf{r}_2) \times \chi$ is antisymmetric overall.

States like $|\uparrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$ are clearly symmetric under exchange of electrons, but when you exchange the two electrons in $|\uparrow\downarrow\rangle$ you get $|\downarrow\uparrow\rangle$ which is not equal to a multiple of $|\uparrow\downarrow\rangle$. Thus the state $|\uparrow\downarrow\rangle$, and also
by an identical argument the state $|\downarrow\uparrow\rangle$, are both neither symmetric nor antisymmetric under exchange of the two electrons. Hence it is not surprising that we will need linear combinations of these two states as our eigenstates. The linear combinations are shown in Table 1.1. $(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)/\sqrt{2}$ is symmetric under exchange of electrons (in common with the other two $s = 1$ states) while $(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)/\sqrt{2}$ is antisymmetric under exchange of electrons.

Another consequence of this asymmetry with respect to exchange is the Pauli exclusion principle, which states that two electrons cannot be in the same quantum state. If two electrons were in precisely the same spatial and spin quantum state (both in, say, spatial state $\phi(\mathbf{r})$ and both with, say, spinup), then their spin wave function must be symmetric under the exchange of the electrons. Their spatial wave function must then be antisymmetric under exchange, so

$$\begin{aligned}\psi_{\text{space}}(\mathbf{r}_1, \mathbf{r}_2) &= \frac{\phi(\mathbf{r}_1)\,\phi(\mathbf{r}_2) - \phi(\mathbf{r}_2)\,\phi(\mathbf{r}_1)}{\sqrt{2}} \\ &= 0\end{aligned} \tag{1.74}$$

Hence the state vanishes, demonstrating that two electrons cannot be in the same quantum state.

Very often we will encounter cases in which two spins are coupled via an interaction which gives an energy contribution of the form $A\hat{\mathbf{S}}^a \cdot \hat{\mathbf{S}}^b$, where A is a constant. If $A > 0$, the lower level will be a singlet (with energy $-3A/4$) with a triplet of excited states (with energy $A/4$) at an energy A above the singlet. This situation is illustrated in Fig. 1.9. A magnetic field can split the triplet state into the three different states with different values of m_s. If $A < 0$, the triplet state will be the lowest level.

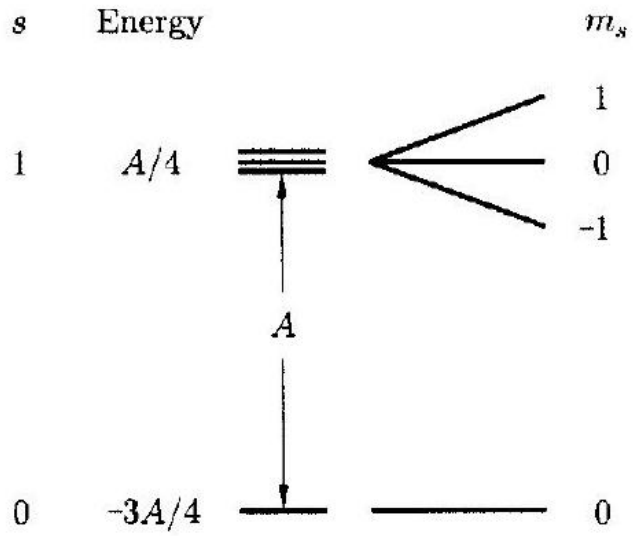


Fig. 1.9 The coupling of two electrons with an interaction of the form $A \hat{\mathbf{S}}^{a} \cdot \hat{\mathbf{S}}^{b}$ gives rise to a triplet ($s = 1$) and a singlet ($s = 0$). If $A > 0$ the singlet is the lower state and the triplet is the upper state. The triplet can be split into three components with a magnetic field B.}

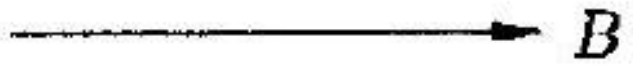


Fig. 1.9 The coupling of two electrons with an interaction of the form $A \hat{\mathbf{S}}^{a} \cdot \hat{\mathbf{S}}^{b}$ gives rise to a triplet ($s = 1$) and a singlet ($s = 0$). If $A > 0$ the singlet is the lower state and the triplet is the upper state. The triplet can be split into three components with a magnetic field B.}

9.1.4 Further reading

- B. I. Bleaney and B. Bleaney, Electricity and Magnetism, OUP 1989, contains a comprehensive treatment of electromagnetism (see also Appendix B).
- A. I. Rae, Introduction to Quantum Mechanics, IOP Publishing 1992 is a clear exposition of Quantum Mechanics at an introductory level.
- A good account of quantum angular momentum can be found in Chapters 1-3 of volume 3 of the Feynman lectures in Physics, R. P. Feynman, Addison-Wesley 1975.
- An excellent description of quantum mechanics may be found in J. J. Sakurai, Modern Quantum Mechanics, 2nd edition 1994, Addison-Wesley.

Exercises

(1.1) Calculate the magnetic moment of an electron (with $g = 2$). What is the Larmor precession frequency of this electron in a magnetic field of flux density 0.3 T? What is the difference in energy of the electron if its spin points parallel or antiparallel to the magnetic field? Convert this energy into a frequency.
(1.2) Using the definition of spin operators in eqn 1.43 , prove eqn 1.53 and the commutation relations, eqns 1.54 and 1.55.
(1.3) Using the definition of the raising and lowering operators in eqns 1.57, prove eqns 1.58, 1.61 .
(1.4) Using the commutation relation for spin, namely that $\left[\hat{S}_x, \hat{S}_y\right] = \mathrm{i}\hat{S}_z$ (and cyclic permutations), prove that

$$[\hat{\mathbf{S}} \cdot \mathbf{X}, \hat{\mathbf{S}}] = \mathrm{i}\hat{\mathbf{S}} \times \mathbf{X}, \tag{1.75}$$

where $\mathbf{X}$ is a vector.
(1.5) Using eqns 1.58 and 1.61 , show that

$$\hat{S}_{\pm} |S, S_z\rangle = \sqrt{S(S+1) - S_z (S_z \pm 1)}\, |S, S_z \pm 1\rangle\,, \tag{1.76}$$

where $|S, S_z\rangle$ represents a state with total spin angular momentum $S(S+1)\hbar^2$ and z component of spin angular momentum $S_z\hbar$. Hence prove the following special cases of eqn 1.76 :

$$\hat{S}_-|S, S\rangle = \sqrt{2}S|S, S-1\rangle \tag{1.77}$$

$$\hat{S}_+|S, S-1\rangle = \sqrt{2}S|S, S\rangle \tag{1.78}$$

(1.6) If the magnetic field $\mathbf{B}$ is uniform in space, show that this is consistent with writing $\mathbf{A} = \frac{1}{2}(\mathbf{B} \times \mathbf{r})$ and show that $\nabla \cdot \mathbf{A} = 0$. Are there other choices of $\mathbf{A}$ that would produce the same $\mathbf{B}$?
(1.7) The kinetic energy operator for an electron is $\hat{\mathbf{p}}^2/2m$. Use eqn 1.41 to show that this can be rewritten

$$\frac{(\boldsymbol{\sigma} \cdot \hat{\mathbf{p}})^2}{2m_{\mathrm{e}}}. \tag{1.79}$$

If a magnetic field is applied one must replace $\hat{\mathbf{p}}$ by $\hat{\mathbf{p}} + e\mathbf{A}$. With the aid of eqn 1.40 , show that this replacement substituted into eqn 1.79 leads to kinetic energy of the form

$$\frac{(\hat{\mathbf{p}} + e\mathbf{A})^2}{2m\mathrm{e}} + g\mu_{\mathbf{B}}\mathbf{B} \cdot \mathbf{S} \tag{1.80}$$

where the g -factor in this case is $g = 2$. (Note that in this problem you have to be careful how you apply eqn 1.40 and 1.41 because $\hat{\mathbf{p}}$ is an operator and will not commute with A.)
(1.8) An atom has zero orbital angular momentum and a spin quantum number $\frac{1}{2}$. It is found to be in the $|\uparrow_z\rangle$ state. A measurement is performed on the value of its angular momentum in a direction at an angle θ to the z axis. Show that the probability of its angular momentum being parallel to this new axis is $\cos^2(\theta/2)$.
(1.9) Using the basis of eqn 1.72, it is possible to construct matrix representations of operators such as $\hat{\mathbf{S}}_z^a \cdot \hat{\mathbf{S}}_z^b$ remembering that, for example, an operator such as $\hat{S}_z^a$ only operates on the part of the wave function connected with the first spin. Thus we have

$$\hat{S}_z^a = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{1.81}$$

$$\hat{S}_z^b = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{1.82}$$

Construct similar representations for $\hat{S}_x^a, \hat{S}_x^b, \hat{S}_y^a$ and $\hat{S}_y^b$ and hence show that

$$\hat{\mathbf{S}}^a \cdot \hat{\mathbf{S}}^b = \frac{1}{4}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{1.83}$$

Find the eigenvalues and eigenvectors of this operator and check that your results agree with those in Table 1.1.

(1.10) A magnetic field of 0.5 T is applied to a spherical sample of (a) water and (b) $MnSO_4 \cdot 4H_2O$. In each case, evaluate the fraction the H and B fields inside the sample differ from the free space values. (The magnetic susceptibilities of water and $MnSO_4 \cdot 4H_2O$ are listed in Table 2.1.) You should find that the corrections are very small indeed.

(1.11) Show that the operator

$$\hat{S}_{\theta,\phi} = \sin\theta\cos\phi\hat{S}_x + \sin\theta\sin\phi\hat{S}_y + \cos\theta\hat{S}_z, \tag{1.84}$$

which represents the spin operator for the component of spin along a direction determined by the spherical polar angles θ and ϕ, has eigenvalues $\pm\frac{1}{2}$ and eigenstates of the form

$$|\uparrow\rangle = \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2)\mathrm{e}^{\mathrm{i}\phi} \end{pmatrix} \tag{1.85}$$

$$|\downarrow\rangle = \begin{pmatrix} \sin(\theta/2) \\ -\cos(\theta/2)\mathrm{e}^{\mathrm{i}\phi} \end{pmatrix}. \tag{1.86}$$

Convince yourself that these results agree with the Riemann sphere representation in Fig. 1.8. Show further that

$$\hat{S}_{\theta,\phi}^2 = \frac{1}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{1.87}$$

(1.12) An electron in a magnetic field aligned along the z direction has a Hamiltonian (energy) operator

$$\hat{\mathcal{H}} = g\mu_\mathrm{B}\mathbf{B}\cdot\hat{\mathbf{S}} = g\mu_\mathrm{B}\boldsymbol{B}\hat{S}_z. \tag{1.88}$$

The time-dependent Schrödinger equation states that

$$\hat{\mathcal{H}}\psi(t) = \mathrm{i}\hbar\frac{\mathrm{d}\psi(t)}{\mathrm{d}t} \tag{1.89}$$

so that

$$\psi(t) = \exp(-\mathrm{i}\hat{\mathcal{H}}t/\hbar)\psi(0). \tag{1.90}$$

Using eqn 1.41, show that

$$\exp(\mathrm{i}\alpha\sigma_m) = I\cos\alpha + \mathrm{i}\sigma_m \sin\alpha \tag{1.91}$$

where I is the identity matrix, σ_m is one of the Pauli spin matrices and α is a real number. Hence show that if $\psi(t)$ is written as a spinor,

$$\psi(t) = \begin{pmatrix} \exp(-\mathrm{i}g\mu_\mathrm{B}Bt/2\hbar) & 0 \\ 0 & \exp(\mathrm{i}g\mu_\mathrm{B}Bt/2\hbar) \end{pmatrix} \psi(0) \tag{1.92}$$

and using the results from the previous question, show that this corresponds to the evolution of the spin state in such a way that the expected value of θ is conserved but ϕ rotates with an angular frequency given by $geB/2m$. This demonstrates that the phenomenon of Larmor precession can also be derived from a quantum mechanical treatment.

(1.13) Here is another way to derive spin precession. Start with eqn 1.88 and use eqn C. 7 to show that

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\hat{\mathbf{S}}\rangle = \frac{1}{\mathrm{i}\hbar}\langle[\hat{\mathbf{S}}, \hat{\mathcal{H}}]\rangle \tag{1.93}$$

$$= -\frac{g\mu_\mathrm{B}}{h}\langle\hat{\mathbf{S}}\rangle \times \mathbf{B} \tag{1.94}$$

which is similar to eqn 1.6 with

$$\gamma = -\frac{g\mu_\mathrm{B}}{h} = -\frac{ge}{2m_\mathrm{e}}. \tag{1.95}$$

The minus sign comes from the negative charge of the electron.

(1.14) This problem is about the corresponding case of an electric dipole. (a) An electric dipole with electric dipole moment $\mathbf{p}$ and moment of inertia I is placed in an electric field $\mathcal{E}$. Show classically that the angle θ, measured between $\mathbf{p}$ and $\mathcal{E}$, obeys the differential equation

$$\ddot{\theta} = -\frac{p\varepsilon\sin\theta}{I} \tag{1.96}$$

Show that this equation leads to simple harmonic motion when θ is very small.

(b) Now repeat the problem quantum mechanically. Consider the Hamiltonian

$$\hat{\mathcal{H}} = -\frac{\hbar^2}{2I}\frac{\partial^2}{\partial\theta^2} - p\varepsilon\cos\theta \tag{1.97}$$

and justify why this might be an appropriate Hamiltonian to use in this case. Using eqn C.7, show that

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\hat{\theta}\rangle = \frac{\langle\hat{L}\rangle}{I} \tag{1.98}$$

where $\hat{L} = -\mathrm{i}\hbar\partial/\partial\theta$ and that

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\hat{L}\rangle = -p\mathcal{E}\langle\sin\theta\rangle. \tag{1.99}$$

Hence deduce that

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\langle\hat{\theta}\rangle = -\frac{p\varepsilon}{I}\langle\sin\theta\rangle, \tag{1.100}$$

which reduces to the classical expression in the appropriate limit. Compare these results to the case of the magnetic dipole. Why are they different? Why does spin precession not result in the electric case?
We have shown that electric dipoles in an electric field oscillate backwards and forwards in the plane of the electric field, while magnetic dipoles precess around a magnetic field. In each case, what is wrong with our familiar idea that if you apply a field (electric or magnetic) then dipoles (electric or magnetic) just line up with the field?
2.1 An atom in a magnetic field 18
2.2 Magnetic susceptibility 19
2.3 Diamagnetism 20
2.4 Paramagnetism 23
2.5 The ground state of an ion and Hund's rules 30
2.6 Adiabatic demagnetization 36
2.7 Nuclear spins 38
2.8 Hyperfine structure 40

9.2 Isolated magnetic moments

In this chapter the properties of isolated magnetic moments will be examined. At this stage, interactions between magnetic moments on different atoms, or between magnetic moments and their immediate environments, are ignored. All that remains is therefore just the physics of isolated atoms and their interaction with an applied magnetic field. Of course that doesn't stop it being complicated, but the complications arise from the combinations of electrons in a given atom, not from the fact that in condensed matter there is a large number of atoms. Using this simplification, the large number of atoms merely leads to properties like the magnetic susceptibility containing a factor of n, the number of atoms per unit volume.

9.2.1 2.1 An atom in a magnetic field

In Section 1.1 (see eqn 1.35) it was shown that an electron spin in a magnetic field parallel to the z axis has an energy equal to

$$E = g\mu_{\mathrm{B}} B m_s \tag{2.1}$$

where $g \approx 2$ and $m_s = \pm\frac{1}{2}$. Hence $E \approx \pm\mu_{\mathrm{B}} B$. In addition to spin angular momentum, electrons in an atom also possess orbital angular momentum. If the position of the i^{th} electron in the atom is $\mathbf{r}_i$, and it has momentum $\mathbf{p}_i$, then the total angular momentum is $\hbar\mathbf{L}$ and is given by

$$\hbar\mathbf{L} = \sum_i \mathbf{r}_i \times \mathbf{p}_i \tag{2.2}$$

where the sum is taken over all electrons in an atom. Let us now consider an atom with a Hamiltonian $\hat{\mathcal{H}}_0$ given by

$$\hat{\mathcal{H}}_0 = \sum_{i=1}^{Z} \left(\frac{p_i^2}{2m} + V_i \right) \tag{2.3}$$

which is a sum (taken over the Z electrons in the atom) of the electronic kinetic energy ($p_i^2/2m_{\mathrm{e}}$ for the i^{th} electron) and potential energy (V_i for the i^{th} electron). Let us assume that the Hamiltonian $\hat{\mathcal{H}}_0$ has known eigenstates and known eigenvalues.

We now add a magnetic field $\mathbf{B}$ given by

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{2.4}$$

where $\mathbf{A}$ is the magnetic vector potential. We choose a gauge [1] such that

$$\mathbf{A}(\mathbf{r}) = \frac{\mathbf{B} \times \mathbf{r}}{2}. \tag{2.5}$$

Then the kinetic energy must be altered according to the prescription described in Section 1.2. Since the charge on the electron is $-e$, the kinetic energy is $\left[\mathbf{p}_i + e\mathbf{A}\left(\mathbf{r}_i\right)\right]^2 / 2m_\mathrm{e}$ and hence the perturbed Hamiltonian must now be written

$$\hat{\mathcal{H}} = \sum_{i=1}^{Z} \left(\frac{\left[\mathbf{p}_i + e\mathbf{A}\left(\mathbf{r}_i\right)\right]^2}{2m_\mathrm{e}} + V_i \right) + g\mu_\mathrm{B}\mathbf{B}\cdot\mathbf{S} \tag{2.6}$$

$$= \sum_i \left(\frac{p_i^2}{2m_\mathrm{e}} + V_i \right) + \mu_\mathrm{B}(\mathbf{L} + g\mathbf{S})\cdot\mathbf{B} + \frac{e^2}{8m_\mathrm{e}} \sum_i \left(\mathbf{B}\times\mathbf{r}_i\right)^2 \tag{2.7}$$

$$= \hat{\mathcal{H}}_0 + \mu_\mathrm{B}(\mathbf{L} + g\mathbf{S})\cdot\mathbf{B} + \frac{e^2}{8m_\mathrm{e}} \sum_i \left(\mathbf{B}\times\mathbf{r}_i\right)^2 \tag{2.8}$$

The dominant perturbation to the original Hamiltonian $\hat{\mathcal{H}}_0$ is usually the term $\mu_\mathrm{B}(\mathbf{L}+g\mathbf{S})\cdot\mathbf{B}$ but, as we shall see, it sometimes vanishes. This is the effect of the atom's own magnetic moment and is known as the paramagnetic term. The third term, $(e^2/8m_\mathrm{e})\sum_i \left(\mathbf{B}\times\mathbf{r}_i\right)^2$, is due to the diamagnetic moment. These contributions will be discussed in greater detail in Section 2.3 (diamagnetism) and Section 2.4 (paramagnetism). In the following section we outline the effects which will need explaining.

9.2.2 2.2 Magnetic susceptibility

As shown in Section 1.1.4, for a linear material $\mathbf{M} = \chi\mathbf{H}$ where $\mathbf{M}$ is the magnetic moment per volume (the magnetization) and χ is the magnetic susceptibility (dimensionless). Note that the definition of $\mathbf{M}$ means that χ represents the magnetic moment induced by a magnetic field $\mathbf{H}$ per unit volume. Magnetic susceptibilities are often tabulated in terms of the molar magnetic susceptibility, χ_m, where

$$\chi_\mathrm{m} = \chi V_\mathrm{m} \tag{2.9}$$

In this equation V_m is the molar volume, the volume occupied by 1 mole (6.022×10^{23} formula units) of the substance. The molar volume (in m^3) is the relative atomic mass [2] of the substance (in kg) divided by the density ρ (in kgm^{-3}). The mass susceptibility χ_g is defined by

$$\chi_\mathrm{g} = \frac{\chi}{\rho} \tag{2.10}$$

and has units of $\mathrm{m}^3\ \mathrm{kg}^{-1}$. The values of magnetic susceptibility for various substances are listed in Table 2.1. If the susceptibility is negative then the material is dominated by diamagnetism, if it is positive then the material is dominated by paramagnetism.

The magnetic susceptibilities of the first 60 elements in the periodic table are plotted in Fig. 2.1. Some of these are negative, indicative of the dominant rôle of diamagnetism as discussed in Section 2.3. However, some of the values are positive, indicative of paramagnetism and this effect will be discussed in Section 2.4.

[1] Equation 2.4 relates B and A. However, for a given magnetic field $\mathbf{B}$, the magnetic vector

potential **A** is not uniquely determined; one can add to **A** the gradient of a scalar potential and still end up with the same B. The choice of **A** that we make is known as a choice of gauge.

[2] The relative atomic mass is the mass of 1 mole. Note that relative atomic masses are usually tabulated in grams.

Table 2.1 The magnetic susceptibility χ and the molar magnetic susceptibility χ_{m} for various substances at 298 K . Water, benzene and NaCl are weakly diamagnetic (the susceptibility is negative). $CuSO_4 \cdot 5H_2O, MnSO_4 \cdot 4H_2O$, Al and Na are paramagnetic (the susceptibility is positive).

	$\chi/10^{-6}$	$\chi_{\mathrm{m}}/10^{-10}$ ($\mathrm{m^3\ mol^{-1}}$)
water	-90	-16.0
benzene	-7.2	-6.4
NaCl	-13.9	-3.75
graphite (||)	-260	-31
graphite ($\perp$)	-3.8	-4.6
Cu	-1.1	-0.078
Ag	-2.4	-0.25
$CuSO_4 \cdot 5H_2O$	176	192
$MnSO_4 \cdot 4H_2O$	2640	2.79×10^3
Al	22	2.2
Na	7.3	1.7

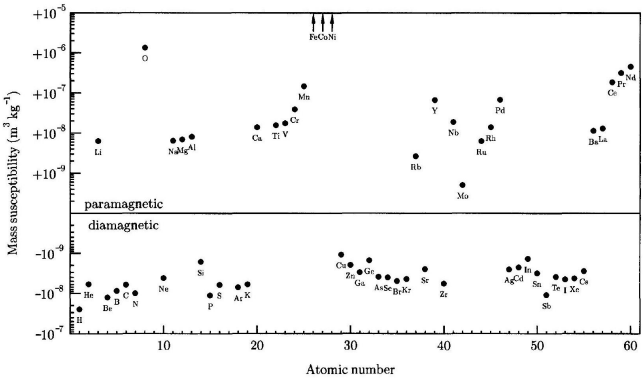


Fig. 2.1 The mass susceptibility of the first 60 elements in the periodic table at room temperature, plotted as a function of the atomic number. Fe , Co and Ni are ferromagnetic so that they have a spontaneous magnetization with no applied magnetic field.

9.2.3 2.3 Diamagnetism

[3] The prefix dia means 'against' or 'across' (and leads to words like diagonal and diameter).
[4] electromotive force
[5] See the further reading.

All materials show some degree of diamagnetism, [3] a weak, negative magnetic susceptibility. For a diamagnetic substance, a magnetic field induces a magnetic moment which opposes the applied magnetic field that caused it.

This effect is often discussed from a classical viewpoint: the action of a magnetic field on the orbital motion of an electron causes a back e.m.f., [4] which by Lenz's law opposes the magnetic field which causes it. However, the Bohrvan Leeuwen theorem described in the previous chapter should make us wary of such approaches which attempt to show that the application of a magnetic field to a classical system can induce a magnetic moment. [5] The phenomenon of diamagnetism is entirely quantum mechanical and should be treated as such.

We can easily illustrate the effect using the quantum mechanical approach. Consider the case of an atom with no unfilled electronic shells, so that the paramagnetic term in eqn 2.8 can be ignored. If $\mathbf{B}$ is parallel to the z axis, then $\mathbf{B} \times \mathbf{r}_i = B\left(-y_i, x_i, 0\right)$ and

$$\left(\mathbf{B} \times \mathbf{r}_i\right)^2 = B^2\left(x_i^2 + y_i^2\right) \tag{2.11}$$

so that the first-order shift in the ground state energy due to the diamagnetic term is

$$\Delta E_0 = \frac{e^2 B^2}{8 m_{\mathrm{e}}} \sum_{i=1}^{Z} \langle 0| \left(x_i^2 + y_i^2\right) |0\rangle \tag{2.12}$$

where $|0\rangle$ is the ground state wave function. If we assume a spherically symmetric atom, [6] $\langle x_i^2 \rangle = \langle y_i^2 \rangle = \frac{1}{3} \langle r_i^2 \rangle$ then we have

$$\Delta E_0 = \frac{e^2 B^2}{12 m_{\mathrm{e}}} \sum_{i=1}^{Z} \langle 0| r_i^2 |0\rangle. \tag{2.13}$$

Consider a solid composed of N ions (each with Z electrons of mass m) in volume V with all shells filled. To derive the magnetization (at $T = 0$), one can follow Appendix E, obtaining

$$M = -\frac{\partial F}{\partial B} = -\frac{N}{V} \frac{\partial \Delta E_0}{\partial B} = -\frac{N e^2 B}{6 m_{\mathrm{e}}} \sum_{i=1}^{Z} \left\langle r_i^2 \right\rangle, \tag{2.14}$$

where F is the Helmholtz function. Hence we can extract the diamagnetic susceptibility $\chi = M/H \approx \mu_0 M/B$ (assuming that $\chi \ll 1$). Following this procedure, we have the result that

$$\chi = -\frac{N}{V} \frac{e^2 \mu_0}{6 m_{\mathrm{e}}} \sum_{i=1}^{Z} \left\langle r_i^2 \right\rangle. \tag{2.15}$$

This expression has assumed first-order perturbation theory. (The second-order term will be considered in Section 2.4.4.) As the temperature is increased above zero, states above the ground state become progressively more important in determining the diamagnetic susceptibility, but this is a marginal effect. Diamagnetic susceptibilities are usually largely temperature independent.

This relation can be rather crudely tested by plotting the experimentally determined diamagnetic molar susceptibilities for various ions against $Z_{\mathrm{eff}}\, r^2$, where Z_{eff} is the number of electrons in the outer shell of an ion [7] and r is the measured ionic radius. The assumption is that all the electrons in the outer shell of the ion have roughly the same value of $\langle r_i \rangle^2$ so that

$$\sum_{i=1}^{Z_{\mathrm{eff}}} \left\langle r_i^2 \right\rangle \approx Z_{\mathrm{eff}} r^2 \tag{2.16}$$

The diamagnetic susceptibility of a number of ions is shown in Fig. 2.2. The experimental values are deduced by comparing the measured diamagnetic susceptibility of a range of ionic salts: $\mathrm{NaF}, \mathrm{NaCl}, \mathrm{NaBr}, \mathrm{KCl}, \mathrm{KBr}, \ldots$. The approach is inaccurate since not all the electrons

in an ion have the same mean radius squared (so that eqn 2.16 is by no means exact), but the agreement is nevertheless quite impressive. Ions are chosen because, for example, Na and Cl atoms have unpaired electrons but $Na^+$and Cl^-ions are both closed shell structures, similar to those of Ne and Ar (see the periodic table in Fig. 2.13 below for reference). Thus paramagnetic effects, which would dominate the magnetic response of the atoms, can be ignored in the ions.

Relatively large and anisotropic diamagnetic susceptibilities are observed in molecules with delocalized π electrons, such as naphthalene and graphite. Napthalene consists of two benzene molecules joined along one side (Fig. 2.3(a)). The π electrons are very mobile and induced currents can run round the edge of the ring, producing a large diamagnetic susceptibility which is largest if the magnetic field is applied perpendicular to the plane of the ring.

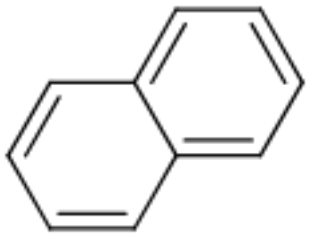
[6] This is a good assumption if the total angular momentum J is zero.

H. L. F. von Helmholtz (1821-1894)

[7] For an ion, this value is different from the atomic number Z, so we use the symbol Z_{eff} for an 'effective' atomic number. We are ignoring electrons in inner shells.

Fig. 2.3 (a) Naphthalene consists of two fused benzene rings. (b) Graphite consists of sheets of hexagonal layers. The carbon atoms are shown as black blobs. The carbon atoms are in registry in alternate, not adjacent planes (as shown by the vertical dotted lines).

(a)



(b)

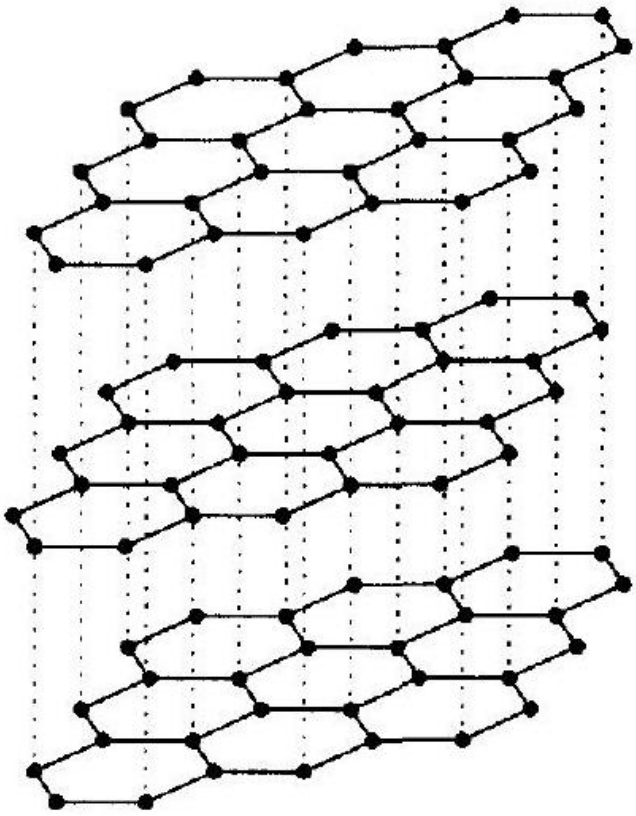
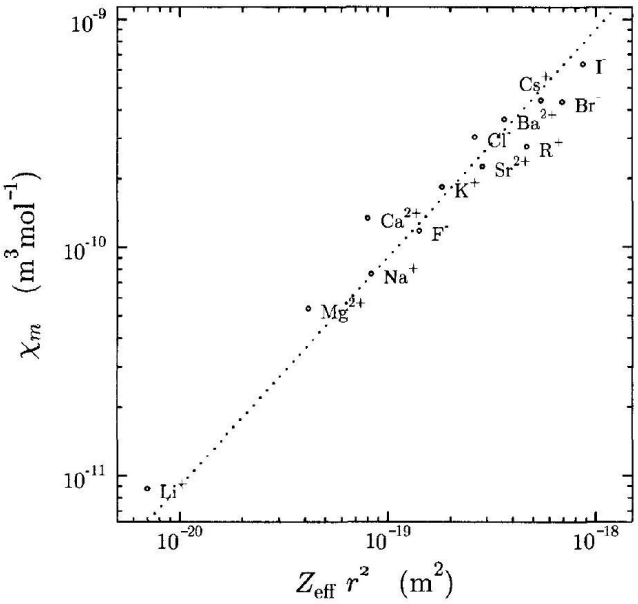


Fig. 2.2 The measured diamagnetic molar susceptibilities χ_{m} of various ions plotted against $Z_{\mathrm{eff}}\, r^2$, where Z_{eff} is the number of electrons in the ion and r is a measured ionic radius.}



The effective ring diameter is several times larger than an atomic diameter and so the effect is large. This is also true for graphite which consists of loosely bound sheets of hexagonal layers (Fig. 2.3(b)). The diamagnetic susceptibility is much larger if the magnetic field is applied perpendicular to the layers than if it is applied in the parallel direction.

Diamagnetism is present in all materials, but it is a weak effect which can either be ignored or is a small correction to a larger effect.

9.2.4 2.4 Paramagnetism

Paramagnetism [8] corresponds to a positive susceptibility so that an applied magnetic field induces a magnetization which aligns parallel with the applied magnetic field which caused it. In the previous section we considered materials which contained no unpaired electrons, and thus the atoms or molecules had no magnetic moment unless a field was applied. Here we will be concerned with atoms that do have a non-zero magnetic moment because of unpaired electrons. Without an applied magnetic field, these magnetic moments point in random directions because the magnetic moments on neighbouring atoms interact only very weakly with each other and can be assumed to be independent. The application of a magnetic field lines them up, the degree of lining up (and hence the induced magnetization) depending on the strength of the applied magnetic field.

The magnetic moment on an atom is associated with its total angular momentum $\mathbf{J}$ which is a sum of the orbital angular momentum $\mathbf{L}$ and the spin angular momentum $\mathbf{S}$, so that

$$\mathbf{J} = \mathbf{L} + \mathbf{S}. \tag{2.17}$$

Here, as throughout this book, these quantities are measured in units of $\hbar$. The way in which the spin and orbital parts of the angular momentum combine will be considered in detail in the following sections. In this section we will just assume that each atom has a magnetic moment of magnitude μ.

Although an increase of magnetic field will tend to line up the spins, an increase of temperature will randomize them. We therefore expect that the magnetization of a paramagnetic material will depend on the ratio B/T. The paramagnetic effect is in general much stronger than the diamagnetic effect, although the diamagnetism is always present as a weak negative contribution.

2.4.1 Semiclassical treatment of paramagnetism

We begin with a semiclassical treatment of paramagnetism (which as we will see below corresponds to $J = \infty$) in which we ignore the fact that magnetic moments can point only along certain directions because of quantization. Consider magnetic moments lying at an angle between θ and $\theta + \mathrm{d}\theta$ to the applied field $\mathbf{B}$ which is assumed without loss of generality to be along the z direction. These have an energy $-\mu B \cos\theta$ and have a net magnetic moment along $\mathbf{B}$ equal to $\mu\cos\theta$. If the magnetic moments could choose any direction to point along at random, the fraction which would have an angle between θ and $\theta + \mathrm{d}\theta$ would be proportional to the area of the annulus shown in Fig. 2.4 which is $2\pi\sin\theta\ \mathrm{d}\theta$ if the sphere has unit radius. The total surface area of the unit sphere is 4π so the fraction is $\frac{1}{2}\sin\theta\ \mathrm{d}\theta$. The probability of having angle between θ and $\theta+\mathrm{d}\theta$ at temperature T is then simply proportional to the product of this statistical factor, $\frac{1}{2}\sin\theta\ \mathrm{d}\theta$, and the Boltzmann factor $\exp(\mu B\cos\theta/k_{\mathrm{B}}T)$ where k_{B} is Boltzmann's constant. The average moment

[8] The prefix para means 'with' or 'along' and leads to English words such as parallel.

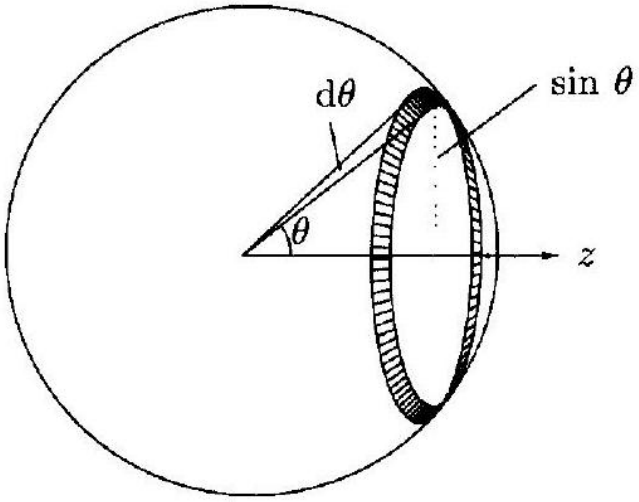
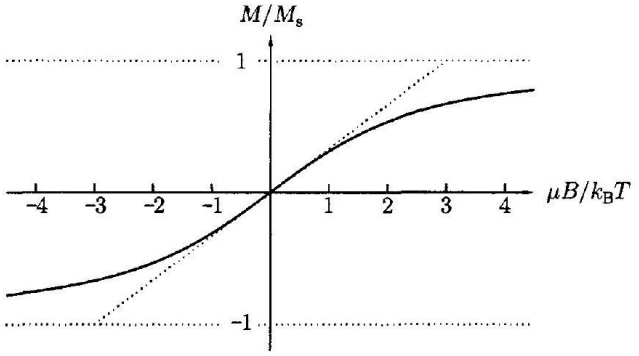


Fig. 2.4 To calculate the average magnetic moment of a paramagnetic material, consider the probability that the moment lies between angles θ and $\theta+\mathrm{d}\theta$ to the z axis. This is proportional to the area of the annulus on the unit sphere, shown shaded, which is $2\pi \sin\theta \, \mathrm{d}\theta$.}

Fig. 2.5 The magnetization of a classical paramagnet is described by the Langevin function, $L(y)=\operatorname{coth} y-\frac{1}{y}$. For small y, $L(y) \approx y/3$, as indicated by the line which is tangential to the curve near the origin. As the magnitude of the magnetic field is increased, or the temperature decreased, the magnitude of the magnetization increases.}



along $\mathbf{B}$ is then

$$\langle \mu_z \rangle = \frac{\int_0^\pi \mu \cos\theta \exp\left(\mu B \cos\theta / k_\mathrm{B}T\right) \frac{1}{2} \sin\theta \, \mathrm{d}\theta}{\int_0^\pi \exp\left(\mu B \cos\theta / k_\mathrm{B}T\right) \frac{1}{2} \sin\theta \, \mathrm{d}\theta} \tag{2.18}$$

$$= \mu \frac{\int_{-1}^1 x \mathrm{e}^{yx} \, \mathrm{d}x}{\int_{-1}^1 \mathrm{e}^{yx} \, \mathrm{d}x} \tag{2.19}$$

where I have defined $y = \mu B / k_\mathrm{B}T$ and $x = \cos\theta$. This leads to

$$\frac{\langle \mu_z \rangle}{\mu} = \coth y - \frac{1}{y} \equiv L(y) \tag{2.20}$$

Paul Langevin (1872-1946)

[9] For small fields, $\chi \ll 1$, so $B \approx \mu_0 H$.

where $L(y) = \coth y - 1/y$ is the Langevin function. It is shown in Fig. 2.5. For small y,

$$\coth(y) = \frac{1}{y} + \frac{y}{3} + O\left(y^3\right) \tag{2.21}$$

so that

$$L(y) = \frac{y}{3} + O\left(y^3\right). \tag{2.22}$$

We will use n to denote the number of magnetic moments per unit volume. The saturation magnetization, M_s, is the maximum magnetization we could obtain when all the magnetic moments are aligned, so that $M_\mathrm{s} = n\mu$. The magnetization that we actually obtain is $M = n\langle\mu_z\rangle$ and the ratio of the magnetization to the saturation magnetization is a useful quantity. Thus we have

$$\frac{M}{M_\mathrm{s}} = \frac{\langle\mu_z\rangle}{\mu} \approx \frac{y}{3} = \frac{\mu B}{3k_\mathrm{B}T} \tag{2.23}$$

and using $\chi = M/H \approx \mu_0 M/B$ which is valid in small fields, [9] we have

$$\chi = \frac{n\mu_0\mu^2}{3k_\mathrm{B}T} \tag{2.24}$$

This demonstrates that the magnetic susceptibility is inversely proportional to the temperature, which is known as Curie's law (after its discoverer, Pierre Curie). [10]

2.4.2 Paramagnetism for $J = \frac{1}{2}$

The calculation above will now be repeated, but this time for a quantum mechanical system. The classical moments are replaced by quantum spins with $J = \frac{1}{2}$. There are now only two possible values of the z component of the magnetic moments: $m_J = \pm\frac{1}{2}$. They can either be pointing parallel to B or antiparallel to B. Thus the magnetic moments are either $-\mu_\mathrm{B}$ or μ_B (assuming $g = 2$) with corresponding energies $\mu_\mathrm{B}B$ or $-\mu_\mathrm{B}B$. (These two solutions are sketched in Fig. 2.6.) Thus

$$\langle g\mu_\mathrm{B} m_J\rangle = \frac{-\mu_\mathrm{B}\mathrm{e}^{\mu_\mathrm{B}B/k_\mathrm{B}T} + \mu_\mathrm{B}\mathrm{e}^{-\mu_\mathrm{B}B/k_\mathrm{B}T}}{\mathrm{e}^{\mu_\mathrm{B}B/k_\mathrm{B}T} + \mathrm{e}^{-\mu_\mathrm{B}B/k_\mathrm{B}T}} \tag{2.25}$$

$$= \mu_\mathrm{B}\tanh\left(\frac{\mu_\mathrm{B}B}{k_\mathrm{B}T}\right) \tag{2.26}$$

so writing $y = \mu_\mathrm{B}B/k_\mathrm{B}T = g\mu_\mathrm{B}JB/k_\mathrm{B}T$ (where $J = \frac{1}{2}$ and $g = 2$) one has that

$$\frac{M}{M_\mathrm{s}} = \frac{\langle m_J\rangle}{J} = \tanh y \tag{2.27}$$

This function is different from the Langevin function, but actually looks pretty similar (see Fig. 2.7). In small applied fields $\tanh\left(\mu_\mathrm{B}/k_\mathrm{B}T\right) \approx \mu_\mathrm{B}/k_\mathrm{B}T$ and

$$\chi = \frac{n\mu_0\mu_\mathrm{B}^2}{k_\mathrm{B}T} \tag{2.28}$$

Equation 2.27 can be derived very efficiently using an alternative method. The partition function Z is the sum of the Boltzmann probabilities weighted by any degeneracy. The partition function for one spin is

$$Z = \mathrm{e}^{\mu_\mathrm{B} B/k_\mathrm{B}T} + \mathrm{e}^{-\mu_\mathrm{B} B/k_\mathrm{B}T} = 2\cosh\left(\frac{\mu_\mathrm{B} B}{k_\mathrm{B}T}\right) \tag{2.29}$$

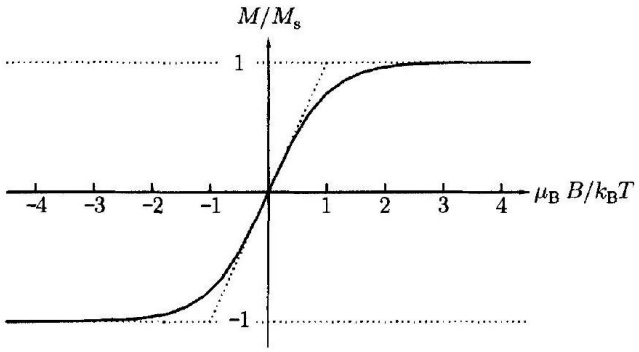


Fig. 2.7 The magnetization of a spin- $\frac{1
{2}$ paramagnet follows a $\tanh y$ function. For small $y, \tanh y \approx y$, as indicated by the line which is tangential to the curve near the origin.}

Pierre Curie (1859-1906)

[10] Often people write

$$\chi = \frac{C_\mathrm{Curie}}{T},$$

where C_Curie is the Curie constant.

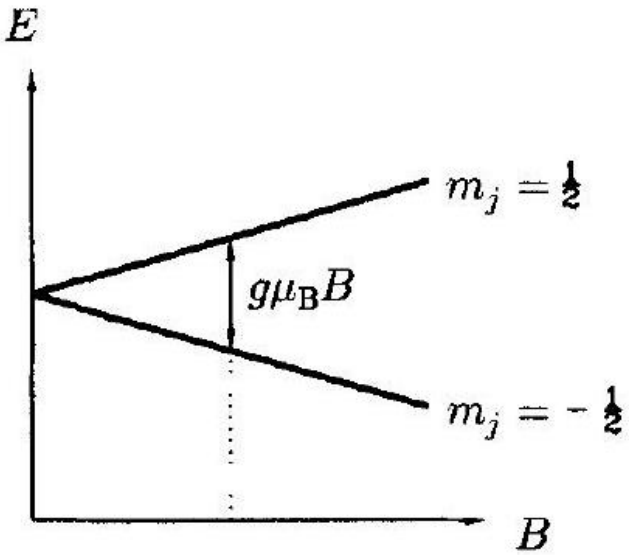
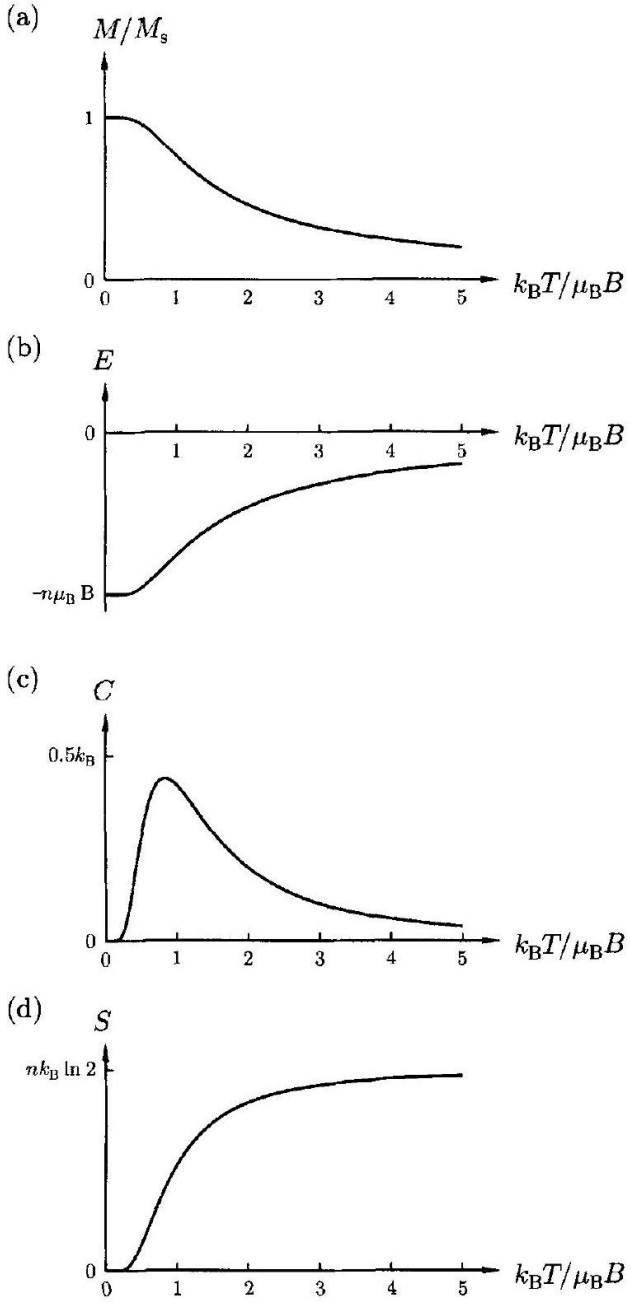


Fig. 2.6 The energy of a spin- $\frac{1
{2}$ magnetic moment as a function of magnetic field.}

See Appendix E for more details on Z, F and expressions such as $M = -(\partial F/\partial B)_T$.

Fig. 2.8 The (a) magnetization M (normalized by the saturation magnetization), (b) energy E, (c) heat capacity C (at constant applied magnetic field) and (d) entropy S of a paramagnetic salt containing n noninteracting spin- $\frac{1}{2}$ ions per unit volume as a function of $k_{\mathrm{B}}T/\mu_{\mathrm{B}}B$. The quantities E, C and S are therefore plotted per unit volume of paramagnetic salt.}



and the Helmholtz free energy can be evaluated using the expression $F = -k_{\mathrm{B}}T \ln Z$ yielding the Helmholtz free energy for n spins per unit volume as

$$F = nk_{\mathrm{B}}T \ln \left[2 \cosh \left(\frac{\mu_{\mathrm{B}}B}{k_{\mathrm{B}}T} \right) \right]. \tag{2.30}$$

The magnetization is then given by $M = -(\partial F/\partial B)_T$ which again yields

$$\frac{M}{M_{\mathrm{s}}} = \tanh \left(\frac{\mu_{\mathrm{B}}B}{k_{\mathrm{B}}T} \right), \tag{2.31}$$

in agreement with eqn 2.27.

This approach can also be used to derive other thermodynamic quantities for this model (see Exercise 2.4), the results of which are plotted in Fig. 2.8 as a function of $k_{\mathrm{B}}T/\mu_{\mathrm{B}}B$. Figure 2.8(a) thus shows the same information as that in Fig. 2.7 but with the horizontal axis inverted. This is because to understand some of the thermal properties of a material we are really interested in the effects of increasing temperature for a fixed magnetic field. As the sample is warmed, the magnetization decreases as the moments randomize but this produces an increase in energy density $E = -M_{\mathrm{s}}B$ (see Fig. 2.8(b)). When $T \to \infty$, the energy is zero since the moments are then completely random with respect to the applied field with the energy gains cancelling the energy losses. Cooling corresponds to an energy decrease (a point we will return to in Section 2.6).

The heat capacity, $C = (\partial E/\partial T)_B$ has a broad maximum close to $k_{\mathrm{B}}T \sim \mu_{\mathrm{B}}B$ which is known as a Schottky anomaly (see Fig. 2.8(c)). This arises because at this temperature, it is possible to thermally excite transitions between the two states of the system. At very low temperature, it is hard to change the energy of the system because there is not enough energy to excite transitions from the ground state and therefore all the spins are 'stuck', all aligned with the magnetic field. At very high temperature, it is hard to change the energy of the system because both states are equally occupied. In between there is a maximum. Peaks in the heat capacity can therefore be a useful indicator that something interesting may be happening. Note however that the Schottky anomaly is not a very sharp peak, cusp or spike, as might be associated with a phase transition, but is a smooth, broad maximum.

The entropy $S = -(\partial F/\partial T)_B$ rises as the temperature increases (see Fig. 2.8(d)), as expected since it reflects the disorder of the spins. Conversely, cooling corresponds to ordering and a reduction in the entropy. This fact is very useful in magnetic cooling techniques, as will be described in Section 2.6.

In the following section we will consider the general case of a paramagnet with total angular momentum quantum number J. This includes the two situations, classical and quantum, considered above as special cases.

2.4.3 The Brillouin function

The general case, where J can take any integer or half-integer value, will now be derived. Many of the general features of the previous cases ($J = \frac{1}{2}$ and $J = \infty$) are found in this general case, for example an increase in magnetic field will tend to align the moments while an increase in temperature will tend to disorder them.

The partition function is given by

$$Z = \sum_{m_J=-J}^{J} \exp\left(m_J g_J \mu_{\mathrm{B}} B/k_{\mathrm{B}}T\right) \tag{2.32}$$

Writing $x = g_J\mu_{\mathrm{B}}B/k_{\mathrm{B}}T$, we have

$$\langle m_J \rangle = \frac{\sum_{m_J=-J}^{J} m_J \mathrm{e}^{m_J x}}{\sum_{m_J=-J}^{J} \mathrm{e}^{m_J x}} = \frac{1}{Z}\frac{\partial Z}{\partial x} \tag{2.33}$$

Walter Schottky (1886-1976)

so that

$$M = n g_J \mu_{\mathrm{B}} \langle m_J \rangle = \frac{n g_J \mu_{\mathrm{B}}}{Z}\frac{\partial Z}{\partial B}\frac{\partial B}{\partial x} = n k_{\mathrm{B}} T \frac{\partial \ln Z}{\partial B}. \tag{2.34}$$

Now the partition function Z is a geometric progression with initial term $a = \mathrm{e}^{-Jx}$ and multiplying term $r = \mathrm{e}^{x}$. This can therefore be summed using the well-known formula

$$a + ar + ar^2 + \cdots + ar^{M-1} = \sum_{j=1}^{M} ar^{j-1} = \frac{a\left(1 - r^M\right)}{1 - r} \tag{2.35}$$

where M is the number of terms in the series, which in this case is $M = 2J + 1$. After a few manipulations, this leads to

$$Z = \frac{\sinh\left[(2J+1)\frac{x}{2}\right]}{\sinh\left[\frac{x}{2}\right]} \tag{2.36}$$

so that with the substitution

$$y = xJ = g_J \mu_{\mathrm{B}} J B / k_{\mathrm{B}} T, \tag{2.37}$$

we find

$$M = M_{\mathrm{s}}\, \mathrm{B}_J(y) \tag{2.38}$$

where the saturation magnetization M_{S} is

$$M_{\mathrm{s}} = n g_J \mu_{\mathrm{B}} J \tag{2.39}$$

Léon Brillouin (1889-1969)
and where $\mathrm{B}_J(y)$ is the Brillouin function given by

$$\mathrm{B}_J(y) = \frac{2J+1}{2J} \coth\left(\frac{2J+1}{2J} y\right) - \frac{1}{2J} \coth\frac{y}{2J}. \tag{2.40}$$

This function is plotted in Fig. 2.9 for various values of J. The Brillouin function has the appropriate limits. For example, when $J = \infty$ it reduces to a Langevin function:

$$\mathrm{B}_\infty(y) = L(y), \tag{2.41}$$

and when $J = \frac{1}{2}$ it reduces to a tanh function:

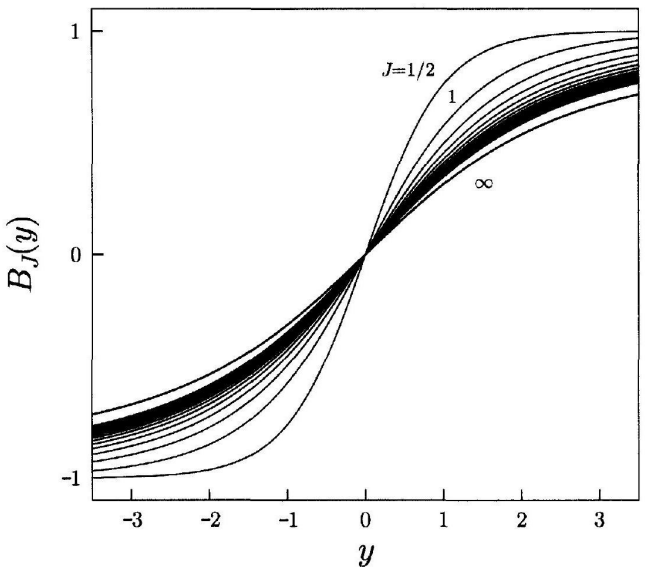
$$\mathrm{B}_{1/2}(y) = \tanh(y) \tag{2.42}$$

Hence it reduces to the cases considered in the previous sections.
A typical value of y can be estimated as follows: for $J = \frac{1}{2}, g_J = 2$ with $B = 1$ T, $y \sim 2 \times 10^{-3}$ at room temperature. Thus except at very low temperature and/or in extremely large magnetic fields, the experimental situation will correspond to $y \ll 1$ (and hence $\chi \ll 1$). For small y the following result can be derived by using the Maclaurin expansion of $\coth y$:

$$\mathrm{B}_J(y) = \frac{(J+1)y}{3J} + O\left(y^3\right). \tag{2.43}$$

Hence for low magnetic fields the susceptibility is given by

$$\chi = \frac{M}{H} \approx \frac{\mu_0 M}{B} = \frac{n \mu_0 \mu_{\mathrm{eff}}^2}{3 k_{\mathrm{B}} T} \tag{2.44}$$



which looks like a classical Curie law. [11] A measurement of χ therefore allows one to deduce μ_{eff} , the value of the effective moment,

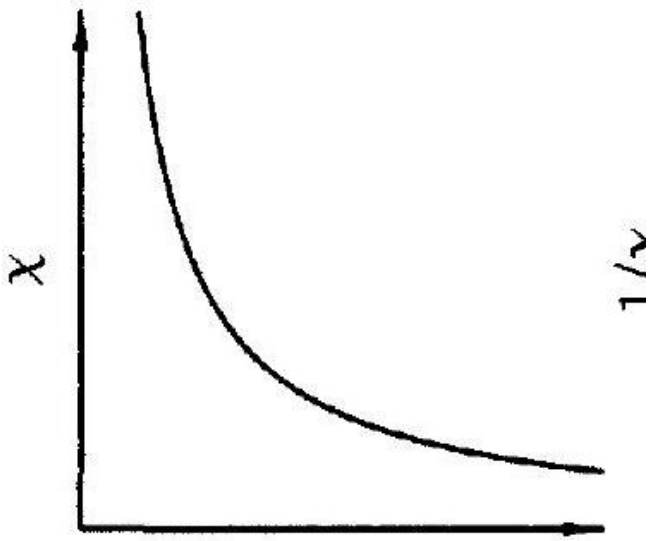
$$\mu_{\text{eff}} = g_J \mu_{\text{B}} \sqrt{J(J+1)} \tag{2.45}$$

where

$$g_J = \frac{3}{2} + \frac{S(S+1) - L(L+1)}{2J(J+1)} \tag{2.46}$$

The constant g_J is known as the Landé g-value (see Appendix C).
The Curie's law dependence of the susceptibility leads to $\chi \propto 1/T$ so that a graph of $1/\chi$ against T is a straight line and a graph of χT is constant against T (see Fig. 2.10). These points will be useful to keep in mind when in later chapters we consider the rôle of interactions. It is important to note

(a)



T

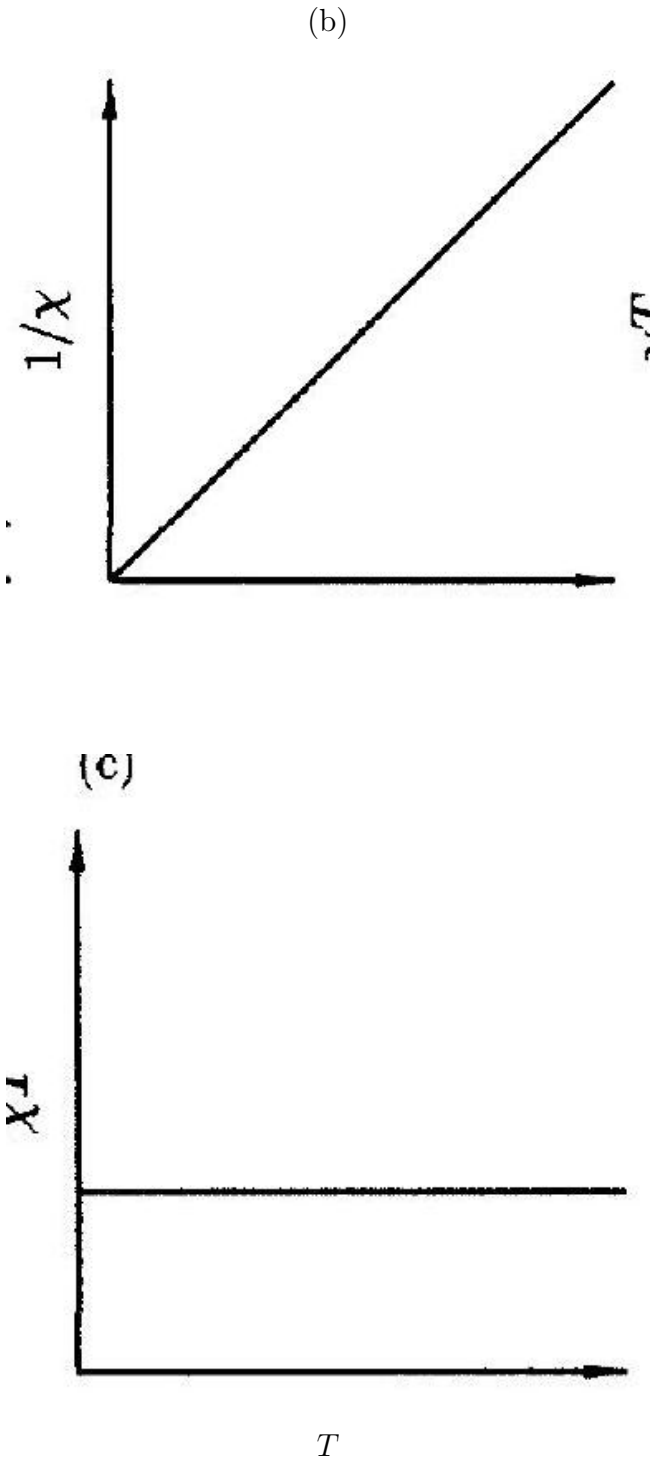


Fig. 2.9 The magnetization of a paramagnet with magnetic moment quantum number J follows a Brillouin function, $B_J(y)$, which is plotted here for different values of J. The values of J are $\frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \ldots$ and $J = \infty$.

[11] i.e. $\chi = C_{\text{Curie}}/T$ where $C_{\text{Curie}} = n\mu_0 g_J^2 J(J+1)/3k_B$ is the Curie constant.

Alfred Landé (1888-1975)

Fig. 2.10 Curie's law states that $\chi \propto 1/T$ as shown in (a). Thus a straight-line graph is obtained by plotting $1/\chi$ against T as shown in (b). A graph of χT against T is constant, as shown in (c).

that the susceptibility is evaluated in the limit of vanishing applied magnetic field and is then

given by eqn 2.44 with $\mu_{\text{eff}} = g_J \mu_{\text{B}} \sqrt{J(J+1)}$. However, at high applied magnetic fields, the magnetization saturates to M_{s} which, using eqn 2.39 , is equivalent to a moment of $g_J \mu_{\text{B}} J$ per ion. These two values, $\mu_{\text{eff}} = g_J \mu_{\text{B}} \sqrt{J(J+1)}$ and $M_{\text{s}}/n = g_J \mu_{\text{B}} J$, are unequal except when $J \to \infty$ (the classical limit).

2.4.4 Van Vleck paramagnetism

If $J = 0$ in the ground state $|0\rangle$, then there is no paramagnetic effect because

$$\langle 0|\hat{\mu}|0\rangle = g_J \mu_{\text{B}} \langle 0|\hat{\mathbf{J}}|0\rangle = 0 \tag{2.47}$$

This implies that the ground state energy of the system does not change if a magnetic field is applied and so that there is no paramagnetic susceptibility. However, this conclusion is only correct in first-order perturbation theory. Second-order perturbation theory nevertheless predicts a change in the ground state energy E_0 because it takes account of excited states with $J \neq 0$ being mixed in. The change of the ground state energy E_0 for an ion with $J = 0$ is

$$\Delta E_0 = \sum_n \frac{|\langle 0|(\mathbf{L} + g\mathbf{S}) \cdot \mathbf{B}|n\rangle|^2}{E_0 - E_n} + \frac{e^2}{8m_{\text{e}}} \sum_i (\mathbf{B} \times \mathbf{r}_i)^2 \tag{2.48}$$

where the second term is due to the diamagnetism and the sum in the first term is taken over all the excited states of the system. The magnetic susceptibility is then

$$\chi = \frac{N}{V} \left(2\mu_{\mathbf{B}}^2 \sum_n \frac{|\langle 0|\left(L_z + gS_z\right)|n\rangle|^2}{E_n - E_0} - \frac{e^2 \mu_0}{6m_{\text{e}}} \sum_{i=1}^{Z} \langle r_i^2 \rangle \right), \tag{2.49}$$

where the first term is positive (because $E_n > E_0$) and is called the van Vleck paramagnetism. The second term is negative and is the conventional diamagnetic susceptibility that we have already considered (see eqn 2.15). Van Vleck paramagnetism is, like diamagnetism, both small and temperature independent.

9.2.5 2.5 The ground state of an ion and Hund's rules

A typical atom does not contain one electron, but many. A lot of these will be in filled shells which have no net angular momentum. However, there may be unfilled shells and the electrons in these unfilled shells can combine to give non-zero spin and orbital angular momentum. In Section 1.3.4 we saw how two spin- $\frac{1}{2}$ electrons could combine into a joint entity with spin equal to either 0 or 1. In an atom, all the spin angular momentum from the electrons in the unfilled shells can combine together and so can all their orbital angular momenta. Thus an atom will have total orbital angular momentum $\hbar\mathbf{L}$ and total spin angular momentum $\hbar\mathbf{S}$. The orbital and spin angular momenta can therefore combine in

$$(2L+1)(2S+1) \tag{2.50}$$

ways. This is the total number of choices of the z component of $\mathbf{L}$ (which is the number of terms in the series $-L, -L+1, \ldots, L-1, L+1$, i.e. $(2L+1)$) multiplied by the total number of choices of the z component of $\mathbf{S}$ (i.e. $(2S+1)$ by a similar argument). These different configurations, obtained by differently combining together the angular momentum (both spin and orbital) from the electrons in the unfilled shells, will cost different amounts of energy. This difference in energy occurs because the choice of spin angular momentum affects the spatial part of the wave function and the orbital angular momentum affects how electrons travel around the nucleus: both therefore affect how well the electrons avoid each other and thus influence the

electrostatic repulsion energy. Below we will see how to find the configuration that minimizes the energy.

2.5.1 Fine structure

So far we have kept the spin and orbital angular momenta separate since they are independent of one another. However they do weakly couple, via the spin-orbit interaction (see Appendix C), which acts as a perturbation on the states with well defined $\mathbf{L}$ and $\mathbf{S}$. Because of this, $\mathbf{L}$ and $\mathbf{S}$ are not separately conserved but the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$ is conserved. If the relativistic effects are considered as a perturbation (which usually can be done) then one can consider $\mathbf{L}^2 = L(L+1)$ and $\mathbf{S}^2 = S(S+1)$ as being conserved. Thus states with L and S are split into a number of levels with differing J; this is known as fine structure. J takes the values from $|L-S|$ to $L+S$. From the definition of $\mathbf{J}$ (eqn 2.17),

$$\mathbf{J}^2 = \mathbf{L}^2 + \mathbf{S}^2 + 2\mathbf{L}\cdot\mathbf{S}, \tag{2.51}$$

and since the spin-orbit interaction takes the form $\lambda\mathbf{L}\cdot\mathbf{S}$ (see Appendix C), where λ is a constant, the expected value of this energy is

$$\langle\lambda\mathbf{L}\cdot\mathbf{S}\rangle = \frac{\lambda}{2}[J(J+1) - L(L+1) - S(S+1)] \tag{2.52}$$

The energy of the atom is mainly determined by the values of S and L via electrostatic considerations and so the energy eigenstates can be labelled with values of S and L. The precise value that J takes in the range $|L-S|$ to $L+S$ is immaterial in the absence of the spin-orbit interaction. Each level is a multiplet of $(2S+1)(2L+1)$ states. When adding the spin-orbit interaction as a perturbation, the multiplets split up into different fine structure levels labelled by J. Each of these levels themselves has a degeneracy of $2J+1$, so that these levels can be split up into their different m_J values by applying a magnetic field. The splitting of the different fine structure levels follows a relationship known as the Landé interval rule which will now be described. The energy separation between adjacent levels $E(J)$ and $E(J-1)$ of a given multiplet is given by

$$\begin{aligned} E(J) - E(J-1) =& \frac{\lambda}{2}[J(J+1) - L(L+1) - S(S+1)] \\ & - \frac{\lambda}{2}[(J-1)J - L(L+1) - S(S+1)] \\ =& \lambda J \end{aligned} \tag{2.53}$$

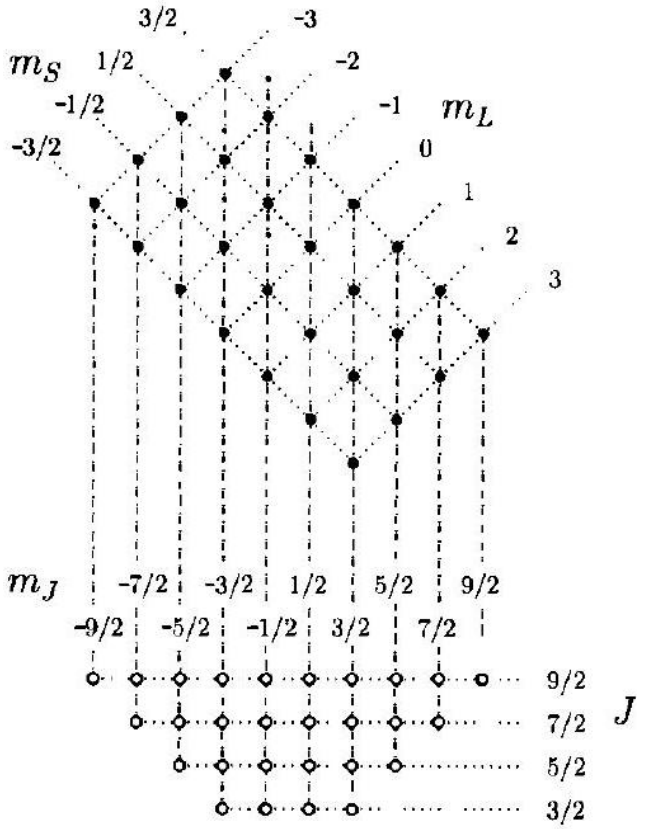


Fig. 2.11 The combination of angular momentum $L = 3$ and $S=\frac{3}{2}$ leads to values of J from $\frac{3}{2}$ to $\frac{9}{2}$.}

[12] Hund's rules are only applicable to the ground state configuration, and do not imply anything about the ordering of levels above the lowest level. They also assume that there is only one subshell which is incomplete.

Friederich Hund (1896-1997)
Thus the splitting is proportional to J (considering the separation of levels $J-1$ and J).

Example 2.1

An example of this combination of angular momentum is shown in Fig. 2.11 for the particular example of $L = 3$ and $S = \frac{3}{2}$. Here $(2S + 1)(2L + 1) = 28$ and the figure demonstrates that if one considers the states for which J takes the values from $|L - S| = \frac{3}{2}$ to $L + S = \frac{9}{2}$, since the degeneracy of each J state is $2J + 1$, there are also 28 possible states. This graphically illustrates the result that

$$\sum_{J=|L-S|}^{L+S} 2J + 1 = (2L + 1)(2S + 1) \tag{2.54}$$

There are clearly a lot of combinations of angular momentum quantum numbers that are possible; which one is the ground state for a particular ion?

2.5.2 Hund's rules

The combination of angular momentum quantum numbers which are found to minimize the energy can be estimated using Hund's rules. [12] These three empirical rules are listed in order of decreasing importance, so that one first satisfies the first and then, having done this, attempts to satisfy the second, and so on for the third.
(1) Arrange the electronic wave function so as to maximize S. In this way the Coulomb energy is minimized because of the Pauli exclusion principle, which prevents electrons with parallel spins being in the same place, and this reduces Coulomb repulsion between electrons.
(2) The next step is, given the wave function determined by the first rule, to maximize L. This also minimizes the energy and can be understood by imagining that electrons in orbits rotating in the same direction can avoid each other more effectively and therefore reduce Coulomb repulsion.
(3) Finally the value of J is found using $J = |L - S|$ if the shell is less than half full and $J = |L + S|$ if it is more than half full. This third rule arises from an attempt to minimize the spin-orbit energy. One should note that the third rule is only applicable in certain circumstances. As will be shown in the following chapter, in many systems, transition metal ions being good examples, the spin-orbit energies are not as significant as some other energy term such as the crystal field so that Hund's third rule is disobeyed. However, as shown below, for rare earth ions Hund's third rule works very well.

Having found values for S, L and J, this ground state can be summarized using a term symbol of the form ${}^{2S+1}L_J$. Here L is written not as a number, but using a letter according to the sequence

L	0	1	2	3	4	5	6	...
	S	P	D	F	G	H	I	...

and $2S + 1$ is the spin multiplicity.

Example 2.2

As an example, consider the rare earth ion Dy^{3+}, which has outer shell $4f^9$: f electrons have $l = 3$, so to satisfy Hund's first rule, $2l + 1 = 7$ of them are spin-up, and we then have 2 left for spin-down (see Fig. 2.12). This gives the value of S as $S = 7 \times \frac{1}{2} - 2 \times \frac{1}{2} = \frac{5}{2}$ (which implies that the spin degeneracy is $2S + 1 = 6$). The spin-up electrons give no net orbital angular momentum, so we only get an orbital contribution from the 2 spin-down electrons and it is this which we have to maximize. This then implies that $I = 3 + 2 = 5$ and hence we must use the symbol H . The shell is more than half full, so $J = \left|5 + \frac{5}{2}\right| = \frac{15}{2}$. Hence the term symbol is ${}^6\mathrm{H}_{15/2}$.

Hund's rules lead to a prediction of the ground state but tell us nothing about the excited states or how close they arc to the ground state. They therefore allow us to estimate the magnetic moment of an ion assuming that only this ground state is populated.

In this and the following chapter we will want to compare these predictions with experimental values and we will concentrate on compounds containing 3d and 4f ions since these are important in many magnetic systems. The 3d elements are the first row of transition metals ($\mathrm{Sc} - \mathrm{Zn}$) and the $4f$ elements are known as the lanthonides or as the rare earths (La-Lu). They are shown in the periodic table in Fig. 2.13.

H																
Li	Be											B Al				
Na													^{S}i	$^{1+}$P	S	^{17}Cl
K	Ca	Sc	^{22}	$\sqrt[3]{3}$	^{24}Cr	^{25}Mn	^{26}	Co	^{24}	^{22}Cu	^{30}	^{3}Ga	Ge	As	^{3}Se	Br
$\bar{r}$ Rb	Sr^{S}	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I
Cs	Ba	*	Hf	Ta	W	Re	Os	Ir	Pt	A_u	Hg	T	$^{+}$Pb	Bi	Po	At
Fr	Ra	†	mit Rf	Db	^{tab}	Bh	use	Mt								

*	La	Ce	Pr	Nd	Pm	Sm	^{5}	Gid	Tb	Dy	Ho	Er	Tm	Yb	Lu
†	Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cr	Es	Fm	Md	No	Lr

The application of Hund's rules to 3d and 4f ions is shown in Fig. 2.14 (although note that, as discussed in more detail in the following chapter, the Hund's rule predictions for $4f$ ions agree more with experiment than those for

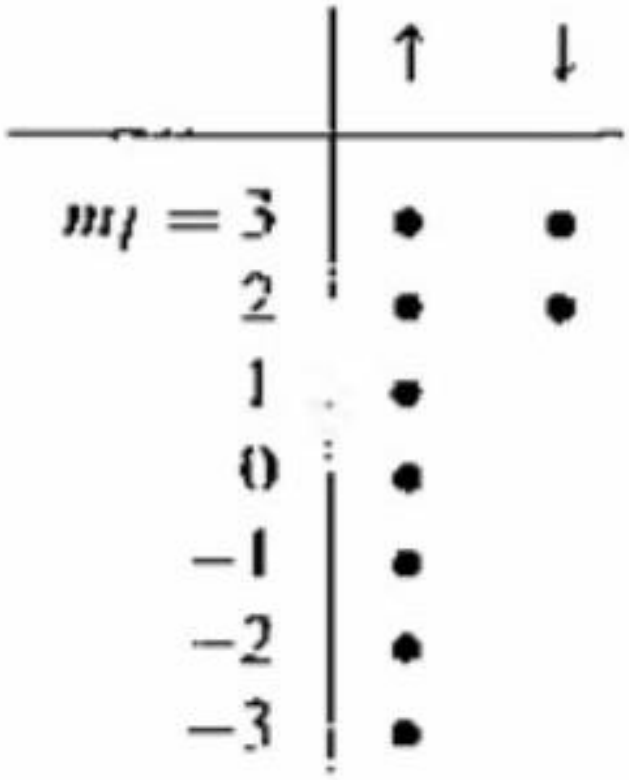


Fig. 2.12 The ground state of $\mathrm{Dy
ˆ{3+}$.}

Fig. 2.13 The periodic table. The 3d elements (Sc − Zn) and the 4 f elements (La − Lu) are shown shaded. The number by each element is the atomic number Z the number of protons in the nucleus).

[13] The present chapter deals only with free atoms or ions. Things will change when the atoms are put in a crystalline environment. The changes are quite large for 3d ions, as may be seen in chapter 3.

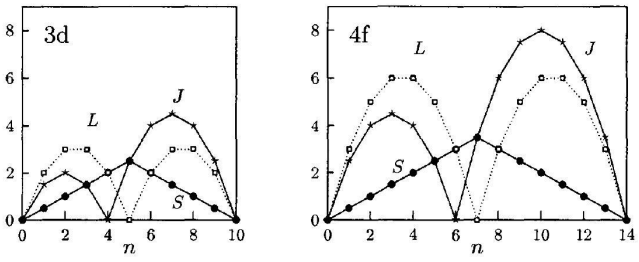
Fig. 2.14S, L and J for 3d and 4f ions according to Hund's rules. In these graphs n is the number of electrons in the subshell (3d or $4f$).

Table 2.2 Magnetic ground states for 4f ions using Hund's rules. For each ion, the shell configuration and the predicted values of S, L and J for the ground state are listed. Also shown is the calculated value of $p=\mu_{\text {eff

} / \mu_{\mathbf{B}}=g_{J}[J(J+1)]^{1 / 2}$ using these Hund's rules predictions. The next column lists the experimental value p_{exp} and shows very good agreement, except for Sm and Eu . The experimental values are obtained from measurements of the susceptibility of paramagnetic salts at temperatures $k_{\text{B}}T \gg E_{\text{CEF}}$ where E_{CEF} is a crystal field energy.}

ion	shell	S	L	J	term	p	p_{exp}
Ce^{3+}	$4f^1$	$\frac{1}{2}$	3	$\frac{5}{2}$	$^2F_{5/2}$	2.54	2.51
Pr^{3+}	$4f^2$	1	5	4	$^3H_4$	3.58	3.56
Nd^{3+}	$4f^3$	$\frac{3}{2}$	6	$\frac{9}{2}$	$^4\mathbf{I_{9/2}}$	3.62	3.3-3.7
$\mathbf{Pm^{3+}}$	$\mathbf{4}f^4$	2	6	4	$^5\mathbf{I_4}$	2.68	-
Sm^{3+}	$4f^5$	$\frac{5}{2}$	5	$\frac{5}{2}$	$^6I_{5/2}$	0.85	1.74
Eu^{3+}	$4f^6$	3	3	0	$^7F_0$	0.0	3.4
Gd^{3+}	$4f^7$	$\frac{7}{2}$	0	$\frac{7}{2}$	$^8S_{7/2}$	7.94	7.98
Tb^{3+}	$4f^8$	3	3	6	$^7F_6$	9.72	9.77
Dy^{3+}	$4f^9$	$\frac{5}{2}$	5	$\frac{15}{2}$	$^6H_{15/2}$	10.63	10.63
Ho^{3+}	$4f^{10}$	2	6	8	$^5I_8$	10.60	10.4
Er^{3+}	$4f^{11}$	$\frac{3}{2}$	6	$\frac{15}{2}$	$^4\mathbf{I_{15/2}}$	9.59	9.5
Tm^{3+}	$4f^{12}$	1	5	6	$^3H_6$	7.57	7.61
Yb^{3+}	$4f^{13}$	$\frac{1}{2}$	3	$\frac{7}{2}$	$^2F_{7/2}$	4.53	4.5
Lu^{3+}	$4f^{14}$	0	0	0	$^1S_0$	0	0

the 3d ions). ^{13}S rises and becomes a maximum in the middle of each group. L and J have maxima at roughly the quarter and three-quarter positions, although for J there is an asymmetry between these maxima which reflects the differing rules for being in a shell which is less than or more than half full.



From eqn 2.44 we have found that a measurement of the susceptibility allows one to deduce the effective moment. This effective moment can be expressed in units of the Bohr magneton μ_{B} as

$$p = \mu_{\text{eff}}/\mu_{\text{B}}, \tag{2.55}$$

and the Hund's rule predictions would suggest that, using eqn 2.45 , we should expect to measure $p = g_J[J(J+1)]^{\frac{1}{2}}$. Extremely good experimental agreement is usually found between this prediction and the measured values of $p = \mu_{\text{eff}}/\mu_{\text{B}}$ for 4 f ions in the solid state, as shown in Table 2.2. A discrepancy does occur for Sm and Eu but this is due to low-lying excited states with different J from the ground states which, because of their close proximity to the ground state in energy, are also significantly populated and cause p to shift from its ground state value. Although the $4f$ ions generally fit quite well, as will be seen in the following chapter much poorer agreement is found for the 3d ions because of the effect of the local crystal environment. This effect is not so important for 4 f ions because the partially filled 4 f shells lie deep within

the ion, beneath the filled 5 s and 5 p shells.

In this chapter we have considered the magnetic moment on each atom being associated with the electrons which are localized on that atom. In metals the conduction electrons are not localized on a particular atom but exist in band states and are delocalized over the entire material. Conduction electrons can show both a diamagnetic and paramagnetic effect and this will be considered in chapter 7.

2.5.3 L-S and j-j coupling

The atomic model that we have considered above has assumed L-S coupling (also known as Russell-Saunders coupling), namely that the spin-orbit interaction is a weak perturbation and the main energy terms are determined by the electrostatic interactions that control the values of L and S, i.e. by combining, separately, the orbital and spin angular momenta for the electrons. Only then, when the total orbital and spin angular momenta of the atom as a whole are known, do we consider applying the spin-orbit interaction as a weak perturbation which splits each term into fine structure levels labelled by J.

For atoms with high atomic number Z, this will not work because the spinorbit interaction energy is proportional [14] to Z^4 (see Appendix C) and therefore cannot be treated as a small perturbation for these atoms. A better scheme is known as $i-i$ coupling: here the spin-orbit interaction is the dominant energy and we couple the spin and orbital angular momentum of each electron separately, and consequently the weaker electrostatic effect may then couple the total angular angular momentum from each electron.

Example 2.3

As an example of these approaches, take carbon (C) and lead (Pb). They both have two p electrons in their outer shell, but Pb is much further down the periodic table than C and has a value of Z^4 which is nearly five orders of magnitude larger. Carbon (configuration: $2p^2$) can be treated with L-S coupling, so using Hund's rules we deduce that $S = 1$ (rule 1 is 'maximize S ' so combining the two electrons gives $\frac{1}{2} + \frac{1}{2}$), $L = 1$ (rule 2 is 'maximize L' and because the electrons are p electrons one of them can have $m_l = 1$, but then the next one must be $m_l = 0$) and $J = 0$ (rule 3 , less than half full so $J = |1 - 1|$), so that the term symbol for the ground state of C is $^3P_0$.

Henry Norris Russell (1877-1957)
Frederick A. Saunders (1875-1963)

[14] The Z^4 dependence works only for hydrogen-like atoms. For neutral atoms it is not so simple, but is close to Z^2.

In Pb (configuration: $6p^2$) the spin-orbit interaction is the dominant energy and it is more appropriate to use j−j coupling. Each electron has $s = \frac{1}{2}$ and $l = 1$ and so the combination j can be $\frac{1}{2}$ or $\frac{3}{2}$, depending on the sign of the spin-orbit energy. It turns out that this actually leads to $j = \frac{1}{2}$. Both electrons will have $j = \frac{1}{2}$ and these can couple via weaker electrostatic effects to make a resultant J which could be 0 or 1 . The ground state is $J = 0$ because of the Pauli exclusion principle (because the two electrons have otherwise identical quantum numbers and therefore must be kept apart in an antisymmetric state), so that the good quantum numbers for the ground state are $j_1 = \frac{1}{2}, j_2 = \frac{1}{2}$ and $J = 0$. Note that L and S are not good quantum numbers in this case. (Good quantum numbers are defined in Appendix C.)

9.2.6 2.6 Adiabatic demagnetization

In this section we will describe a technique which can be used to cool samples down to very low temperature, and which uses some of the ideas we have developed about paramagnetic materials. This brings out the idea of entropy rather clearly. Consider a sample of a paramagnetic salt, which contains N independent magnetic moments. Without a magnetic field applied, the magnetic moments will point in random directions (because we are assuming that they do not interact with each other) and the system will have no net magnetization. An applied field will however tend to line up the magnetic moments and produce a magnetization. Increasing temperature reduces the magnetization, increasing magnetic field increases the magnetization, emphasized by the fact that the magnetization is a function of B/T (see eqns 2.37 and 2.38).

At very high temperatures, the magnetic moments all point in random directions and the net magnetization is zero. The thermal energy $k_{\mathrm{B}}T$ is so large that all states are equally populated, irrespective of whether or not the state is energetically favourable. If the magnetic moments have $J = \frac{1}{2}$ they can only point parallel or antiparallel to the magnetic field: hence there are $W = 2^N$ ways of arranging up and down magnetic moments. Hence the magnetic contribution to the entropy, S, is

$$S = k_{\mathrm{B}} \ln W = N k_{\mathrm{B}} \ln 2 \tag{2.56}$$

In the general case of $J > \frac{1}{2}, W = (2J+1)^N$ and the entropy is

$$S = N k_{\mathrm{B}} \ln(2J+1) \tag{2.57}$$

To calculate the entropy at a lower temperature, one needs a different approach. The probability $p\left(m_J\right)$ that the z axis component of the total angular momentum of an ion takes the value m_J is proportional to the Boltzmann factor $\exp\left(-m_J g \mu_{\mathrm{B}} B / k_{\mathrm{B}} T\right)$ so that

$$p\left(m_J\right) = \frac{1}{Z} \exp\left(m_J g \mu_{\mathrm{B}} B / k_{\mathrm{B}} T\right) \tag{2.58}$$

where Z is the partition function (given in eqn 2.32). The amount of order of this system can be described by the entropy S and because we are now treating
the system probabilistically, we use the expression:

$$S = -N k_{\mathrm{B}} \sum_{m_J} p\left(m_J\right) \ln p\left(m_J\right). \tag{2.59}$$

Alternatively, the equation for the entropy can be generated by computing the Helmholtz free energy, F, via $F = -N k_{\mathrm{B}} T \ln Z$ and then using $S = -(\partial F / \partial T)_B$.

Let us now explore the consequences of eqn 2.59. In the absence of an applied magnetic field, or at high temperatures, the system is completely disordered and all values of m_J are equally likely with probability $p\left(m_J\right) = 1/(2J+1)$ so that the entropy S reduces to

$$S = N k_{\mathrm{B}} \ln(2J+1) \tag{2.60}$$

in agreement with eqn 2.57 . As the temperature is reduced, states with low energy become increasingly probable; the degree of alignment of the magnetic moments parallel to an applied magnetic field (the magnetization) increases and the entropy falls. At low temperatures, all the magnetic moments will align with the magnetic field to save energy. In this case there is only one way of arranging the system (with all spins aligned) so $W = 1$ and $S = 0$.

The principle of magnetically cooling a sample is as follows. The paramagnet is first cooled to a low starting temperature using liquid helium. The magnetic cooling then proceeds via two steps (see also Fig. 2.15).

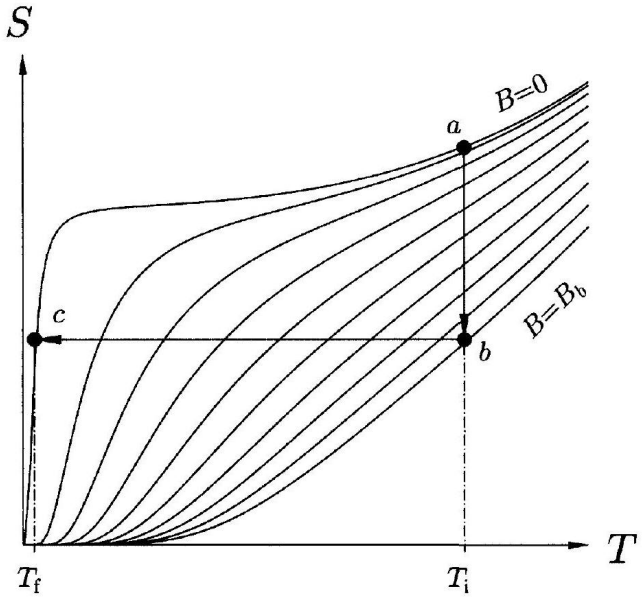


Fig. 2.15 The entropy of a paramagnetic salt as a function of temperature for several different applied magnetic fields between zero and some maximum value which we will call B_{b}. Magnetic cooling of a paramagnetic salt from temperature T_i to T_f is accomplished as indicated in two steps: first, isothermal magnetization from a to b by increasing the magnetic field from 0 to B_b at constant temperature T_1; second, adiabatic demagnetization from b to c. The $S(T)$ curves have been calculated assuming $J = \frac{1}{2}$ (see eqn 2.76). A term $\propto T^3$ has been added to these curves to simulate the entropy of the lattice vibrations. The curve for $B = 0$ is actually for B small but non-zero to simulate the effect of a small residual field.}

The first step is isothermal magnetization. The energy of a paramagnet is reduced by alignment of the moments parallel to a magnetic field. At a given temperature the alignment of the moments may therefore be increased by increasing the strength of an applied magnetic field. This is performed isothermally (see Fig. 2.15, step $a \rightarrow b$) by having the sample thermally connected to a bath of liquid helium (the boiling point of helium at atmospheric pressure is 4.2 K), or perhaps at reduced pressure so that the temperature can be less than 4.2 K . The temperature of the sample does not change and the helium bath absorbs the heat liberated by the sample as its energy and entropy decrease. The thermal connection is usually provided by low pressure helium gas in the sample chamber which conducts heat between the sample and the chamber walls, the chamber itself sitting inside the helium bath. (The gas is often called 'exchange' gas because it allows the sample and the bath to exchange heat.)

The second step is to thermally isolate the sample from the helium bath (by pumping away the exchange gas). The magnetic field is then slowly reduced to zero, slowly so that the process is quasi-static and the entropy is constant. This step is called adiabatic demagnetization (see Fig. 2.15, step $b \rightarrow c$) and it reduces the temperature of the system. During the adiabatic demagnetization the entropy of the sample remains constant; but the entropy of the magnetic moments increases (as the moments randomise as the field is turned down) which is precisely balanced by the entropy of the phonons (the lattice vibrations) which decreases as the sample cools. Entropy is thus exchanged between the phonons and the spins.

Does this method of cooling have a limit? At first sight it looks like the entropy for $B = 0$ would be $S = Nk_\mathrm{B} \ln(2J+1)$ for all temperatures $T > 0$ and therefore would fall to zero only at absolute zero. Thus adiabatic demagnetization looks like it might work as a cooling method all the way to absolute zero. However, in real systems there is always some small residual internal field due to interactions between the moments which ensures that the entropy falls prematurely

towards zero when the temperature is a little above absolute zero (see Fig. 2.15). The size of this field puts a limit on the lowest temperature to which the salt can be cooled. In certain salts which have a very small residual internal field, temperatures of a few milliKelvin can be achieved.

9.2.7 2.7 Nuclear spins

It is not only the electrons in an atom which have a magnetic moment. The nucleus often has a non-zero spin resulting from the angular momentum of the nucleus. For each nucleus there is a quantum number, I, called the nuclear spin quantum number which represents the total angular momentum of the nucleus in units of $\hbar$. However this magnetic moment is very small since its size scales with the inverse of the mass of the particles involved: nuclear moments are thus typically a thousand times smaller than electronic moments (most are between 10^{-3} and $10^{-4}\mu_{\mathrm{B}}$). The small size and absence of any strong interaction between nuclei in adjacent atoms precludes the ordering of nuclear spin systems at ordinary laboratory temperatures.

Table 2.3 Properties of some common nuclear spins. Listed first are the neutron (n), proton (p), deuteron (d) and triton (t), Z is the atomic number (the number of protons), N is the number of neutrons, and nucleus X is listed as ${

ˆ{A} \mathrm{X}$ where $A = Z + N$ is the mass number. I is the nuclear spin quantum number, μ is the nuclear magnetic moment measured in nuclear magnetons (μ_{N}) and g_I is the nuclear g-factor. The frequency ν at which these moments precess in a field of 1 T is shown in the final column.}

Nucleus	Z	N	I	μ/μ_{N}	g_1	ν (in MHz for $B = 1$ T)
n	0	1	$\frac{1}{2}$	-1.913	-3.826	29.17
p = ^{1}H	1	0	$\frac{1}{2}$	2.793	5.586	42.58
d = ^{2}H	1	1	1	0.857	0.857	6.536
t = ^{3}H	1	2	$\frac{1}{2}$	-2.128	4.255	32.43
^{12}C	6	6	0	0	0	0
^{13}C	6	7	$\frac{1}{2}$	0.702	1.404	10.71
^{14} N	7	7	1	0.404	0.404	3.076
^{16}O	8	8	0	0	0	0
^{17}O	8	9	$\frac{5}{2}$	-1.893	-0.757	5.772
^{19} F	9	10	$\frac{1}{2}$	2.628	5.257	40.05
^{31}P	15	16	$\frac{1}{2}$	1,132	2.263	17.24
^{33} S	16	17	$\frac{3}{2}$	0.643	0.429	3.266

The unit of nuclear magnetism is the nuclear magneton μ_{N} defined by

$$\mu_{\mathrm{N}} = \frac{e\hbar}{2m_{\mathrm{p}}} = 5.0508 \times 10^{-27}\mathrm{Am}^2, \tag{2.61}$$

where m_{p} is the mass of the proton. (This is much smaller than the magnetic moment of a 1 s electron which is given by the Bohr magneton $\mu_{\mathrm{B}} = e\hbar/2m_{\mathrm{e}} = 9.27 \times 10^{-24}$ A m^2, where m_{e} is the mass of the electron, see eqn 1.15.) The nuclear spin quantum number I takes one of the following values: $0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$ The component of angular momentum in the z-direction is given by m_I in units of $\hbar$ where m_I can take one of the following values:

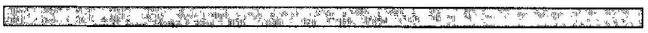
$$-I, -I+1, \cdots, I-1, I. \tag{2.62}$$

The magnetic moment of a nucleus takes a value, resolved along the z-direction for the spin state with the largest value of m_I, of

$$\mu = g_I \mu_N I \tag{2.63}$$

where g_I is the nuclear g-factor, a number of the order of unity which reflects the detailed structure of the nucleus. Values of I, g_I and μ are shown for selected nuclei in Table 2.3. Neither the magnetic moments of the neutron or the proton are simple multiples of μ_N. This is indicative of the complicated internal structure of each particle; both the neutron and the proton are each composed of three quarks. For a proton $I = \frac{1}{2}$ and $g_I = 5.586.$ and so its magnetic moment is $\mu_p = 1.410 \times 10^{-26}\text{Am}^2$.

Normally we ignore nuclear magnetism: the magnetic moments which exist from the protons in water do not make water stick to the poles of magnets (in fact water is mainly weakly diamagnetic, due to the clouds of electrons in the water molecules). However we can detect the magnetic moments of nuclei by doing experiments which use very sensitive resonant techniques. This experimental method, known as nuclear magnetic resonance, is discussed in Section 3.2.1.



Example 2.4

Many common nuclei have no magnetic moment at all, for example the commonest forms of carbon and oxygen, ^{12}C and ^{16}O. This is because both of these nuclei have an even number of protons and neutrons; pairs of identical particles in a nucleus tend to pair up as singlets. Non-zero magnetic moments are found in nuclei with unpaired nucleons. Thus ^{13}C has an unpaired neutron and has $I = \frac{1}{2}$, ^{14} N has an unpaired neutron and an unpaired proton and these combine to give $I = 1$. Though the commonest isotopes of carbon and oxygen have no nuclear moments, the next most common isotopes of these two nuclei do have a magnetic moment (about 1.1% of natural carbon is ^{13}C and about 0.04% of natural oxygen is ^{17}O; both of these do have a non-zero moment, see Table 2.3).

9.2.8 2.8 Hyperfine structure

Inside an atom the nuclear moment can magnetically interact with the electronic moment but only very weakly. This leads to energy splittings which are even smaller than the fine structure discussed in Section 2.5.1, and so is known as hyperfine structure. The weak magnetic interactions between the nucleus and the electrons are known as hyperfine interactions. Their origin is rather complex, but the essential principle can be understood in the following way. Consider a nuclear magnetic moment μ which sits in a magnetic field $\mathbf{B}_{\text{electrons}}$ which is produced by the motion and the spin of all the electrons. This produces an energy term $-\boldsymbol{\mu} \cdot \mathbf{B}_{\text{electrons}}$. Now $\mathbf{B}_{\text{electrons}}$ is expected to be proportional to the total angular momentum of all the electrons, $\mathbf{J}$, so that the Hamiltonian for the hyperfine interaction can be written as

$$\hat{\mathcal{H}}_{\text{hf}} = A\mathbf{I} \cdot \mathbf{J}. \tag{2.64}$$

Here $\mathbf{I}$ is the nuclear angular momentum and A is a parameter which can be determined from experiment and measures the strength of the hyperfine interaction.

The precise form of the hyperfine interaction is complicated for a general atom, but can be worked out in detail for a single-electron atom interacting with a point nucleus (see Exercise 2.8). Here we treat the problem in outline only to illuminate the basic mechanism of the interaction.

There are two types of magnetic interaction that are important. The first is the magnetic dipolar interaction which will be discussed in more detail in

Section 4.1. The dipolar interaction between a nuclear moment $\mu_I = g_I \mu_I \mathbf{I}$ and an electronic moment $\mu_c = g_c \mu_B S$ is given by $\hat{\mathcal{H}}_{\text{dipole}}$ where

$$\hat{\mathcal{H}}_{\text{dipole}} = \frac{\mu_0}{4\pi r^3}\left[\mu_c \cdot \mu_I - \frac{3}{r^2}(\mu_e \cdot \mathbf{r})(\boldsymbol{\mu}_I \cdot \mathbf{r})\right] \tag{2.65}$$

$$= \frac{\mu_0)\, g_e g_I \mu_B \mu_I}{4\pi r^3}\left[\mathbf{I}\cdot\mathbf{S} - \frac{3}{r^2}(\mathbf{S}\cdot\mathbf{r})(\mathbf{I}\cdot\mathbf{r})\right]. \tag{2.66}$$

If both moments are aligned with the z axis. this reduces to $\hat{\mathcal{H}}_{\text{dipole}} \propto (1 - 3\cos^2\theta)\, I_2 S_- / r^3$ and hence leads to an interaction energy which is proportional to

$$\int d^3\mathbf{r} |\psi(\mathbf{r})|^2 \left(1 - 3\cos^2\theta\right) m_s m_I / r^3. \tag{2.67}$$

an expression which vanishes in the case of an s orbital. This is essentially because a dipole field averages to zero over a spherical surface, as can bc understood from Fig. 2.16.

The spherically symmetric s orbital wave function averages this to zero. The integral does not vanish however for an unpaired electron in an orbital with $l > 0$ such as a porbital. Our treatment has also only considered the spin part of the electronic moment; the orbital moment of the electron also produces a magnetic field at the nucleus which gives a term proportional to $\mathbf{I}\cdot\mathbf{I}$. (All of the terms come out in the wash in the treatment of Exercisc 2.8.)

The second mechanism is the Fermi contact interaction, which in contrast to the previous mechanism which vanished for an s orbital, vanishes for every orbital except an s orbital. Here we consider what happens to eqn 2.66 when $r \to 0$. The nucleus is not of course a point dipole, but has a tinite volume, let us call it V. We further assume [15] that the nucleus is a sphere of uniform magnetization $\mathbf{M} = \mu_1 / V$. The magnetic flux density inside this sphere is then $\mu_0 \mathbf{M}$, but we must first subtract the demagnetization field which for a sphere is $\mu_0 \mathbf{M}/3$ (see Appendix D) leaving a flux density equal to $2\mu_0 \mathbf{M}/3$. Thus an electron which ventures inside the nuclear sphere would experience a magnetic flux density

$$\mathbf{B} = \frac{2\mu_0 \mu_I}{3V}. \tag{2.68}$$

so that the energy cost can be obtained by multiplying this by $\mu_e |\psi(0)|^2 V$, the amount of electronic moment which is in the nucleus (note that $|\psi(0)|^2 V$ is the probability of finding the electron inside the nucleus). Thus the energy cost E_{contact} is then

$$E_{\text{contact}} = \left(\frac{2\mu_0}{3}\right) 2\mu_B g_i \mu_N \mathbf{I}\cdot\mathbf{S} \tag{2.69}$$

where I have used $\mu_e = -2\mu_B S$ and $\mu_I = g_I \mu_I \mathbf{I}$. The net field experienced by the electron inside the nucleus can also be understood with reference to Fig. 2.16. Only s-electrons have any amplitude at $\mathbf{r} = 0$ and so only they show this effect. The hyperfine splitting caused by the Fermi contact interaction for s-electrons tends to be much larger than the hyperfine splitting due to the magnetic dipolar effect for electrons with $l > 0$.

As shown in Exercise 2.8 the form of the hyperfine interaction which includes both dipolar and Fermi contact terms is quite complicated but essentially consists of a constant times $\mathbf{I}\cdot\mathbf{S}$ for s-electrons ($l = 0$) and $\mathbf{I}\cdot\mathbf{N}$ for

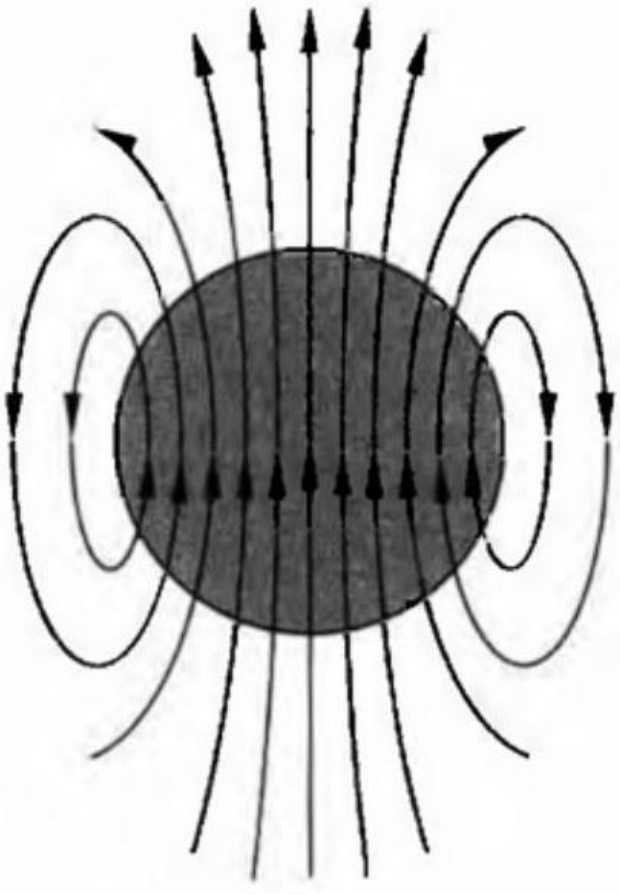


Fig. 2.16 A schematic diagram of the magnetic field distribution inside and outside of a magnetic nucleus (assumcd spherical and shaded in the diagram). The magnetic ficld distribution is caused by a corrent which flows around the equator of the sphere shown shaded in the figure. Outside the sphere the magnetic field appears to be due to a point dipole at the centre of the sphere and averaging the tield over a spherically symmetric volume gives rero: the field is as often pointing up as it is pointing down. Inside the nuclear sphere, the ficld is on average pointing up and does not average to zero over a spherically symmetric volume. This is the origin of the Fermi contact interaction.

[15] This assumption is used just to make the calculation simple. It actually gives the right answer, but a better treatment is given in Exercise 2.8.

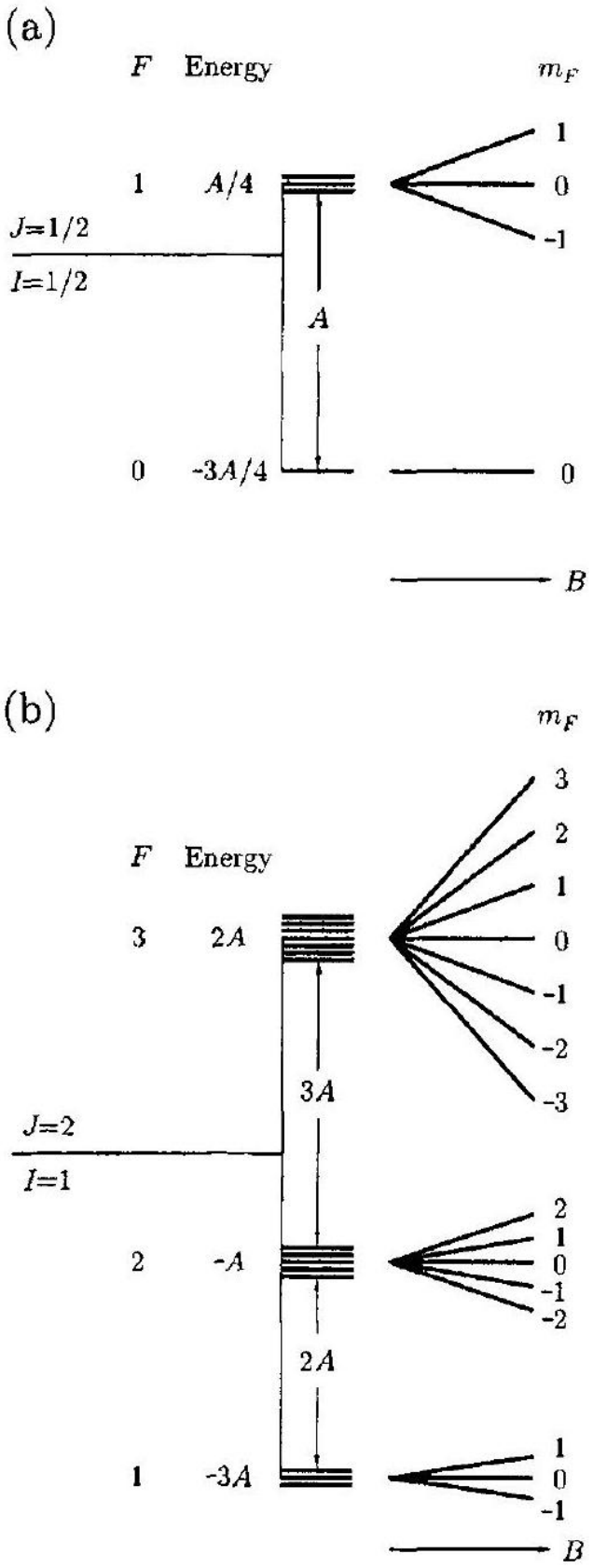


Fig. 2.17 An illustration of the hypefine splitting for nuclei with (a) $I=\frac{1}{2}$, $J=\frac{1}{2}$ and (b) $I = I, J = 2$. The quantum number F can take values $|J - I|, \ldots, J + I - 1, J + I$. The energy of each state is obtained by $\frac{A}{2}[F(F + 1) - I(I + 1) - J(J + 1)]$. Each hyperfine state, labelled by F, has a degeneracy of $2F + 1$, and these states can be split up into their different m_F values by applying a magnetic field B. Note that the Landé interval rule is obeyed. The size of the splitting in a magnetic field depends on the value of F, I and J via a g -factor for the quantum number F which is given approximately by $g_F = g_J[F(F + 1) + J(J + 1) - I(I + 1)/2F(F + 1)]$.}

electrons with $l > 0$, where $\mathbf{N} = \mathbf{L} - \mathbf{S} + 3(\mathbf{S} \cdot \mathbf{r})/r^2$. For s-electrons, $\mathbf{J} = \mathbf{S}$ and so the interaction agrees with eqn 2.64. It turns out that $\mathbf{N}$ can be projected on to $\mathbf{J}$ so that the expected value of $\mathbf{I}\cdot\mathbf{N}$ is proportional to $\mathbf{I}\cdot\mathbf{J}$ (the constant of proportionality is $L(L+1)/J(J+1)$ for a single electron atom) so that eqn 2.64 is also appropriate for electrons with $l > 0$.

The total angular momentum $\mathbf{F}$ of the atom, i.e. of the combination of the nucleus and the electron, is given by

$$\mathbf{F} = \mathbf{J} + \mathbf{I}. \tag{2.70}$$

The quantum number F can take values $|J - I|, \ldots, J + I - 1, J + I$. Most of the next few lines closely follow the treatment of the spin-orbit interaction. Using eqn 2.70, we can show that

$$\mathbf{F}^2 = \mathbf{J}^2 + \mathbf{I}^2 + 2\mathbf{I} \cdot \mathbf{J}, \tag{2.71}$$

and since the hyperfine interaction takes the form $A\mathbf{I} \cdot \mathbf{J}$, the expected value of this energy is

$$\langle A\mathbf{I} \cdot \mathbf{J}\rangle = \frac{A}{2}[F(F + 1) - I(I + 1) - J(J + 1)]. \tag{2.72}$$

The hyperfine interaction can be added as a perturbation, and it splits the electronic levels up into different hyperfine structure levels labelled by F. Each of these levels has a degeneracy of $2F+1$, so that these states can be split up into their different m_F values by applying a magnetic field. The splitting of the different hyperfine structure levels also follows a relationship known as the Landé interval rule by direct analogy with eqn 2.53 . The energy separation between adjacent levels $E(F)$ and $E(F - 1)$ is therefore given by

$$E(F) - E(F - 1) = AF. \tag{2.73}$$

Thus the splitting is proportional to F, the larger of the two quantum numbers of the adjacent levels being considered. This is illustrated in the following example.

Example 2.5

This effect is demonstrated for two examples in Fig. 2.17. Atomic hydrogen in the ground state contains one proton ($I = \frac{1}{2}$) and one electron ($J = S = \frac{1}{2}$) so that $F = 0$ or 1 (this example is shown in Fig. 2.17(a)). The energy splitting between these states is extremely small, about 6 orders of magnitude smaller than typical electronic energies, and corresponds to an electromagnetic frequency of about 1420 MHz , or a wavelength of about 21 cm . This transition is the basis of the atomic hydrogen maser and is also important in astronomy. Astronomers call this HI emission. If the concentration of atomic hydrogen in space is not so great that H_2 molecules can form, emission of radiation at 21 cm can be produced thermally (since the energy required is less than the temperature of the cosmic microwave background, see Fig. 3.8). This is Doppler shifted if the gas is moving with respect to the Earth, so that the rotation of gas within different regions of distant galaxies can be measured from the Earth. It can also be used for our own Galaxy, the Milky Way.

A more complex example is shown in Fig. 2.17(b), for an atom with $I = 1$ and $J = 2$. This shows that the Landé interval rule is obeyed as the size of the splitting between two adjacent levels scales with the value of F for the upper level.

9.2.9 Further reading

- B. H. Bransden and C. J. Joachain, Physics of atoms and molecules, Longman 1983, provides extensive information on isolated atoms.
- Useful background information may also be found in P. W. Atkins, Molecular quantum mechanics, OUP 1983.
- Also useful is the comprehensive book by A. Abragam and B. Bleaney, Electron paramagnetic resonance of transition ions, Dover 1986.
- D. J. Griffiths, Introduction to electromagnetism, Prentice Hall 1989 provides a readable account of magnetostatic fields in matter.
- A discussion of the merits of classical versus quantum mechanical derivations of diamagnetism is given in *S.L.* O'Dell and R. K. P. Zia, American Journal of Physics 54, 32 (1986).

Exercises

(2.1) Calculate the diamagnetic orbital susceptibility of a gas of hydrogen atoms (with number density 10^{20} m^{-3}) in the ground state, and compare this with the paramagnetic spin susceptibility at 100 K .
(2.2) Estimate the diamagnetic susceptibility of a duck (assume it is composed entirely of water). What magnetic field would be necessary to induce the same magnetic moment in the duck as is contained in a magnetized iron filing? Repeat the calculation for a cow.
(2.3) Calculate the paramagnetic moment of a crystal (with dimensions 2 mm $\times$ 2 mm $\times$ 2 mm) of $CuSO_4 \cdot 5H_2O$ (see Table 2.1, density 2286 kg m^{-3}, relative molecular mass 249.7 g) in a field of 1 T at 10 K .
(2.4) Using the expressions for the partition function and the Helmholtz function of a spin-$\frac{1}{2}$ particle in a magnetic field B from eqns 2.29 and 2.30 respectively, show that for n such non-interacting particles per unit volume, the energy E per unit volume is given by

$$E = -n\mu_B B \tanh\left(\frac{\mu_B B}{k_B T}\right) \tag{2.74}$$

the heat capacity per unit volume is given by

$$C = nk_B \left(\frac{\mu_B B}{k_B T}\right)^2 \operatorname{sech}^2\left(\frac{\mu_B B}{k_B T}\right) \tag{2.75}$$

and the entropy per unit volume is given by

$$S = nk_B \left[\ln\left(2\cosh\left(\frac{\mu_B B}{k_B T}\right)\right) - \frac{\mu_B B}{k_B T}\tanh\left(\frac{\mu_B B}{k_B T}\right)\right]. \tag{2.76}$$

These results are plotted in Fig. 2.8.
(2.5) Generalize the results of the previous problem to the case of general J. Writing the partition function

$$Z_J(y) = \frac{\sinh\left[(2J+1)\frac{y}{2J}\right]}{\sinh\left[\frac{y}{2J}\right]} \tag{2.77}$$

where $y = xJ = g_J\mu_B JB/k_BT$ (see eqn 2.37), show that the entropy $S(y)$ and the heat capacity at constant field $C(y)$ as a function of y are given by

$$S(y) = nk_B\left[\ln Z_J(y) - y\ \mathrm{B}_J(y)\right] \tag{2.78}$$

$$C(y) = nk_B y^2 \frac{\mathrm{dB}_J(y)}{\mathrm{d}y} \tag{2.79}$$

where $\mathbf{B}_J(y)$ is the Brillouin function. Hence show that for $y \ll 1$,

$$Z_J(y) = 2J + 1 \tag{2.80}$$

$$S(y) = nk_B\left[\ln(2J+1) - \frac{J+1}{3J}y^2\right] \tag{2.81}$$

$$C(y) = nk_B\frac{J+1}{3J}y^2 \tag{2.82}$$

and as $y \to \infty$,

$$S(y) \to 0 \tag{2.83}$$

$$C(y) \to 0 \tag{2.84}$$

(2.6) Show that Hund's rules for a shell of angular momentum l and containing n electrons can be summarized by

$$S = \frac{2l + 1 - |2l + 1 - n|}{2}$$
$$L = S|2l + 1 - n|$$
$$J = S|2l - n|$$

(2.7) Find the term symbols for the ground states of the ions (a) Ho^{3+} ($4f^{10}$), (b) Er^{3+} ($4f^{11}$), (c) Tm^{3+} ($4f^{12}$), and (d) Lu^{3+} ($4f^{14}$).
(2.8) The magnetic vector potential $\mathbf{A}$ at position $\mathbf{r}$ due to a point dipole μ at the origin is

$$\mathbf{A} = \frac{\mu_0}{4\pi r^3}\boldsymbol{\mu} \times \mathbf{r} = \frac{\mu_0}{4\pi}\nabla \times \left(\frac{\boldsymbol{\mu}}{r}\right) \tag{2.85}$$

Using this in the Hamiltonian for an electron in an atom

$$\hat{\mathcal{H}} = \frac{1}{2m_e}(\mathbf{p} + e\mathbf{A})^2 + 2\mu_B\mathbf{S}\cdot\mathbf{B} + V(r) \tag{2.86}$$

and neglecting terms in A^2, show that the perturbation, $\hat{\mathcal{H}}' = \hat{\mathcal{H}} - \hat{\mathcal{H}}_0$, to the Hamiltonian $\hat{\mathcal{H}}_0 = p^2/2m_e + V(r)$ due to the presence of a nuclear dipole $\mu = g_I\mu_N\mathrm{I}$ is

$$\hat{\mathcal{H}}' = \frac{\mu_0}{4\pi}2g_I\mu_B\mu_N\left[(\mathbf{S}\cdot\nabla)(\mathbf{I}\cdot\nabla)\frac{\mathbf{I}}{r} - (\mathbf{S}\cdot\mathbf{I})\nabla^2\frac{1}{r} + \frac{1}{r^3}\mathbf{L}\cdot\mathbf{I}\right] \tag{2.87}$$

To prove this you will need the vector identity

$$\nabla \times (\nabla \times \mathbf{Q}) = \nabla(\nabla\cdot\mathbf{Q}) - \nabla^2\mathbf{Q} \tag{2.88}$$

where $\mathbf{Q}$ is a vector field. Using $\nabla^2(1/r) = -4\pi\delta(\mathbf{r})$, where $\delta(\mathbf{r})$ is a delta function, show that

$$\begin{aligned}&(\mathbf{S}\cdot\nabla)(\mathbf{I}\cdot\nabla)\frac{1}{r}\\&= -\frac{1}{r^3}\left[\mathbf{S}\cdot\mathbf{I} - \frac{3(\mathbf{S}\cdot\mathbf{r})(\mathbf{I}\cdot\mathbf{r})}{r^2}\right] - \frac{4\pi}{3}\mathbf{S}\cdot\mathbf{I}\delta(\mathbf{r})\end{aligned} \tag{2.89}$$

and using these results deduce that

$$\begin{aligned}\hat{\mathcal{H}}' =&\frac{\mu_0}{4\pi}2g_I\mu_\mathrm{B}\mu_\mathrm{N}\mathbf{I}\\&\cdot\left[\frac{\mathbf{L}}{r^3} - \frac{\mathbf{S}}{r^3} + \frac{3\mathbf{r}(\mathbf{S}\cdot\mathbf{r})}{r^5} + \frac{8\pi}{3}\mathbf{S}\delta(\mathbf{r})\right]\end{aligned} \tag{2.90}$$

The first term in the square brackets is the coupling of the orbital moment of the electron with the nuclear moment. The second and third terms are the dipolar coupling of the spin of the electron with the nuclear moment. The final term in this equation is equal to the Fermi contact energy (eqn 2.69).

(2.9) The magnetic susceptibility of platinum is 2.61×10^{-4}. The density of platinum is 21450 kg m^{-3} and its relative atomic mass is given by 195.09 g mol^{-1}. Calculate its molar susceptibility (in m^3 mol^{-1}) and its mass susceptibility (in m^3 kg^{-1}). Using Appendix A, translate these results into cgs units to find the molar susceptibility in emu mol^{-1} and the mass susceptibility in emu g^{-1}. To understand why the magnetic susceptibility of platinum, a metal, is temperature independent (in contrast with Curie's law), see chapter 7.

(2.10) Consider a set of spins described by the partition function in eqn 2.32. The Helmholtz free energy density is $F = U - TS = -nk_\mathrm{B}T\log Z$ where $U = -\mathbf{M}\cdot\mathbf{B}$ is the potential energy density and S is the entropy density. This can be rearranged to show that

$$TS = -\mathbf{M}\cdot\mathbf{B} + nk_\mathrm{B}T\log Z. \tag{2.91}$$

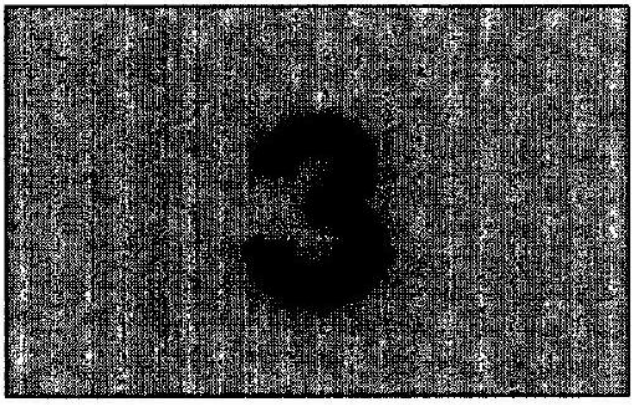
Show that if the magnetic field $\mathbf{B}$ is changed by $\delta\mathbf{B}$, with the temperature held fixed, then

$$T\delta S = -\mathbf{B}\cdot\delta\mathbf{M} \tag{2.92}$$

so that

$$\delta F = \delta U - T\delta S = -\mathbf{M}\cdot\delta\mathbf{B} \tag{2.93}$$

9.3 Environments



We have seen in the previous chapter that the magnetic properties of many crystals containing rare earths can be deduced by considering the rare earth ions to be behaving as completely free ions without interacting with each other or their surroundings. However, for magnetic ions in certain crystals one cannot usually ignore such interactions and for many materials they are large and significant. In this chapter we will consider the interactions between an atom and its immediate surroundings. In the following chapter we shall consider the direct magnetic interactions between a magnetic atom in a crystal and its neighbouring magnetic atoms.

9.3.1 3.1 Crystal fields

To understand the effect of the local environment due to the crystal on the energy levels of an atom, it is necessary to first review the shapes of the atomic orbitals. The angular dependences of the electron density of the s , p and d orbitals are shown in Fig. 3.1. This figure only shows the angular part of the wave functions for each orbital; there is also a radial part (for details, see Appendix C). Only s orbitals are spherically symmetric; the others have a pronounced angular dependence. This is crucial because the local environments are often not spherically symmetric so that different orbitals will behave in different ways.

3.1.1 Origin of crystal fields

The crystal field is an electric field derived from neighbouring atoms in the crystal. In crystal field theory the neighbouring orbitals are modelled as negative point charges; an improvement on this approximation is ligand field theory which is essentially an extension of molecular orbital theory that focusses on the rôle of the d orbitals on the central ion and their overlap with orbitals on surrounding ions (ligands). The size and nature of crystal field effects depend crucially on the symmetry of the local environment. A common case to consider is the octahedral environment. This is because in many transition metal compounds a transition metal ion sits at the centre of an octahedron with an ion such as oxygen on each corner. The crystal field in this case arises mainly from electrostatic repulsion from the negatively charged electrons in the oxygen orbitals. Diagrams of both an octahedral and a tetrahedral environment are shown in Fig. 3.2.

The d orbitals fall into two classes, the t_{2g} orbitals which point between the x, y and z axes (these are the d_{xy}, d_{xz} and d_{yz} orbitals) and the e_g orbitals which

3.1 Crystal fields 45
3.2 Magnetic resonance techniques 52

Fig. 3.1 The angular distribution of the s. p and d orbitals. The $d_{z

$ and $d_{\gamma_2} \ldots \Sigma_2 \geq$ levels are grouped together and called the e_g levels. The $\mathbf{d_{xy}}, \mathbf{d_1}$, and $\mathbf{d_{yz}}$ levels are grouped wgether and called the t_{2g} levels. The d_{-2} orbital is sometimes referred to as $d_{3-2} \ldots$.}

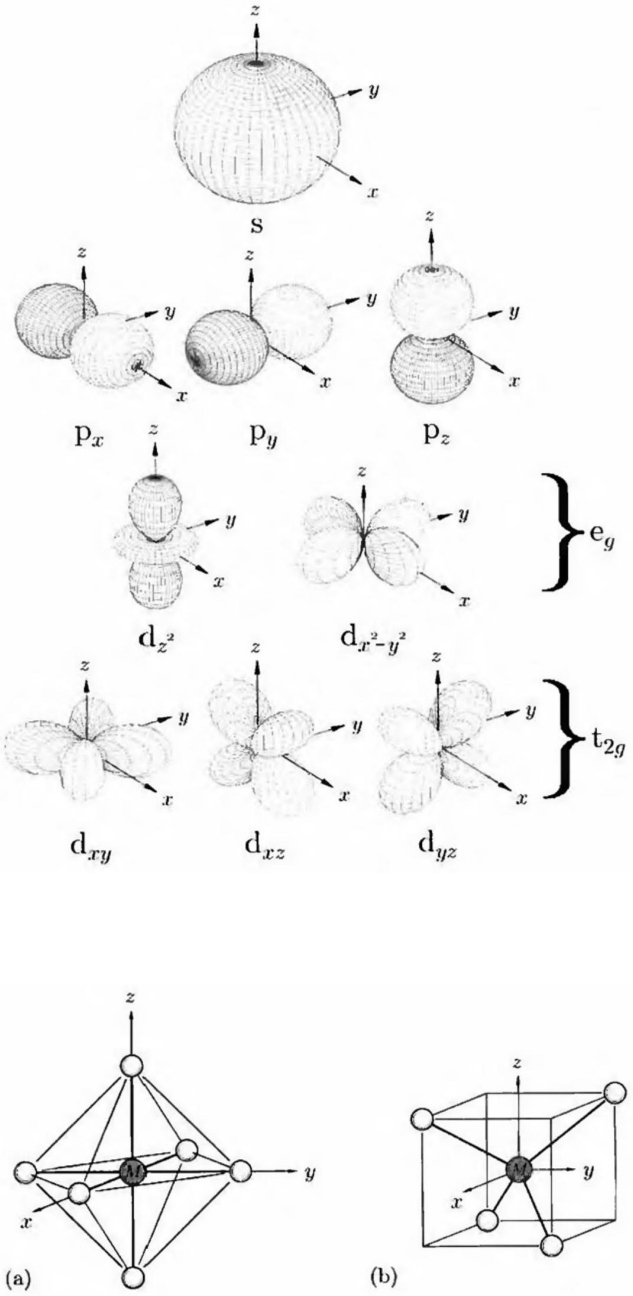
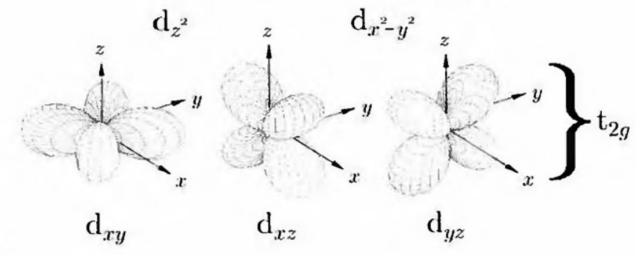


Fig. 3.2 A metal atom M in an (a) octuhedral and (b) tetrahedral environment. The octahedral environment is found in many transition metal oxides where oxygen anions sit on the corners of an octahedron with the metal atom at the centre. The tetrahedral environment can be conveniently described by considering alternate comers of a cube as shown.

Fig. 3.1 The angular distribution of the s. p and d orbitals. The $d_{z

$ and $d_{\gamma_2} \ldots \Sigma_2 \geq$ levels are grouped together and called the e_g levels. The $\mathbf{d_{xy}}, \mathbf{d_1}$, and $\mathbf{d_{yz}}$ levels are grouped wgether and called the t_{2g} levels. The d_{-2} orbital is sometimes referred to as $d_{3-2} \ldots$.}



point along these axes (the d_2 orbital, which has lobes which point along the 2 axis and the $d_{x^2-y^2}$ orbital, which has lobes which point along both the x and y-axes). Suppose a cation containing ten delectrons is placed in the centre of a sphere of radius r which is uniformly negatively charged. The d orbitals will all be degenerate in this spherically symmetric environment. though the presence of the charge will raise the energy of the entire system. Now imagine the charge on the sphere to collect into six discrete point charges, each lying at the vertex of an octahedron, but still on the surface of the sphere. The total electronic energy of all the d orbitals will not change, but the d orbitals will no longer be degenerate. What has now been created is an octahedral environment.

To demonstrate that the environment affects the orbitals in different ways. consider Fig. 3.3 which shows, in plan view, wo different d orbitals in an octahedral environment (this is the projection of Fig. 3.2(a) on to the xy plane). The crystal ficld is largely produced by p orbitals on neighbouring atoms. It is clear that the d_{xy} orbital (Fig. 3.3(a)) has a lower overlap with these neighbouring p orbitals than the $d_{x^2-y^2}$ orbital (Fig. 3.3(b)) and hence will have a lower electrostatic energy.

In an octahedral environment, neighbouring positive charges congregate at the points ($\pm r.0.0$), ($0. \pm r, 0$) and ($0, 0, \pm r$), the three orbitals d_{xy}, d_{yz}, d_{xx} which point between the x, y and z axes will be lowered in energy, but the d_z. and $d_{x^2-y^2}$ which point along the x, y and z axes will be raised in energy. The five levels therefore split as shown in Fig. 3.4(a), with the threefold t_{2g} levels lowered in energy and the twofold e_g levels raised in energy.

If the local environment is something other than octahedrally symmetric, the crystal field may even work in the opposite sense. For example, in a tetrahedral environment, the orbitals which point along the axes now maximally avoid the charge density associated with the atoms situated on four of the corners of the cube which describe a tetrahedron (see Fig. 3.2(b)). Thus in this tetrahedral case the two-fold e_g levels are lower in energy (see Fig. 3.4(b)).

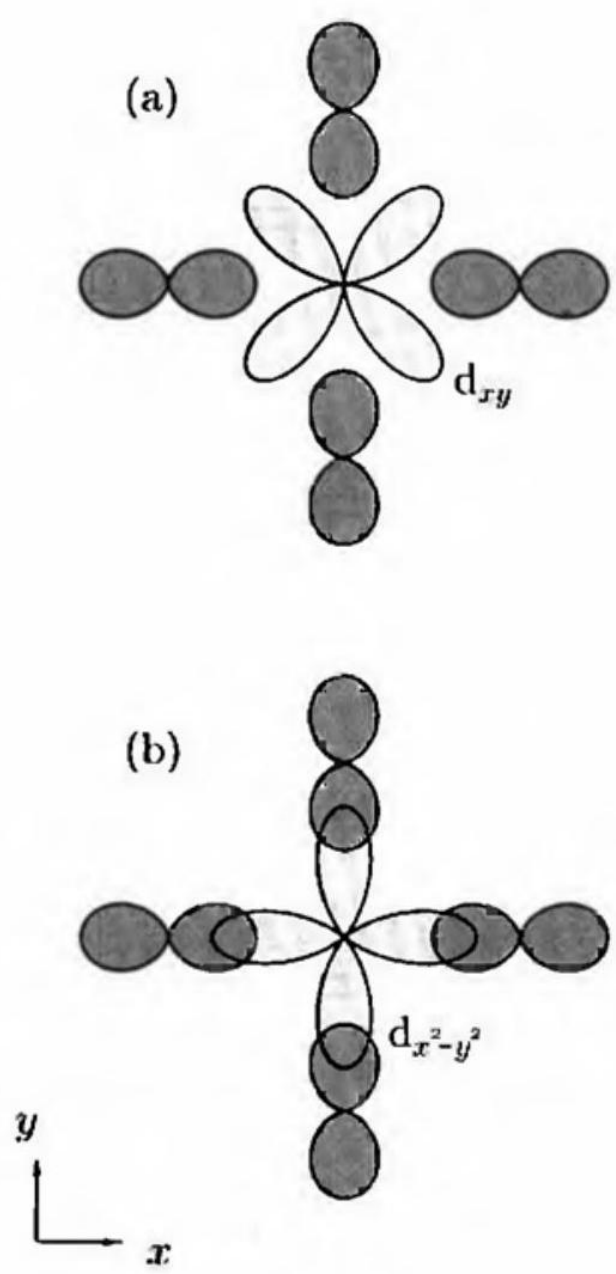


Fig. 3.3 The crystal lield originates from an electrostatic interaction. (as) The $\mathrm{d _{x y}$ orbital is lowered in energy with respect to tb) the $d_{12\ldots2}$ orbital in an uctahedral environment.}

(b)

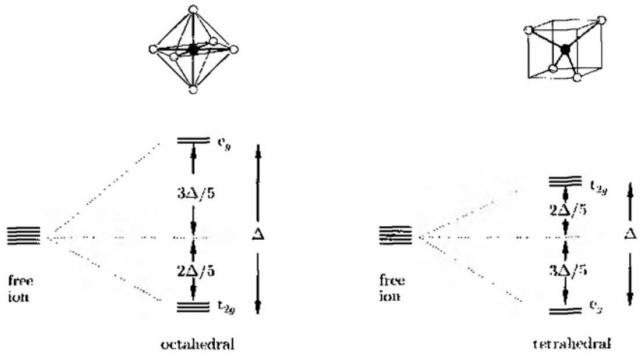
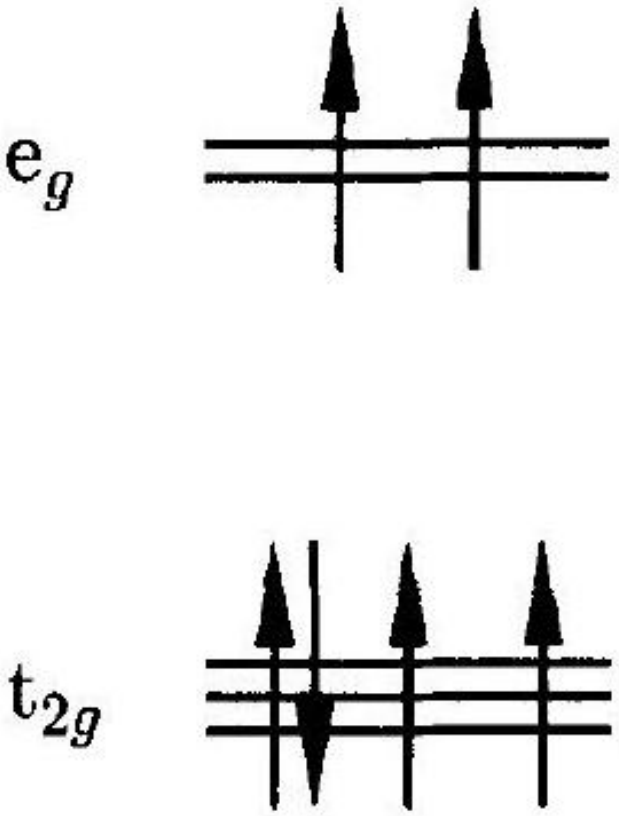
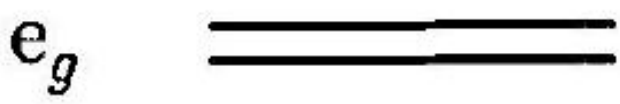


Fig. 3.4 The crystal field in an (a) octahedral and (b) terrahedral environment.

(a) $S = 2$



(b) $S = 0$



If one is dealing with a transition metal ion in which not all 3d electrons are present, those that are present will fill the lowest (in this case the t_{2g}) levels before filling the e_g levels. However, the precise order in which the orbitals fill depends on the competition between the crystal field energy and the Coulomb energy cost of putting two electrons in the same orbital, which is known as the pairing energy and is usually positive. If the crystal field energy is lower than the pairing energy (known as the weak-field case), then as the electrons are added to the system, they will first singly occupy each orbital before any orbital becomes doubly occupied. If on the other hand the crystal field energy is larger than the pairing energy (the strong-field case), electrons will doubly occupy the lower energy orbitals before they consider scaling the dizzy energetic heights of the higher energy orbitals. These cases are illustrated in Fig. 3.5.

If one is dealing with octahedral environments, there is never any doubt about how to add the electrons if you are dealing with adding $1, 2, 3, 8, 9$ or 10 electrons. The interesting cases occur with $4, 5, 6$ or 7 . This will now be illustrated for the case of adding six electrons, as would be appropriate for an Fe^{2+} ion (see also Fig. 3.5).

Example 3.1

The Fe^{2+} ion has a 3 d^6 shell. The pairing energy is positive. In the weak-field case, one would fill each orbital once which would leave one electron to spare; this could then be reluctantly paired up with one of the t_{2g} electrons. This would leave four unpaired electrons and a $S = 2$ state (see Fig. 3.5(a)). This is known as a high-spin configuration. In the strong field case, all six electrons are shoe-horned into the three t_{2g} orbitals leaving the e_g orbitals unfilled. Since there are no unpaired electrons, the system has $S = 0$ (see Fig. 3.5(b)). This is known as a low-spin configuration.

In some materials with Fe^{2+} it is possible to initiate a spin transition between the low-spin and high-spin configurations using temperature or pressure or even light irradiation.

ESW.E

3.1.2 Orbital quenching

The expected magnetic ground states for 3d ions are shown in Table 3.1. It is a simple matter to calculate the values of S, L and J by following Hund's rules as outlined in Section 2.5. However, as mentioned there, the values of the predicted moment given by $g_J[J(J+1)]^{1/2}$ do not always appear to agree with experiment. The exception to this is the case of $3d^5$ and $3d^{10}$ for which there is a half or completely full shell of electrons so that $L = 0$.

The reason for this discrepancy is that for 3d ions the crystal field interaction is much stronger than the spin-orbit interaction. Hence Hund's third rule as stated in Section 2.5, which is based upon the fact that spin-orbit interaction is the next most significant energy term after Coulombic effects, is actually wrong. The data appear to suggest that these systems instead choose a ground

Table 3.1 Magnetic ground states for 3d ions using Hund's rules. For each ion, the shell configuration and the predicted values of S, L and J for the ground state are listed. Also shown is the calculated value of $p=\mu_{\text {eff

} / \mu_{\mathrm{B}}}$ for each ion using Hund's rules predictions. This is given the symbol $p_1 = g_J[J(J+1)]^{1/2}$ and the next column lists the experimental values p_{exp} which are derived from measurements on paramagnetic salts containing the relevant ions. This agrees much better with $p_2 = 2[S(S+1)]^{1/2}$, which assumes orbital quenching, so that $L = 0, J = S$ and $g_J = 2$.}

ion	shell	S	L	J	term	p_1	p_{exp}	p_2
Ti^{3+}, V^{4+}	3 d^1	$\frac{1}{2}$	2	$\frac{3}{2}$	$^2D_{3/2}$	1.55	1.70	1.73
v^{3+}	3 d^2	1	3	2	$^3F_2$	1.63	2.61	2.83
Cr^{3+}, V^{2+}	3 d^3	$\frac{3}{2}$	3	$\frac{3}{2}$	$^4F_{3/2}$	0.77	3.85	3.87
Mn^{3+}, Cr^{2+}	3 d^4	2	2	0	$^5D_0$	0	4.82	4.90
Fe^{3+}, Mn^{2+}	3 d^5	$\frac{5}{2}$	0	$\frac{5}{2}$	$^6S_{5/2}$	5.92	5.82	5.92
Fe^{2+}	3 d^6	2	2	4	$^5D_4$	6.70	5.36	4.90
Co^{2+}	3 d^7	$\frac{3}{2}$	3	$\frac{9}{2}$	$^4F_{9/2}$	6.63	4.90	3.87
Ni^{2+}	3 d^8	1	3	4	$^3F_4$	5.59	3.12	2.83
Cu^{2+}	3 d^9	$\frac{1}{2}$	2	$\frac{5}{2}$	$^2D_{5/2}$	3.55	1.83	1.73
Zn^{2+}	3 d^{10}	0	0	0	$^1S_0$	0	0	0

state such that $L = 0$ (so that $J = S, g_J = 2$) and hence

$$\mu_{\text{eff}} = 2\mu_{\mathbf{B}}\sqrt{S(S+1)}. \tag{3.1}$$

As shown in Table 3.1, this produces a much better degree of agreement with experiment. This effect is known as orbital quenching and the orbital moment is said to be quenched. For 4 f ions, the orbitals are much less extended away from the nucleus and lie beneath the 5 s and 5 p shells so that the crystal field terms are much less important and Hund's third rule is obeyed. The situation in the higher transition metal ions (the 4 d and 5 d series) is less clear-cut because the heavier ions have a larger spin-orbit splitting and the effects of the crystal field and the spin-orbit interaction can be comparable.

For the case of the 3d ions where spin-orbit interactions can be effectively ignored, let us consider the orbital quenching effect in more detail. The crystal field in an octahedral environment is given by a constant plus a term proportional to $x^4 + y^4 + z^4 - \frac{3}{5}r^4 + O\left(r^6/a^6\right)$ (see Exercise 3.2), and is therefore a real function. This is typical of crystal-field Hamiltonians which can be expressed by real functions (without differential operators) and therefore the eigenfunctions of the Hamiltonian are all real. Now the total angular momentum operator $\hat{\mathbf{L}}$ is Hermitian and so has real eigenvalues but the operator itself is purely imaginary. [1] Thus if we have a non-degenerate ground state $|0\rangle$ (non-degenerate so that we can't play tricks by making linear combinations with it and another ground state) which is realized by the crystalfield splitting it must be a real function. Therefore $\langle 0|\hat{\mathbf{L}}|0\rangle$ must be purely imaginary because $\hat{\mathbf{L}}$ is purely imaginary. But since $\hat{\mathbf{L}}$ is Hermitian, $\langle 0|\hat{\mathbf{L}}|0\rangle$ must be purely real. It can only be purely real and purely imaginary if

$$\langle 0|\hat{\mathbf{L}}|0\rangle = 0 \tag{3.2}$$

[1] Equation C. 23 shows that the angular momentum operator is $\hat{\mathbf{L}} = -i\hat{\mathbf{r}} \times \nabla$ which is purely imaginary.

[2] This is to make states with azimuthal dependence such as ($\cos m_l\phi$) or $\sin(m_l\phi)$. Remember that

$$\cos(m_l\phi) = \frac{1}{2}\left(e^{\mathrm{i}m_l\phi} + e^{-\mathrm{i}m_l\phi}\right)$$
$$\sin(m_l\phi) = \frac{1}{2\mathrm{i}}\left(e^{\mathrm{i}m_l\phi} - e^{-\mathrm{i}m_l\phi}\right).$$

so that all components of the orbital angular momentum of a non-degenerate state are quenched.

This result can be seen in another way. Orbital states which are eigenstates of $\hat{L}_z$ with eigenvalue m_l have an azimuthal dependence of the form $\mathrm{e}^{\mathrm{i}m_l\phi}$. The requirement for real eigenfunctions implies that linear combinations of states with $\pm m_l$ must be formed, [2] so the resulting state has $\langle L_z\rangle = 0$. A semiclassical interpretation of orbital quenching is that the orbital angular momentum precesses in the crystal field, so that its magnitude is unchanged but its components all average to zero.

In fact the orbital angular momentum may not be completely quenched because the spin-orbit interaction is not completely ignorable, even in the 3d ions. The spin-orbit interaction can be included as a perturbation and in this case it can mix in states with non-zero angular momentum. This results in a quenched ground state but with a g-factor which is not quite equal to the spinonly value of 2 , the difference from 2 reflecting the mixed-in $L > 0$ states. It can also result in the g-factor being slightly anisotropic so that its value slightly depends on which direction you apply the magnetic field with respect to the crystal axes.

3.1.3 The Jahn-Teller effect

We have assumed so far that all we need to do is to work out what kind of symmetry the local environment has, then deduce the electronic structure

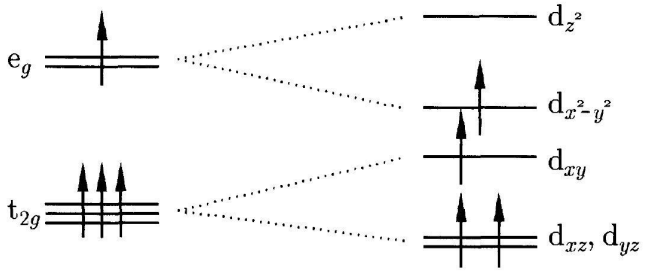
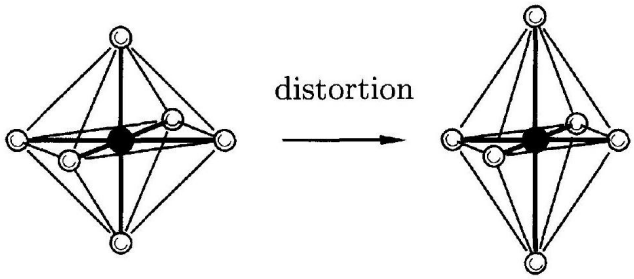


Fig. 3.6 The Jahn-Teller effect for $\mathrm{Mn}^{3+}(3\mathrm{d}^4)$. An octahedral complex (left) can distort (right), thus splitting the t_{2g} and e_g levels. The distortion lowers the energy because the singly occupied e_g level is lowered in energy. The saving in energy from the lowering of the d_{xz} and d_{yz} levels is exactly balanced by the raising of the d_{xy} level.}



and hence figure out the magnetic properties based on how many electrons there are to fill up the energy levels. Sometimes, however, the magnetic properties themselves can influence the symmetry of the local environment! This comes about because it can sometimes be energetically

favourable for, say, an octahedron to spontaneously distort as shown in Fig. 3.6 because the energy cost of increased elastic energy is balanced by a resultant electronic energy saving due to the distortion. This phenomenon is known as the JahnTeller effect. For example, Mn^{3+} ions (which have a configuration 3 d^4) in an octahedral environment show this kind of behaviour (see Fig. 3.6). In contrast, Mn^{4+} ions (3 d^3) would not show this effect because there is no net lowering of the electronic energy by a distortion.

To describe the effect, at least at the phenomenological level, we will assume that the distortion of the system can be quantified by a parameter Q, which denotes the distance of distortion along an appropriate normal mode coordinate. This gives rise to an energy cost which is quadratic in Q and can be written as

$$E(Q) = \frac{1}{2}M\omega^2Q^2, \tag{3.3}$$

where M and ω are respectively the mass of the anion and the angular frequency corresponding to the particular normal mode. This relation is plotted in Fig. 3.7(a). Clearly the minimum distortion energy is zero and is obtained when $Q = 0$ (no distortion).

The distortion also raises the energy of certain orbitals while lowering the energy of others. If all orbitals are either completely full or completely empty, this does not matter since the overall energy is simply given by eqn 3.3. However, in the cases of partially filled orbitals this effect can be highly significant since the system can have a net reduction in total energy. The electronic energy dependence on Q could be rather complicated, but one can write it as a Taylor series in Q and provided the distortion is small it is legitimate to keep only the term linear in Q. Let us therefore suppose that the energy of a given orbital has a term either AQ or $-AQ$ corresponding to a raising or a lowering of the electronic energy, where A is a suitable constant, assumed to be positive. Then the total energy $E(Q)$ is given by the sum of the electronic energy and the elastic energy

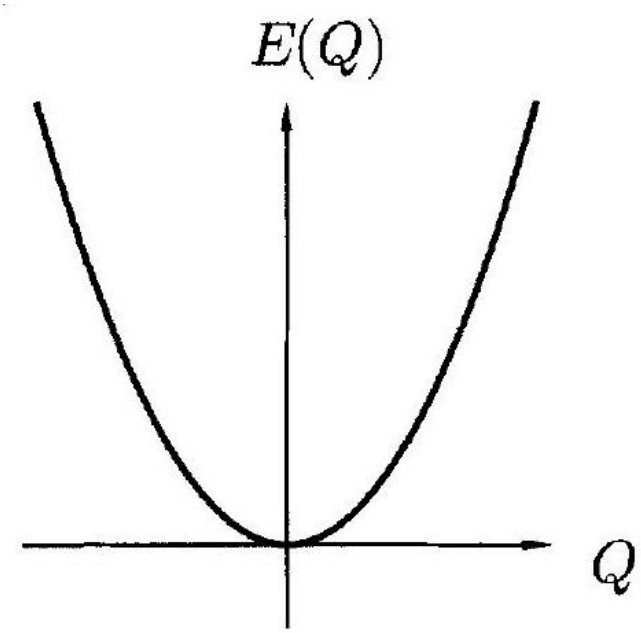
$$E(Q) = \pm AQ + \frac{1}{2}M\omega^2Q^2, \tag{3.4}$$

where the two possible choices of the sign of the AQ term give rise to two separate curves which are plotted in Fig. 3.7(b). If we consider only one of them we can find the minimum energy for that orbital using $\partial E/\partial Q = 0$ which yields a value of Q given by

$$Q_0 = \frac{A}{M\omega^2} \tag{3.5}$$

and a minimum energy which is given by $E_{\min} = -A^2/2M\omega^2$ which is less than zero. If only that orbital is occupied, then the system can make a net energy saving by spontaneously distorting.

What we have been considering so far is essentially a static Jahn-Teller effect because the distortion which can spontaneously occur is fixed on a particular axis of an octahedron. However, the distortion can switch from one
(a)



(b)

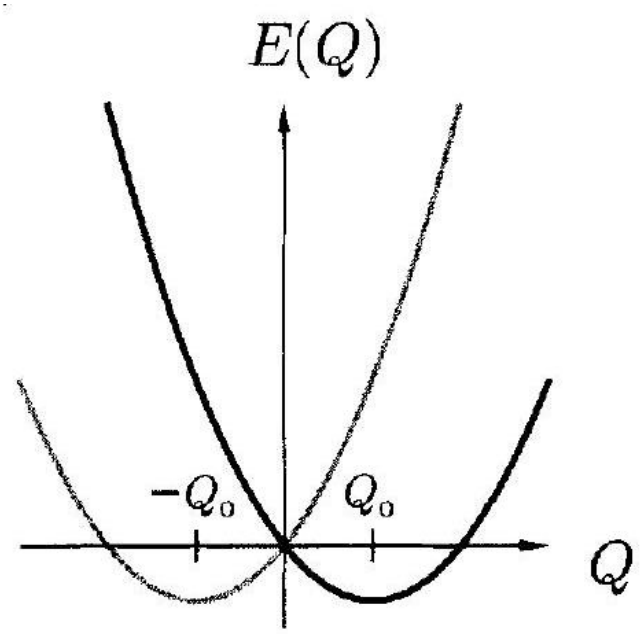


Fig. 3.7 (a) The energy of an octahedral complex as a function of the distortion Q according to eqn 3.3. (b) The energy of an octahedral complex as a function of the distortion Q according to eqn 3.4.

axis to another at higher temperatures, giving rise to a dynamic Jahn-Teller effect. Another type of dynamic effect involves rapid hopping of the distortion from site to site. This is important in, for example, materials which contain a mixture of Mn^{3+} and Mn^{4+} ions. Such phenomena can be detected by their effect on magnetic resonance data.

A further effect is that in certain materials (e.g. $DyVO_4$), below a certain critical temperature, a Jahn-Teller distortion of each complex can occur cooperatively throughout the crystal. This is known as a cooperative JahnTeller transition.

This concludes our discussion of the effect of the crystal field. In the rest of this chapter there follows a description of various experimental techniques which can be used to study the

interaction of a magnetic moment with its environment.

9.3.2 3.2 Magnetic resonance techniques

The local environment of a magnetic moment is determined by the crystal fields, but spin-orbit couplings and hyperfine interactions with the nuclei also play an important rôle in controlling the electronic structure. These effects can be studied by using a variety of experimental techniques which involve magnetic resonance. These experimental techniques will be described in the remainder of this chapter.

We have already seen in Section 1.1.2 how the application of a magnetic field to a magnetic moment can induce precession of that magnetic moment at an angular frequency given by $|\gamma B|$ where γ is the gyromagnetic ratio. A system of magnetic moments in a magnetic field can absorb energy at this frequency and thus one may observe a resonant absorption of energy from an electromagnetic wave tuned to the correct frequency. This is magnetic resonance and it can take a number of different experimental forms, depending on what type of magnetic moment is resonating. The appropriate frequency of the electromagnetic wave depends on the sizes of both the magnetic field and the magnetic moment. Figure 3.8 summarizes the range of electromagnetic radiation as a function of frequency for a range of different units. It can be used for reference and referred to throughout the following discussion as necessary.

3.2.1 Nuclear magnetic resonance

The most commonly used form of magnetic resonance is nuclear magnetic resonance (NMR). This technique is greatly employed in medical imaging, where it goes by the name of magnetic resonance imaging (MRI) to avoid the use of the dreaded word "nuclear". Some people may be devastated to discover that the average human body contains over 10^{27} nuclei; by mass we are $\sim 99.98\%$ nuclear! MRI measures NMR in the protons in a patient's body and provides a safe, non-invasive technique of yielding detailed cross-sectional images.

To perform any NMR experiment, one needs a nucleus with a non-zero spin. Nuclei which are commonly studied include $^{1}\mathbf{H}$ (proton), $^{2}\mathbf{H}$ (deuteron) and ^{13}C. In a simple NMR experiment (see Fig. 3.9), a sample is placed inside

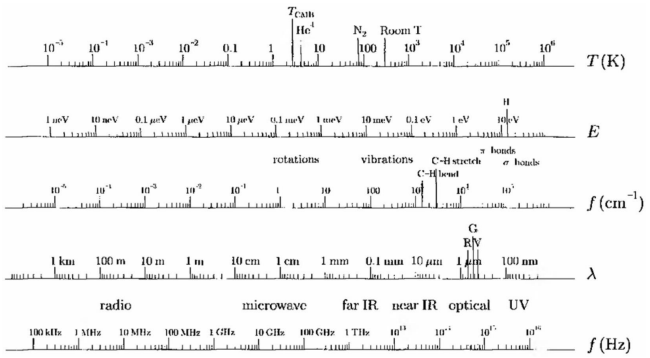


Fig. 3.8 The electromagnetic spectrum. The energy of a photon is shown as a tempcrature $T=E / k_{\mathrm{B }}$ in Kelvin and as an energy E in eV . The corresponding frequency f is shown in Hz and, because the unit is often quoted in spectroscopy, in cm^{-1}. The cm^{-1} scale is marked with some common molecular transitions and excitations (the typical range for molecular motations and vibrations are shown, together with the $\mathrm{C}-\mathrm{H}$ bending and stretching modes). The energy of typical π and σ bonds are also shown. The wavelength $\lambda = c/f$ of the pholon is shown (where c is the speed of light). The particular temperatures marked on the temperature scale are T_{CMB} (the temperature of the cosmic microwave background), the boiling points of liquid Helium (He^4) and nitrogen (N_2), both at atmospheric pressure. and also the value of room temperaturc. Other abbreviations on this diagram are IR = infrared, UV = ultraviolet. $\mathrm{R} = \mathrm{rcd}, \mathrm{G} = \mathrm{green}$. V = violet. The letter H marks 13.6 eV . the magnitude of the energy of the 1 s electron in hydrogen. The frequency axis also contains descriptions of the main regions of the electromagnetic spectrum: radio. microwave, infrared (both 'near' and 'far'), optical and UV.}

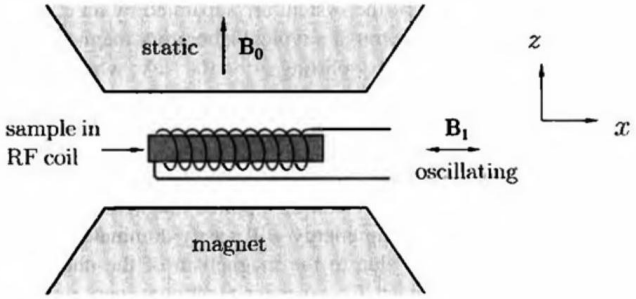


Fig. 3.9 Schematic diagram of an NMR experiment. The sample sits inside a radio frequency (RF) coil which produces an oxcillating RF field. A highly homogenous static magnetic ficld is provided by a magnet. The static field B_{0} and the oscillating field B_1 are perpendicular to each other. In a real experiment, the sample would be much smaller than is shown here so that it experiences a uniform field from the RF coil.}

a coil which is mounted between the pole pieces of a magnet. The magnet produces a magnetic field $\mathbf{B}_0$ along a particular direction, say the z-direction. We have already seen that the quantity m_l, the z component of the angular momentum of the nucleus, can only take integral values between $-I$ and I.

[3] The positive charge of the nuclei and the negative charge of the electron is responsible for the differing signs in the expressions:

$$\mu = +g_{\rm N}\mu_{\rm N}\mathbf{I}$$
$$\mu = -g_J\mu_{\rm B}\mathbf{J}$$

for the nuclear and electronic moments respectively.

[4] For the time being, we set $J = 0$ and thus ignore the electron spin.

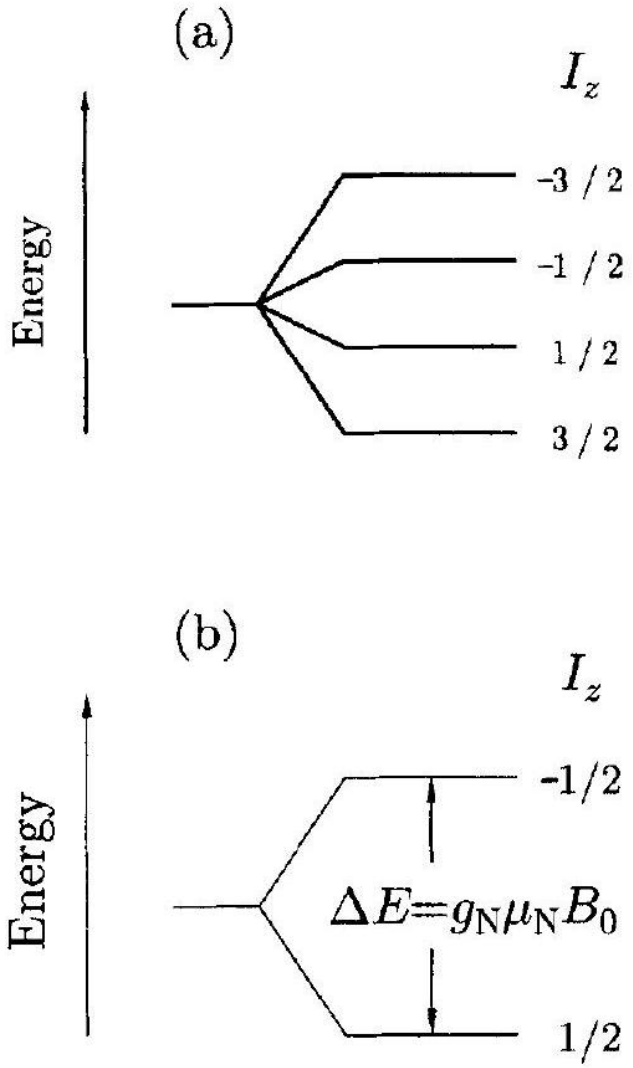


Fig. 3.10 (a) The four levels from an $I=\frac{3}{2}$ nucleus. (b) In the simpler case of an $I = \frac{1}{2}$ nucleus, there are just two levels. These two nuclear levels are separated by an energy $\Delta E = g_{\rm N}\mu_{\rm N}B_0$; the lower (upper) level corresponds to the nuclear magnetic moment lying parallel (antiparallel) to the magnetic field B_0.}

The energy E of the nucleus is the energy of a magnetic moment [3]μ in the magnetic field $\mathbf{B}_0$, and thus $E = -\mu \cdot \mathbf{B}_0$ so that

$$E = -g_{\rm N}\mu_{\rm N}m_I B_0. \tag{3.6}$$

This corresponds to a ladder of $2I+1$ equally spaced levels with mutual separation $g_{\rm N}\mu_{\rm N}B_0$, as illustrated in Fig. 3.10(a) for the case $I = \frac{3}{2}$, as found in Na or Cu.[4] Exciting transitions

[05] Actually you do not have to get it exactly right. You can be plus or minus the linewidth which can be determined by various factors, including variations in the magnetic field at the nucleus due to the dipolar field produced by neighbouring nuclei.

between adjacent pairs of levels with a radiofrequency (RF) field is the basis of nuclear magnetic resonance. The RF field $\mathbf{B}_1$ is applied in the x-direction, and leads to a perturbation of the system which is proportional to $B_1\hat{I}_x$. The matrix element of the perturbation is proportional to $\langle m_I'|\,\hat{I}_x\,|m_I\rangle$ and is zero unless $m_I' = m_I \pm 1$. The allowed transitions are therefore described by the selection rule,

$$\Delta m_I = \pm 1, \tag{3.7}$$

which implies that only transitions between adjacent states may occur. In the most common arrangement, the RF coil not only produces the excitation, but is itself part of a tuned circuit with a large Q factor. As transitions are excited in the nuclei by the RF field, energy is transferred between the RF circuit and the sample, and this results in small changes of the Q factor of the circuit.

There are two ways to perform the experiment: you could keep the frequency of the RF field constant and vary the magnetic field or you could keep the magnetic field constant and vary the frequency of the RF field. It is usually the former which is performed. A crucial factor is to have a highly homogeneous magnet to produce the constant field B_0. If this is not the case, different parts of the sample will sit in slightly different magnetic fields and will come into resonance at slightly different points of the magnetic field sweep, causing the measured resonance to be extremely broad, maybe so much that it is washed out altogether.

Example 3.2

For a proton, $I = \frac{1}{2}$ and m_I can take the values $\pm\frac{1}{2}$ only, as illustrated in Fig. 3.10(b). The two states of the system are separated by an energy $\Delta E = g_{\mathrm{N}}\mu_{\mathrm{N}}B$ which is tiny: for a proton in a typical laboratory magnetic field $B_0 \sim 1$ T this corresponds to an energy splitting of $\sim 10^{-7}$eV; when this energy is expressed as $k_{\mathrm{B}}T$ it is equivalent to a temperature of only ~ 1mK.

Therefore at room temperature, at this magnetic field, the nuclei will show only a minute tendency to line up with the applied magnetic field on average because the thermal randomizing energy will vastly dominate over the weak alignment energy. Any effect due to the magnetism of the nuclei would be practically impossible to detect if a resonant technique were not employed. Because in NMR one is perturbing the system with a RF signal of exactly [5] the right frequency ω given by

$$\hbar\omega = \Delta E \tag{3.8}$$

we can excite transitions across this energy splitting. Why do we need radiofrequencies? The next example shows why.

Example 3.3

For a magnetic field of 1 T , the frequency v is given by

$$v = \frac{\omega}{2\pi} = \frac{g_{\mathrm{N}}\mu_{\mathrm{N}}B_0}{2\pi\hbar} = 42.58\mathrm{MHz} \tag{3.9}$$

which is in the radiofrequency region of the electromagnetic spectrum.

This means one can use oscillators and coils (in comparison with electron spin resonance experiments which have to use Gunn diodes and microwave waveguides, see Section 3.2.2). Therefore at exactly the right frequency of RF excitation the system will absorb energy. This is the nuclear magnetic resonance and is a resonance between the frequency of the RF excitation

and the energy separation of the nuclear levels. Modern NMR spectrometers use 12-15 T magnets, so the frequency of the RF is in the $\sim 500-650$MHz range, still in the radio frequency region of the electromagnetic spectrum.

The NMR resonance frequency at a given magnetic field can be slightly shifted up and down from the value $\gamma B/2\pi$ depending on the chemical environment of the nucleus. These chemical shifts (typically a few parts per million) are due to the fact that the electrons orbiting the nucleus slightly shield the nucleus from the applied field. The amount by which a nucleus is shielded in a given chemical environment is well known, allowing a given molecule to be 'fingerprinted'. NMR lines can split due to magnetic coupling to neighbouring nuclei. This is known as spin-spin coupling and occurs via an indirect contact hyperfine interaction which is mediated by the electrons. This also gives information about the environment of the nucleus.

In order to understand the transitions which we are inducing in a little more detail, let us consider how a two-level spin system ($I = \frac{1}{2}$) will absorb energy from the RF coil. We will label the lower level - and the upper level + . The probability of inducing a transition between the two levels is independent of whether the transition is from the lower to the higher level or vice versa. This is because the perturbing Hamiltonian $\hat{\mathcal{H}} \propto \hat{I}_x$, appropriate for an oscillating RF field, is Hermitian, so that

$$\left|\langle\psi_-|\,\hat{\mathcal{H}}\left|\,\psi_+\right\rangle\,\right|^2 = \left|\langle\psi_+|\,\hat{\mathcal{H}}\left|\,\psi_-\right\rangle\,\right|^2. \tag{3.10}$$

Hence the probability per unit time of these so-called stimulated transitions between levels + and - is independent of the direction of the transition and occurs at a rate W which is proportional to the size of the RF power used to excite transitions. At time t, if there are $N_-(t)$ spins in the lower level, $WN_-(t)$ will be excited per unit time into the upper level. If there are $N_+(t)$ spins in the upper level, $WN_+(t)$ will be excited per unit time into the lower level. Thus

$$\frac{\mathrm{d}N_+(t)}{\mathrm{d}t} = WN_-(t) - WN_+(t) \tag{3.11}$$

$$\frac{\mathrm{d}N_-(t)}{\mathrm{d}t} = WN_+(t) - WN_-(t) \tag{3.12}$$

which when subtracted gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(N_+(t) - N_-(t)\right) = -2W\left(N_+(t) - N_-(t)\right), \tag{3.13}$$

which can be solved to give

$$N_+(t) - N_-(t) = \left(N_+(0) - N_-(0)\right)\mathrm{e}^{-2Wt} \tag{3.14}$$

An initial difference in population tends exponentially to zero when driven by a stimulated electromagnetic transition. Let us now define $n(t)$ by

$$n(t) = N_+(t) - N_-(t), \tag{3.15}$$

so that eqn 3.14 becomes

$$n(t) = n(0)\mathrm{e}^{-2Wt} \tag{3.16}$$

The rate of absorption or emission of electromagnetic energy by the spin system is easily calculated since each transition involves an exchange of energy equal to $\hbar\omega$. The energy of the system at time t is

$$E(t) = N_- E_- + N_+ \left(E_- + \hbar\omega\right) \tag{3.17}$$

where $E_-$is the energy of the lower level. This expression for $E(t)$ can be rearranged to give a constant term plus $\frac{1}{2}\hbar\omega n(t)$. Thus the rate of absorption of energy is given by

$$\frac{\mathrm{d}E}{\mathrm{d}t} = -W\hbar\omega n(t) \tag{3.18}$$

and, because it is proportional to $n(t)$, will tend to zero with a time constant of $1/2W$ as the populations of the upper and lower levels become progressively equalized. This demonstrates also that to absorb energy you need $n(t) \neq 0$, i.e. a population difference. This is produced by interaction with the thermal modes of the system, so that the polarization of the system is brought back towards a Boltzmann probability expression:

$$\left(\frac{N_+}{N_-}\right)_0 = \mathrm{e}^{-\hbar\omega/k_\mathrm{B}T} \tag{3.19}$$

Here, the subscript 0 denotes the equilibrium value of the system. After absorption of energy the polarization may change, but if the nuclear spin system interacts with the thermal motion and excitations of the sample, the polarization of the spin system can return towards its equilibrium value. If the polarization of the spin system has been reduced to zero by rapid electromagnetic transitions, the spin system will take some time to 'recover' once the electromagnetic transitions are switched off; this time is called T_1 and measures the time constant of the interaction between the spin system and its surroundings. T_1 is the spin-lattice relaxation time (sometimes called the longitudinal relaxation time). Thus we would expect that the polarization would be restored as

$$n(t) = n_0 \left(1 - \mathrm{e}^{-t/T_1}\right). \tag{3.20}$$

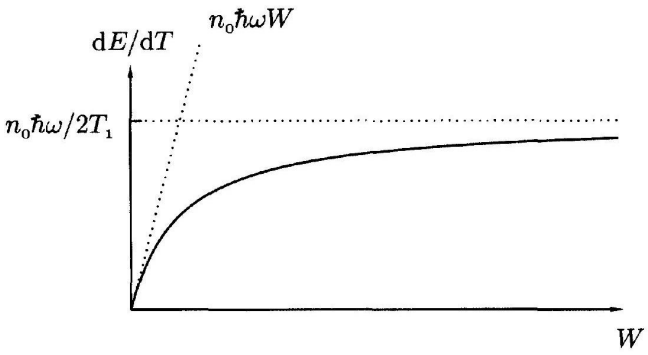


Fig. 3.11 The rate of absorption of electromagnetic energy given by eqn 3.25 as a function of the transition rate W. For low W, the expression is proportional to W, but at high W it saturates at a value governed by $1 / T_{1}$.

Now let us put both processes in together, the stimulated transitions and the relaxation processes. Thus we may combine our equations to arrive at

$$\frac{\mathrm{d}}{\mathrm{d}t}n(t) = -2Wn(t) + \frac{n_0 - n(t)}{T_1}, \tag{3.21}$$

which has a steady state solution when

$$\frac{\mathrm{d}}{\mathrm{d}t}n(t) = 0. \tag{3.22}$$

Thus

$$n(t) = \frac{n_0}{1 + 2WT_1}, \tag{3.23}$$

and the steady state rate of absorption of electromagnetic energy is given by

$$\frac{\mathrm{d}E}{\mathrm{d}t} = n(t)\hbar\omega W \tag{3.24}$$

$$= n_0\hbar\omega\frac{W}{1 + 2WT_1}. \tag{3.25}$$

This relation is plotted in Fig. 3.11. For low-amplitude RF perturbing fields, the rate of absorption of electromagnetic energy is proportional to W. But at large RF perturbing fields, it settles at a level proportional to T_1, and independent of the precise magnitude of W. This is known as saturation.

The rate of absorption of electromagnetic energy is proportional to the population difference between the upper and lower levels. The rate of absorption may therefore be used to monitor the population difference and to observe its time dependence, subject of course to the constraint that such observation will necessarily perturb the population difference to some extent. Equation 3.25 and Fig. 3.11 demonstrate that you can only increase the power of the RF signal so much to improve the signal. This also demonstrates that $\mathrm{d}E/\mathrm{d}t$ is proportional to $\hbar\omega$; more energetic (i.e. higher frequency) photons lead to a larger signal. But of course this means larger magnetic fields to match $g\mu_N B$ to $\hbar\omega$. This will prove to be an advantage for ESR (see Section 3.2.2 below) since very beefy microwave photons are needed to excite transitions across the larger electronic
gaps, and this leads to a higher sensitivity. Because electronic moments are larger than nuclear moments, you can get away with smaller magnetic fields than NMR and still use higher frequency photons.

Let us return to NMR and consider relaxation processes in a little more detail. In a magnetic field a very weak polarization of the spins will exist (weak because the energy separation $\Delta E \ll k_\mathrm{B}T$.) Let us say it takes an equilibrium value M_0 with the static magnetic field $\mathbf{B}_0$ switched on (recall that $\mathbf{B}_0$ is along the z axis). The effect of the radio frequency excitations will be to progressively destroy this magnetization. Also, as we have seen, if the magnetic field is switched off, the magnetization will relax back to the equilibrium value by the weak interactions with the surroundings with a time constant T_1. Thus we expect that

$$\frac{\mathrm{d}M_z}{\mathrm{d}t} = \frac{M_0 - M_z}{T_1} \tag{3.26}$$

This T_1 relaxation must involve interactions with the lattice because energy must be exchanged with it. This is because changing M_z has energetic consequences since relaxing a spin in the z-direction changes its orientation with respect to the applied field, and hence changes its energy.

The M_x and M_y components should be zero, but if they are not they will relax back to zero in a time T_2 such that

$$\frac{\mathrm{d}M_x}{\mathrm{d}t} = -\frac{M_x}{T_2}, \quad \frac{\mathrm{d}M_y}{\mathrm{d}t} = -\frac{M_y}{T_2} \tag{3.27}$$

T_2 is the spin-spin relaxation time and is the characteristic time for dephasing because it corresponds to the interaction between different parts of the spin system. It can also be due to inhomogeneities in the magnetic field B_0. It thus leads to differences in precession frequency due to the interactions between the observed spin and the spins of its neighbours. Changing M_x or M_y has no energetic consequences because the applied field is along M_z.

The effect of the applied magnetic field is also to cause spin precession, so that the equations for M_x, M_y and M_z are

$$\frac{\mathrm{d}M_x}{\mathrm{d}t} = \gamma(\mathbf{M} \times \mathbf{B})_y - \frac{M_x}{T_1} \tag{3.28}$$

$$\frac{\mathrm{d}M_y}{\mathrm{d}t} = \gamma(\mathbf{M} \times \mathbf{B})_y - \frac{M_y}{T_1} \tag{3.29}$$

$$\frac{\mathrm{d}M_z}{\mathrm{d}t} = \gamma(\mathbf{M} \times \mathbf{B})_z + \frac{M_0 - M_z}{T_1} \tag{3.30}$$

which are known as the Bloch equations. They were worked out in 1946 and can be solved for a number of cases of interest, very often by using a rotating reference frame method in which the coordinates are changed to ones which rotate in the xy plane at the resonance frequency.

The spin-spin relaxation time T_2 can be measured using the following technique: in thermal equilibrium with B_1 switched off, there is a weak magnetization parallel to z. A short pulse of RF signal of duration t_p, with the frequency of the RF signal set close to the resonance frequency of the spins, causes the spins to rotate by an angle $\gamma B_1 t_\mathrm{p}$ where B_1 is the amplitude of the RF signal. This angle can be adjusted by changing t_p and if $t_\mathrm{p} = \pi/2\gamma B_1$, it produces what is known as a 90° pulse. The amplitude and time of the pulse are chosen

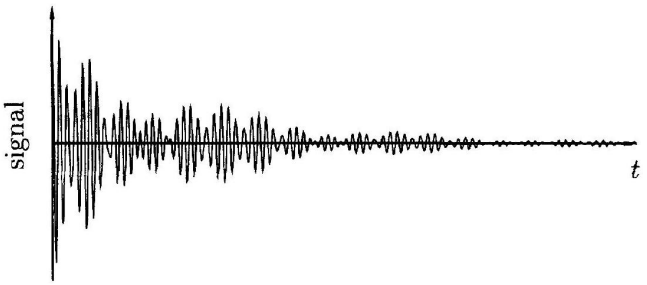


Fig. 3.12 A simulated free induction decay signal. The oscillations die away with time due to spin-spin relaxation ($T_{2
$ processes). The data show an interference pattern which results from three inequivalent nuclei, each of which experiences a slightly different magnetic field because of their different sites within a crystal.}

so that the magnetization just precesses for a short period and rotates into the xy plane. With B_1 switched off, the magnetization now precesses in the static field $\mathbf{B_0}$, rotating steadily in the xy plane at a frequency γB_0, producing an induced voltage in the coil. The oscillation would carry on for ever if it were not for T_2 relaxation processes (spin-spin relaxation), which cause the oscillations to relax, thereby giving a measurement of T_2. The corresponding decay of voltage in the coil is known as free induction decay (see Fig. 3.12).

One problem that occurs is that this method also measures relaxation due to the inhomogeneity of the magnet because this will cause different parts of the sample to give slightly different precession frequencies. What is needed is a method to separate the real spin-spin relaxation from the relaxation due to the inhomogeneity of the magnet. One way this can be

achieved is to use a spin echo technique. The principle of this technique is illustrated in Fig. 3.13 and involves both a 90° pulse (which tips the spins by 90° into the xy plane as before) and the subsequent application of a 180° pulse (which rotates the spins by 180°). The effect can be likened to a marathon in which all athletes run steadily but some athletes are faster than others. After the starter's pistol (the 90° pulse) the athletes begin to run as a pack, but after a while the fitter begin to lead while others begin to trail behind. Then suddenly, at time τ, the rules of the game are changed (the 180° pulse) and the runners are told to turn around and head for the start of the race. Now the slower runners have an advantage since they are closer to the start. But by symmetry it is easy to see that the runners will all arrive at the start of the race simultaneously, at time 2τ, regardless of their individual speeds. In the same way, all the spins realign after 2τ, regardless of their individual dephasing. This is the spin echo.

Thus in the experiment we can use the spin echo to remove the problem that the magnetic field is inhomogeneous causing spins to precess at different rates. Also the spread in fields due to chemical shifts is removed. But we will not have taken away the effect of the spin-spin interaction which is due to time-dependent fluctuating random magnetic fields due to neighbouring nuclei. This relaxation mechanism cannot be refocussed and so the echo signal will be reduced in amplitude by an amount which depends on τ. The NMR intensity of the echo signal should follow $I(2\tau) = I(0)\mathrm{e}^{-2\tau/T_2}$ and so a measurement of the true T_2, due to the spin-spin coupling, is possible.

A measurement of T_1 can be performed using a similar method, though this time a 180° pulse is applied first. This causes the magnetization to rotate from

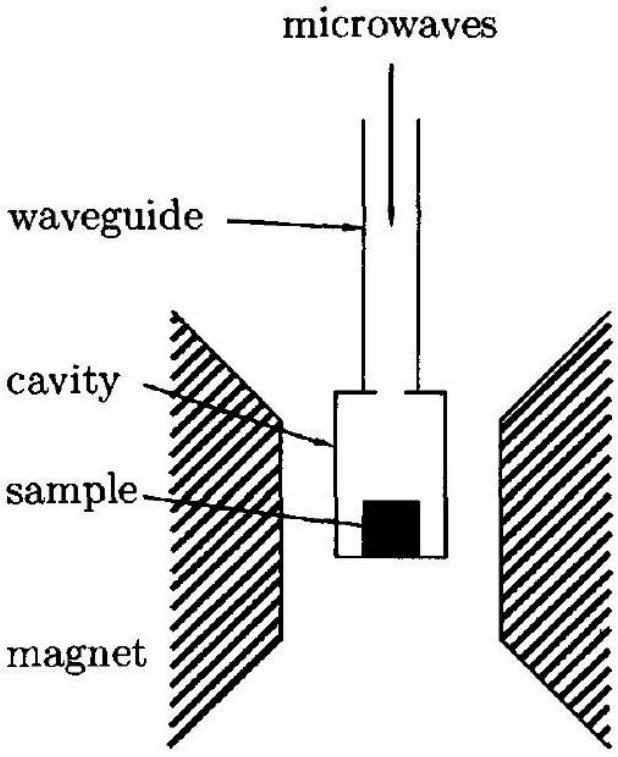


Fig. 3.14 Schematic diagram of an ESR experiment. Microwaves enter a cavity via a waveguide and the absorption of microwaves induced by a resonance is measured by monitoring the Q-factor of the cavity. The sample must be placed in the centre of the magnet, where the field is most clearly uniform.

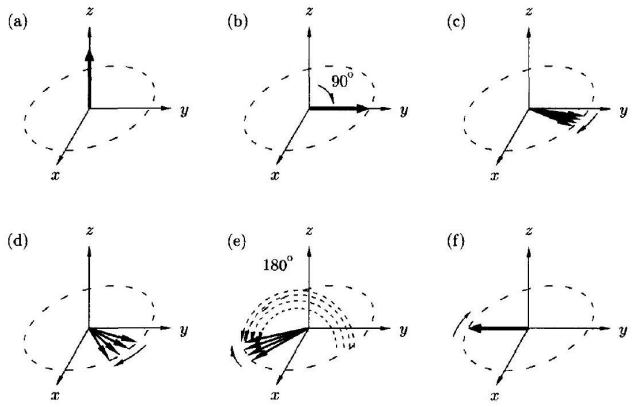


Fig. 3.13 The spin echo effect. (a) The equilibrium magnetization initially lies along the z direction, parallel to $B_{0
$. (b) At a time which we will call $t = 0$, a 90° pulse along the x axis is used to rotate the spins into the xy plane. (c) Because of the steady field B_0 along the z axis, the spins now precess in the xy plane, though at slightly different rates because of inhomogeneities in the magnetic field. (d) This means they progressively dephase with respect to each other, becoming more separated in angle as they precess round. (e) At time $t = \tau$, a 180° pulse along the x axis rotates the spins by 180° around the x axis, and as they subsequently precess in the field B_0, their order has been reversed. (f) They then will come back together at time 2τ, producing a spin echo signal, provided no other relaxation process has occurred in the meantime. Although the figure has been drawn for the case of a short delay between the 90° pulse and the 180°, the effect will work if the delay is long and the spins make a number of rotations before the 180° pulse is applied.}

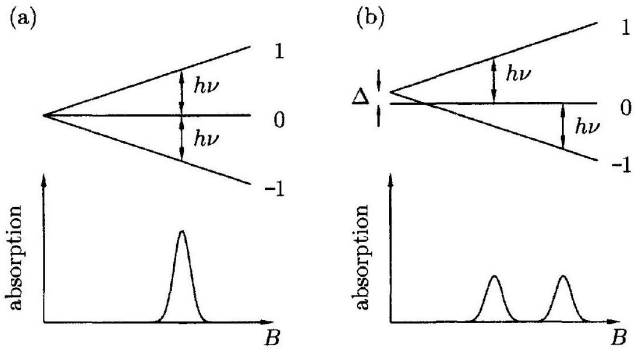
along $\hat{\mathbf{z}}$ to along $-\hat{\mathbf{z}}$. It then relaxes back to $\hat{\mathbf{z}}$ with a time constant T_1 but does not precess. Thus as a function of time τ after the 180° pulse

$$\mathbf{M}(\tau) = \hat{\mathbf{z}} M_z(\tau) = \hat{\mathbf{z}} M_0 \left(1 - 2\mathrm{e}^{-\tau/T_1}\right). \tag{3.31}$$

At a time τ after the 180° pulse, a 90° pulse is used to rotate the magnetization into the xy plane where it now begins to do a free induction decay with initial amplitude $M_z(\tau)$. Thus by measuring the initial amplitude of the free induction decay as a function of the time delay between the 180° pulse and the 90° pulse, T_1 can be deduced. A variety of other pulse sequences can also be used to extract additional types of information.

3.2.2 Electron spin resonance

One can perform the analogous experiment to NMR with electrons. In this case the effect is known as electron spin resonance (ESR), or sometimes as electron paramagnetic resonance (EPR). The electron magnetic moment is much larger than that of nuclei so the precession frequencies are also much higher. For typical laboratory magnetic fields, the electromagnetic radiation is now in the microwave regime. A schematic of the experimental setup is shown in Fig. 3.14. The sample sits in a resonant cavity and microwave radiation enters via a waveguide. It is usually much more convenient to keep the frequency of



the microwave radiation fixed and to sweep the magnetic field. The cavity has a very high quality factor (Q-factor) and is designed to enhance the sensitivity of the detection of the weak ESR signal. The microwave absorption is modified as a function of the applied magnetic field. Very often a set of modulation coils is added to provide a small oscillatory magnetic field at the sample in addition to the static magnetic field provided by the magnet. The measurement can then be further enhanced by a phase sensitive detection technique. [6]

In an ESR transition the selection rule is

$$\Delta m_J = \pm 1, \tag{3.32}$$

the familiar dipole selection rule because one is inducing magnetic dipole transitions. Figure 3.15 shows these transitions in an ion with $J = 1$ such as Ni^{2+} in an octahedral environment. The experiment is performed at fixed frequency hv with the magnetic field swept. Every time two adjacent levels are separated by $h\nu$, there is an ESR transition. Figure 3.15 demonstrates how the position and the number of lines can give information about the crystal field splitting.

Figure 3.16 shows the energy of a free Mn^{2+} ion (spin- $\frac{5}{2}$) as a function of magnetic field. At a magnetic field B this gives a number of possible transitions between the different m_J levels, all of which have size $g\mu_B B$. Thus a single line might be expected to be observed in an ESR experiment at a frequency v given by $hv = g\mu_{\mathrm{B}} B$. However, in a real experiment on a solid containing Mn^{2+} ions, the crystal field significantly complicates this picture, even in a cubic environment.

However, the hyperfine coupling between the nucleus and the electron gives a term in the Hamiltonian equal to $A\mathbf{I} \cdot \mathbf{J}$ and each m_J level is split into $2I+ 1$ hyperfine levels depending on the value of m_I. In an ESR transition the selection rules are therefore given by eqn 3.32 and by

$$\Delta m_I = 0. \tag{3.33}$$

Fig. 3.15 ESR in an ion with $J = 1$ such as Ni^{2+} with (a) no crystal field splitting (leading to a single ESR line) and (b) crystal field splitting Δ (leading to two ESR lines).

[6] A lock-in amplifier is used to produce a signal which drives the oscillatory magnetic field. The microwave signal is then fed into the lock-in amplifier, which rejects any component of this signal which is not at the frequency of the oscillatory magnetic field. In this way the noise is drastically reduced.

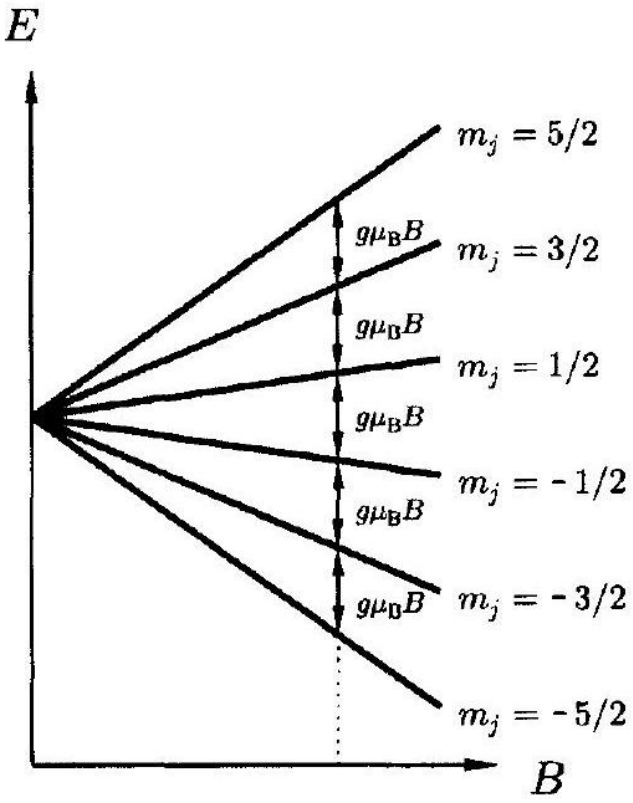


Fig. 3.16 The energy of a spin- $\frac{5
{2}$ magnetic moment, such as a free Mn^{2+} ion, as a function of magnetic field. The possible transitions with the selection rule $\Delta m_J = \pm 1$ are as shown. This picture ignores the crystal field.}

[7] This is because the energy spacing between the nuclear levels is very small in comparison with the energy of the microwave photons

The latter selection rule occurs because the microwave frequency induces only dipole transitions between electronic levelsnot between nuclear levels[7] Since A is very small in comparison with $g\mu_B B$ for typical fields in an ESR experimentone need only compute the components of **J** parallel to **B** so that the energy levels in a magnetic field are

$$E = g\mu_B B m_J + A m_I m_J \tag{3.34}$$

where the energy of the nuclear spin in the applied magnetic field has been ignoredsince it is much smaller even than the hyperfine energyUsing the selection ruleone can now show that the resonances should occur at

$$\begin{aligned} hv &= [g\mu_B B\,(m_J + 1) + A m_I\,(m_J + 1)] - [g\mu_B B m_J + A m_I m_J] \\ &= g\mu_B B + A m_I \end{aligned} \tag{3.35}$$

and each ESR line will split up into $2I + 1$ hyperfine lines

Example 3.4

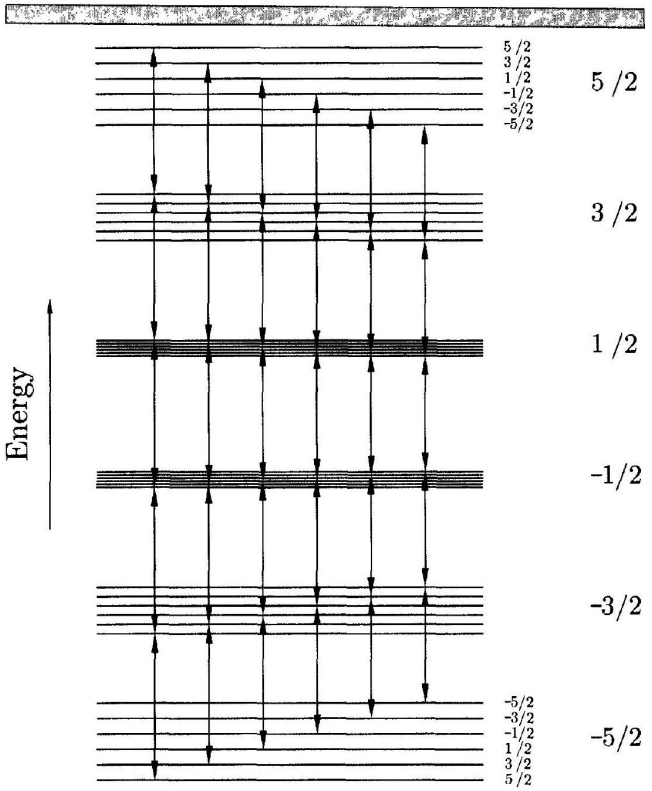
For the case of a free Mn^{2+} ion$I = \frac{5}{2}$ so this means six lines in the ESR spectrumThis situation is depicted in Fig 3.17 although for a Mn^{2+} ion in a crystalthe crystal field causes

the splittings to be more complex and to depend on the angle between the magnetic field and the crystal axes in a nontrivial way

ESR experiments probe electronic spins in a useful range because crystal field splittings are very often at GHz frequenciesESR studies in paramagnetic salts have provided a great deal of detailed information concerning the crystal and ligand fieldsVery often the effect is anisotropic so that lines move as the magnetic field is rotated with respect to the crystal axis

Example 3.5

As an example of how this information can be obtainedconsider Ce^{3+} $(4f^1)$ in an environment with axial symmetrysee Fig318This ion has $L = 3$ and $S = \frac{1}{2}$ so the spinorbit interaction splits this into an eightfold degenerate $J = \frac{7}{2}$ level and a lower sixfold degenerate $J = \frac{5}{2}$ The crystal field is axial and contains only even powers of J_z by symmetryso the sixfold $J = \frac{5}{2}$ manifoldas a group of lines are calledsplits up into three doubletsand the eightfold $J = \frac{7}{2}$ manifold into fourThese splittings are much smaller than the spinorbit interaction in rare earth systemsThe doublets can be split by applying a magnetic field which is almost always the smallest energy term for typical laboratory magnetic fieldsusually around 10^{-4}eV At low temperatureonly the ground state will be occupied and it is this transition which will be observed



The ground state of many systems is, like Ce^{3+}, a doublet. This is a consequence of Kramers theorem, which states that in a system containing an odd number of electrons, at least two-fold degeneracy must remain in the absence of a magnetic field. The pairs of states involved, Kramers doublets, are time conjugate [8] and therefore can be split by a magnetic field but not by an electric field. Ce^{3+} $\left(4f^1\right)$ has an odd number of electrons in the 4 f shells and therefore qualifies as a Kramers ion.

The ground state of Ce^{3+} is a Kramers doublet and, viewed in isolation, looks like a spin- $\frac{1}{2}$ state, although its quantum numbers are very different. It can in fact be assigned an effective spin $\tilde{S} = \frac{1}{2}$. This is a fictitious angular momentum which is chosen so that the degeneracy of the group of levels under consideration is set equal to $(2\tilde{S}+1)$. The motivation for this substitution

is to find an effective spin Hamiltonian that approximately describes the experimental situation in the language of something much simpler, namely a free atom or ion with spin- $\frac{1}{2}$. This Hamiltonian can be written

$$\tilde{H} = g\mu_{\mathrm{B}}\mathbf{B}\cdot\tilde{\mathbf{S}} \tag{3.36}$$

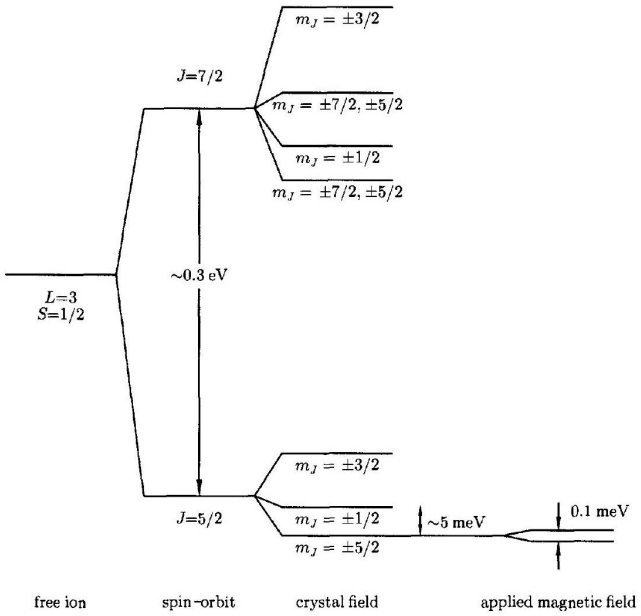
Fig. 3.17 The hyperfine splitting in a free Mn^{2+} ion at a fixed magnetic field. $J = \frac{5}{2}$ and $I = \frac{5}{2}$ so each of the six m_J levels is split into six hyperfine levels. The possible transitions with $\Delta m_J = \pm 1$ and $\Delta m_I = 0$ are shown, giving rise to six distinct lines in the ESR spectrum. This picture ignores the effect of the crystal field which leads to these splittings becoming strongly dependent on the direction of the applied magnetic field with respect to the crystalline axes.

Hendrik A. Kramers (1894-1952)

[8] that is, they are complex conjugates of each other, and are thus time reversed versions of each other.

Fig. 3.18 The energy levels for a $\mathrm{Ce

ˆ{3+}$ ion. The dominant energy splitting is the spinorbit interaction (this energy corresponds to about 3000 K so the upper manifold is not appreciably thermally excited). The two manifolds are split into sets of doublets by an assumed axial crystal field. The crystal field splittings have been exaggerated in this diagram so that they are easily observable. The lower splitting is $\sim$ 5meV which corresponds to about 50 K . Finally an applied magnetic field splits all the doublets (only the lower one is shown split here).}



where g is now an effective g-factor (sometimes called a spectroscopic splitting factor) which would be equal to 2 for a real isolated spin- $\frac{1}{2}$ electron but includes the orbital contribution in this case. Very often one needs to use an effective g-tensor, g, to express the fact that the interaction of the spin with the field depends on the orientation of the magnetic field, so that eqn 3.36 must be replaced with

$$\tilde{H} = \mu_{\mathrm{B}}(\mathbf{B}\cdot\ \mathrm{g}\cdot\tilde{\mathbf{S}}) = \sum_{ij}\ \mathrm{g}_{ij}B_i\tilde{S}_j \tag{3.37}$$

The effect of the crystal field can then be included by adding in a term to the Hamiltonian to represent the energetic preference for the spin to lie along particular crystalline directions

because of the crystal field. This is known as single ion anisotropy and the term in the Hamiltonian, $\tilde{H}_{\rm SI}$, is

$$\tilde{H}_{\rm SI} = -DS_z^2, \tag{3.38}$$

for a uniaxial crystal (i.e. one with a particular axis such that the energy just depends on the angle of the spin with that axis) and for a cubic crystal

$$\tilde{H}_{\rm SI} = -D\left(S_x^4 + S_y^4 + S_z^4\right) \tag{3.39}$$

where D is an anisotropy constant.
ESR is useful not only in studying the physics of salts with paramagnetic ions but is used extensively in chemistry. The resonances can also be extremely
sensitive to atomic position in a molecule and the technique is greatly used in the study of chemical reactions and particularly of free radicals, which are atoms or fragments of molecules with an unpaired electron and are often highly reactive. Free radicals are of great environmental and biological importance. ESR is not as widely used as NMR because it is applicable only to materials with an unpaired spin.

Example 3.6

One example of a field in which ESR is very useful is in studies of various biological molecules which often contain transition metal ions. An important example is haem, an iron porphyrin group, which is found in haemoglobin and myoglobin. These molecules are involved in transfer of oxygen and other small molecules such as carbon monoxide. The Fe (II) ion in the haem converts from $S = 2$ to $S = 0$ after binding to the small molecule (this is an example of a spin transition, see Section 3.1.1).

The spin-lattice relaxation times in ESR are often very short in comparison with those of NMR; this is because the electronic moment is much more strongly coupled to lattice vibrations than are the nuclei. Lattice vibrations cause a modulation in the crystal field which couples to the electronic moment via the spin-orbit interaction, often making T_1 so short that the resulting ESR line is too broad to observe. (The width of the line scales as T_1^{-1}). A favourable case is Mn^{2+} which is an S-state ion ($3\,d^5, {}^6S_{5/2}$) and therefore has no orbital moment to which the spins can couple and ESR can be observed easily in this ion at room temperature. For other transition metal ions, the size of T_1 depends on the degree of orbital quenching. [9] The resonance can be narrowed (T_1 increased) by cooling the sample, since this reduces the lattice vibrations. This also has the benefit of increasing the relative thermal population of the levels which increases the intensity of the resonance.

A variant of ESR is a technique called ENDOR (an acronymn for Electron Nuclear DOuble Resonance) which can be used to measure hyperfine interaction constants and the nuclear Zeeman interaction with great precision. It is often impossible to measure directly NMR transitions between nuclear levels in paramagnetic salts because they are very weak. In ENDOR one uses both an RF signal and a microwave signal and capitalizes on the high sensitivity of ESR. The microwave power is tuned to a $\Delta m_I = 0$ transition and the ESR signal then depends on the relative population of those levels. The populations can be changed by using RF power tuned to the NMR transition. The NMR transition can then be detected via its effect on the ESR signal.

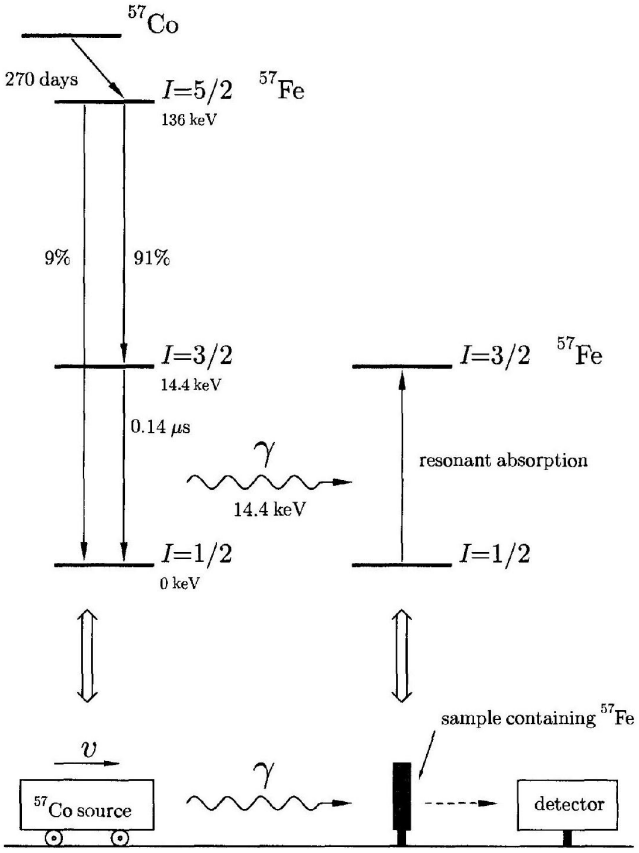
3.2.3 Mössbauer spectroscopy

Another form of spectroscopy is based upon the Mössbauer effect which was discovered in 1957. The principle of the experiment is illustrated in Fig. 3.19. A source containing ^{57}Co nuclei provides a ready supply of excited ^{57}Fe nuclei; these decay to the ground state via a gamma ray cascade which includes a

[9] The more completely the orbital moment is quenched towards zero, the smaller the coupling of the electronic moment to the lattice vibrations and this leads to a longer T_1 and hence a narrower ESR line.

Fig. 3.19 Schematic illustration of the principles of the Mössbauer technique. ${

ˆ{57} \mathrm{Co}$ decays slowly into an excited state of the ^{57}Fe nucleus. The majority (91%) of the subsequent decay is rapidly to the $I = \frac{3}{2}$ state which then decays to the $I = \frac{1}{2}$ ground state. This decay releases a 14.4 keV photon. The experiment is performed by moving the ^{57}Co source at a speed v relative to the sample which must contain some ^{57}Fe nuclei. The detector measures the transmission of γ-rays through the sample which can be used to deduce the absorption.}



14.4 keV gamma ray (corresponding to a frequency $v = 3.5 \times 10^{18}$ Hz). This gamma ray can excite a transition in the sample being studied if it is absorbed resonantly. To do this, its energy must match the energy gap in the sample. By moving the source at speed v one can very slightly adjust the frequency of the gamma ray because of the Doppler effect. Because of the high frequency of the photon, the Doppler shifts can be quite significant: a velocity $v = 1$ mm s^{-1} leads to a shift of $vv/c \approx 12$MHz. Thus one can probe any splittings in the ground state in the source or absorber nucleus which might result from magnetic or other interactions.

The technique would be useless if the γ ray photon emitted by the excited ^{57}Fe nucleus did

not have a well defined frequency. This is why the relatively
slow decay of the ^{57}Fe nucleus is vital: the rather plodding 0.2μ s half-life corresponds to an uncertainty of only about 2 MHz in the frequency. A second vital feature is that the ^{57}Fe atoms are in the solid state. In order to conserve momentum, a free Fe atom would be subject to a recoil velocity of the order of $hv/m_{\mathrm{Fe}}c \sim 80$ m s^{-1} which would ruin the experiment (we have seen that even a relative velocity of 1 mm s^{-1} can produce a measurable effect). However a ^{57}Fe atom which is held rigidly in a solid transmits that momentum to the entire crystal. If the recoil energy is lower than a certain quantity, the probability of phonon emission becomes vanishingly small and the gamma ray can be emitted without any loss of recoil energy. This recoil-free emission and resonant absorption of γ-rays is the essence of the Mössbauer effect.

The conditions described above therefore imply that the effect will be optimized for low-energy γ-rays associated with nuclei strongly bound in a crystal lattice at low temperatures. ^{57}Fe is not the only isotope which is ideal for this purpose (other isotopes which can be used include ^{119}Sn, ^{127}I, ^{151}Eu, and ^{197}Au) but it is the most commonly used. The raw result which one obtains from a Mössbauer measurement is a plot of γ-ray counts (or relative absorption) against the velocity of the source with respect to the absorber.

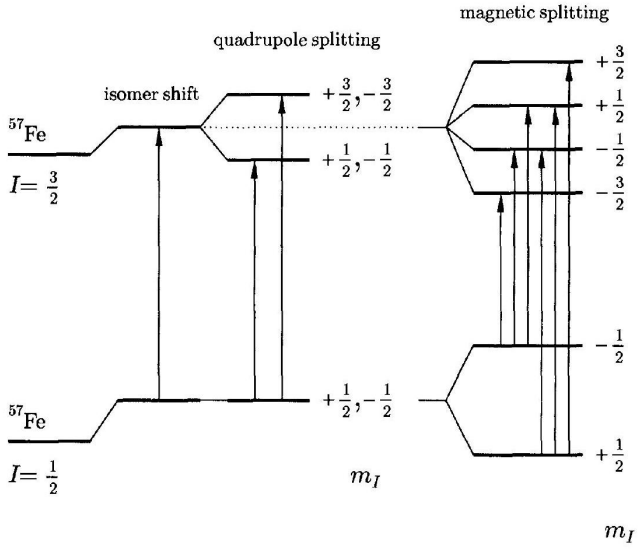
Why should this effect tell us anything about the sample? For a start, the resonant absorption may not occur exactly where you would expect it to (i.e. when the source is stationary), but may be slightly shifted (i.e. to when the source is moving at a particular velocity). This isomer shift is due to the slight change in the Coulomb interaction between the nuclear and electronic charge distributions over the nuclear volume which is associated with the slight increase of size of the ^{57}Fe nucleus in the $I = \frac{3}{2}$ state.

Furthermore, one may not necessarily just observe one resonant absorption line as a function of source velocity, but perhaps a number of lines. This can be due to quadrupole splitting or magnetic splitting. The first effect is due to the electric quadrupole moment of the excited ^{57}Fe nucleus (although the ground state of ^{57}Fe has $I = \frac{1}{2}$ and thus has no electric quadrupole moment, the excited state of interest has $I = \frac{3}{2}$ and nuclei with $I > \frac{1}{2}$ can have a non-zero quadrupole moment). If the nucleus is subjected to an electric field gradient, as may be found in certain crystal environments, the interaction between the nuclear quadrupole moment and the electric field gradient splits the excited $I = \frac{3}{2}$ state into a doublet, so that two lines are produced in the Mössbauer spectrum. The second effect is caused by the interaction between the nucleus and the local magnetic field. This can split the $I = \frac{1}{2}$ ground state into a doublet, and the excited $I = \frac{3}{2}$ into a quadruplet, leading to six possible lines in the Mössbauer spectrum (the selection rule is $\Delta m_1 = 0, \pm 1$). This effect can be used to detect magnetic exchange interactions and local magnetic fields. All these shifts and splittings are illustrated schematically in Fig. 3.20. Note that the figure is not (and could not be) drawn to scale. The observed splittings are typically in the range $10^7 - 10^8$ Hz and are therefore 11 or 12 orders of magnitude smaller than the energy gap between the ground ($I = \frac{1}{2}$) and excited state $\left(I = \frac{3}{2}\right)$ of the ^{57}Fe nucleus. Putting this another way, one is measuring energy splittings of less than 1μeV with a 14.4 keV photon! Were it not for the Mössbauer effect, the remarkable recoil-free resonance absorption and emission of these γ rays, such experiments would be completely impossible.

Rudolf L. Mössbauer (1929-)

Fig. 3.20 The effects of chemical shift, quadrupole splitting and magnetic splitting on the nuclear energy levels of ${

^{57} \mathrm{Fe}$. The arrows show the Mössbauer absorption transitions. The difference in the size of the transitions is greatly exaggerated; in reality they typically differ from each other by less than 1 part in 10^{11}.}



3.2.4 Muon-spin rotation

The muon is a spin- $\frac{1}{2}$ particle (charge $\pm e$, mass $250 m_e$) which has a lifetime of 2.2μ s. The muon is found in nature as the dominant constituent of cosmic rays arriving at sea-level. Muons can be used to study the magnetic properties of samples, but cosmic rays do not provide a sufficiently intense source. For this purpose it is therefore necessary to use the more intense beams of muons available from synchrotrons and cyclotrons.

A technique called muon-spin rotation (often abbreviated to μSR) will now be described in some detail. It is important to realize that in sharp contrast to the neutron and X-ray techniques that will be discussed later in the book, scattering is not involved; muons are implanted into a sample of interest and reside there for the rest of their short lives, never to emerge again. It is the positrons into which they decay that are released from the sample and yield information about the muons from which they came.

The mass of the muon is intermediate between that of the electron and the proton, and thus so are its magnetic moment and gyromagnetic ratio. The muon comes in either charge state, although it is the positive muon, μ^+, which is of particular use in experiments in magnetism. As a small, positively charged particle, it is attracted by areas of large electron density and stops in interstitial sites in inorganic materials or bonds directly on to organic molecules. By contrast the negative muon, μ^-, implants close to an atomic nucleus and is generally much less sensitive to magnetic properties.

Muons can be prepared by colliding a high energy proton beam with a suitable target which produces pions. The pions (π^+) decay very quickly (26 ns) into muons

$$\pi^+ \rightarrow \mu^+ + \upsilon_\mu,$$

so that if one selects the muons arising from pions which have stopped in the target, the muon beam emerges completely spin-polarized. This is because the neutrino (ν_μ) which is also produced has its spin antiparallel to its momentum; the pion has no spin, so the muon must have its spin antiparallel to its momentum and hence emerges spin-polarized. These muons

can then be implanted into a sample but their energy is large, at least 4 MeV . Following implantation they lose energy very quickly (in $0.1 - 1$ ns) to a few keV by ionization of atoms and scattering with electrons. A muon then begins to undergo a series of successive electron capture and loss reactions which reduce the energy to a few hundred eV in about a picosecond. If muonium (see Fig. 3.21), a hydrogen-like state consisting of $\mu^+$and e^-, is ultimately formed then electron capture ultimately wins and the last few eV are lost by inelastic collisions between the muonium atom and the host atoms. All of these effects are very fast so that the muon (or muonium) is thermalized very rapidly. Moreover the effects are all Coulombic in origin and do not interact with the muon-spin so that the muon is thermalized in matter without appreciable depolarization. This is a crucial feature for muon-spin rotation experiments. One may be concerned that the muon may only measure a region of sample which has been subjected to radiation damage by the energetic incoming muon. This does not appear to be a problem since there is a threshold energy for vacancy production, which means that only the initial part of the muon path suffers much damage. Beyond this point of damage the muon still has sufficient energy to propagate through the sample a further distance, thought to be about 1μ m, leaving it well away from any induced vacancies and interstitials.

In a μSR experiment the muons are stopped in the specimen of interest and decay after a time t with probability proportional to e^{-t/τ_μ} where $\tau_\mu = 2.2\mu$ s is the lifetime of the muon. The muon decay is a three-body process

$$\mu^+ \rightarrow e^+ + v_e + \bar{v}_\mu.$$

The decay involves the weak interaction and thus has the unusual feature of not conserving parity. This phenomenon (which also lies behind the negative helicity of the muon neutrino) leads to a propensity for the emitted positron (e^+) to emerge predominantly along the direction that the muon-spin was aligned along when the muon decayed.

The angular distribution of emitted positrons is shown in Fig. 3.22 for the case of the most energetic emitted positrons. In fact positrons over a range of energies are emitted so that the net effect is something not quite as pronounced. In an experiment a magnetic field can be applied perpendicular to the initial muon-spin direction. This can cause the muon-spin to precess. By repeating the experiment for many muons, one can therefore follow the polarization of an ensemble of precessing muons with arbitrary accuracy, provided one is willing to take data for long enough. A schematic diagram of the experiment is shown in Fig. 3.23(a). Consider a muon, with its polarization aligned antiparallel to its momentum, which is implanted in a sample. (It is antiparallel because of

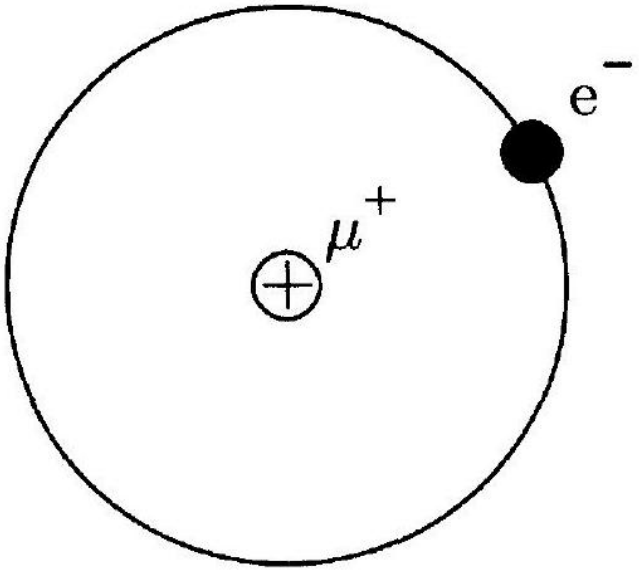


Fig. 3.21 Muonium, a hydrogen-like state consisting of $\mu^{+} $and e^{-}.}

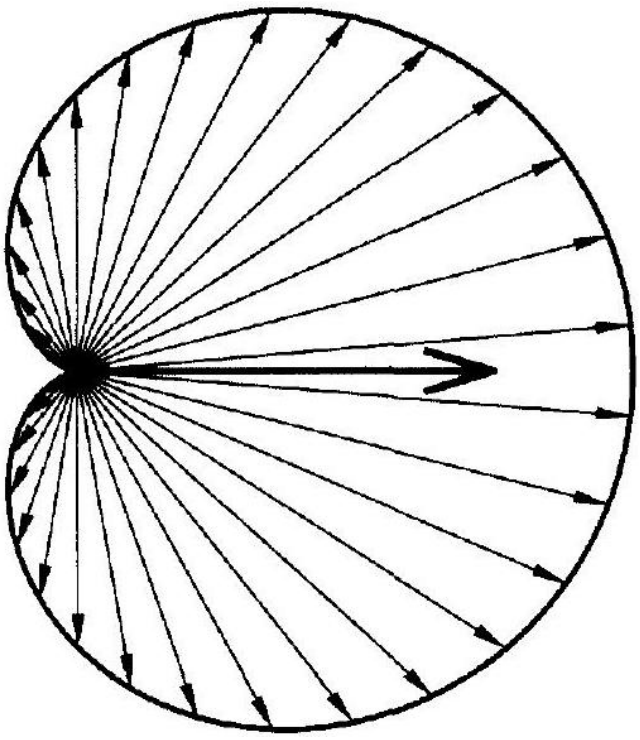


Fig. 3.22 The angular distribution of emitted positrons with respect to the initial muonspin direction. The expected distribution for the most energetically emitted positrons is shown.

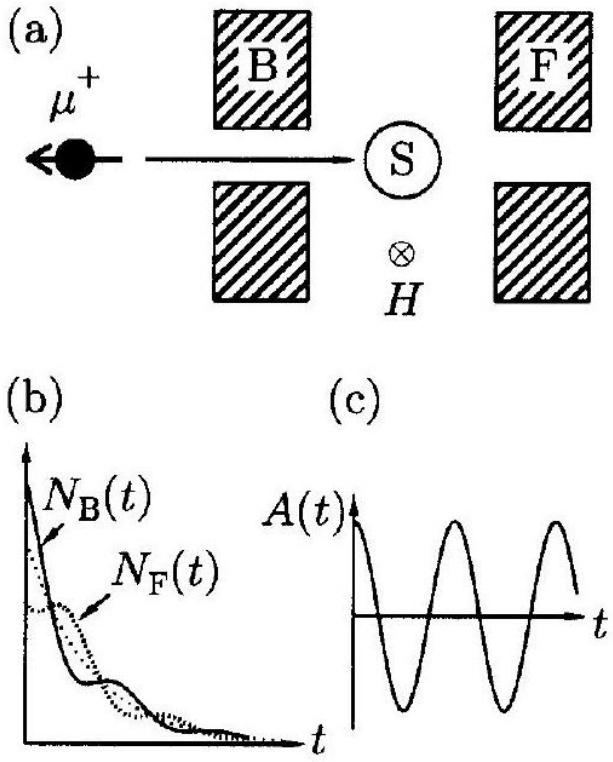


Fig. 3.23 (a) Schematic illustration of a μ SR experiment. A spin-polarized beam of muons is brought to rest in a sample S. Following decay, positrons are detected in either a forward detector F or a backward detector B. If a transverse magnetic field H is applied to the sample as shown then the muons will precess. (b) The number of positrons detected in the forward (broken-line) and backward (solidline) detectors. The dotted line shows the average of the two signals. (c) The asymmetry function.

the way that it was formed, see above, so the muon enters the sample with its spin pointing along the direction from which it came.) If the muon is unlucky enough to decay immediately, then it will not have time to precess and a positron will be emitted preferentially into the backward detector. If it lives a little longer it will have time to precess so that if it lives for half a revolution the resultant positron will be preferentially emitted into the forward detector. Thus the positron beam from an ensemble of precessing muons can be likened to the beam of light from a lighthouse.
The time evolution of the numbers of positrons detected in the forward and backward detector are described by the functions $N_F(t)$ and $N_B(t)$ respectively and these are shown in Fig. 3.23(b). Because the muon decay is a radioactive process these two terms sum to an exponential decay. Thus the time evolution of the muon polarization can be obtained by examining the normalized difference of these two functions via the asymmetry function $A(t)$, given by

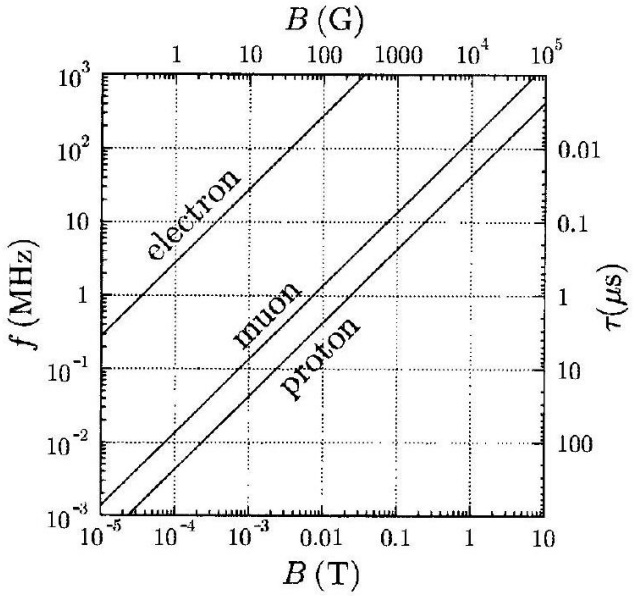
$$A(t) = \frac{N_B(t) - N_F(t)}{N_B(t) + N_F(t)}, \tag{3.40}$$

and is shown in Fig. 3.23(c). The muon is produced with 100% spin polarization so that, in contrast with NMR, one does not need to use tricks with pulses to observe a free induction decay. The muon will undergo Larmor precession in a magnetic field, internal or external. This precession can be followed by measuring the asymmetry. The frequency of precession is directly related to the magnetic field via $\omega = \gamma_\mu B$ where $\gamma_\mu = ge/2m_\mu = 2\pi \times 135.5\text{MHzT}^{-1}$ is the gyromagnetic ratio for the muon and m_μ is its mass (here $g \approx 2$). Implanted muons

in magnetically ordered materials therefore precess in the internal magnetic field and directly yield oscillating signals whose frequency is proportional to the internal magnetic field. In this respect the muon behaves as a microscopic magnetometer. The Larmor precession frequencies for the proton (for NMR), electron (for ESR) and muon (for μSR) are shown in Fig. 3.24.

The large magnetic moment of the muon makes it very sensitive to extremely small magnetic fields (down to $\sim 10^{-5}$ T) and thus is useful in studying

Fig. 3.24 The Larmor precession frequency f in MHz (and the corresponding period $\tau = 1/f$) for the electron, muon and proton as a function of applied magnetic field B.



small moment magnetism. It is also valuable in studying materials in which the magnetic order is random or of very short range. Since muons stop uniformly throughout a sample, each signal appears in the experimental spectrum with a strength proportional to its volume fraction, and thus the technique is helpful in cases where samples may be multiphase or incompletely ordered. Because no spatial information is obtained (in contrast to diffraction techniques considered later) single-crystal samples are not essential (though they can be useful in certain cases) and experiments can often provide information on the magnetic order of certain materials where conventional magnetic neutron diffraction cannot be simply performed. To extract quantitative information from μSR experiments it is necessary to know the muon-site and this can in some cases hinder the search for a straightforward interpretation of the data. (In NMR of course the nucleus localizes the protons rather well so you exactly know the location of your probe.) Usually there is a small set of possible interstitial sites which the muon can occupy and in favourable circumstances only one will be consistent with the observed data. The technique has been widely applied to magnetic materials.

9.3.3 Further reading

- A good introduction to NMR may be found in P. J. Hore, Nuclear magnetic resonance OUP 1995. Also extremely useful is B. Cowan, Nuclear magnetic resonance and relaxation CUP 1997. The classic text on NMR is A. Abragam, Principles of nuclear magnetism OUP 1961.

- A. Abragam and B. Bleaney, Electron Paramagnetic Resonance of Transition Ions, Dover

1986, provides extensive information about crystal fields and ESR experiments in paramagnetic salts.

- The crystal field can be considered using the so-called Stevens operators, see K. W. H. Stevens, Proc. Phys. Soc. A 65, 209 (1952) and M. T. Hutchings, Solid State Physics 16, 227 (1966).
- For further information on the Mössbauer effect see Mössbauer spectroscopy, edited by D. P. E. Dickson and F. J. Berry, CUP 1986.
- For further information on μSR see S. J. Blundell, Contemp. Phys. 40, 175 (1999).

Exercises

(3.1) ASc^{++}ion has one electron in the 3d shell. It is in an anisotropic crystal and the crystal field can be written as a potential acting on the 3 d electron as $A\hat{l}_z^2$. What are the lowest orbital states of the Sc ion if $A > 0$ and if $A < 0$? The spin-orbit coupling $\lambda\hat{\mathbf{I}} \cdot \hat{\mathbf{s}}$ is much smaller than the crystal field. When this is included, what are the approximate ground states of the ion, for $A < 0$ and $A > 0$? Discuss the effect on these states of applying a small magnetic field along the z axis and perpendicular to the z axis, and sketch the temperature dependence of the susceptibility.
(3.2) Equal point positive charges are placed on each of the six corners of an octahedron. Taking the origin of a
set of cartesian coordinates to be at the centre of the octahedron, show that the potential close to the centre is given by

$$V = \frac{q}{4\pi\epsilon_0 a}\left[6 + \frac{35}{4a^4}\left(x^4 + y^4 + z^4 - \frac{3}{5}r^4\right) + O\left(\frac{r^6}{a^6}\right)\right] \tag{3.41}$$

where q is the magnitude of each charge and a is the distance between the origin and each charge.
(3.3) A compound has n transition-metal ions per unit volume, and each ion sits in a crystal field of the form

$$V = D\left(x^4 + y^4 + z^4 - \frac{3}{5}r^4\right) \tag{3.42}$$

which acts as a perturbation on the degenerate d levels. We can choose the unperturbed wave functions in a variety of ways, so let us begin by picking eigenfunctions of $\hat{L}_z$, labelling them by their eigenvalue m_l. Thus

$$\begin{aligned} |\pm 2\rangle &= R(r)\sin^2\theta \mathrm{e}^{\pm 2\mathrm{i}\phi} \\ |\pm 1\rangle &= \mp 2R(r)\sin\theta\cos\theta\mathrm{e}^{\pm\mathrm{i}\phi} \\ |0\rangle &= \sqrt{\frac{2}{3}}R(r)\left(3\cos^2\theta - 1\right) \end{aligned} \tag{3.43}$$

where $R(r)$ is the radial part of the wave function which includes a normalization constant. A general state of the system can be written as $|\psi\rangle = \sum_{m_j=-2}^{2} a_j|j\rangle$ and then specified as a vector:

$$|\psi\rangle = \begin{pmatrix} a_2 \\ a_1 \\ a_0 \\ a_1 \\ a_{-2} \end{pmatrix}. \tag{3.44}$$

In this basis show that

$$V = \begin{pmatrix} A & 0 & 0 & 0 & 5A \\ 0 & -4A & 0 & 0 & 0 \\ 0 & 0 & 6A & 0 & 0 \\ 0 & 0 & 0 & -4A & 0 \\ 5A & 0 & 0 & 0 & A \end{pmatrix} \tag{3.45}$$

where A is a constant and show that the eigenvalues and eigenfunctions are:

$$\begin{array}{lccccc} \text{eigenvalues :} & 6A & 6A & -4A & -4A & -4A \\ \text{eigenfunctions :} & |0\rangle & \frac{|2\rangle+|-2\rangle}{\sqrt{2}} & |1\rangle & |-1\rangle & \frac{|2\rangle-|-2\rangle}{\sqrt{2}} \end{array} \tag{3.46}$$

A magnetic field B induces a splitting $\mu_{\mathrm{B}} B m_l$ so that V becomes

$$V = \begin{pmatrix} A+2\mu_{\mathrm{B}}B & 0 & 0 & 0 & 5A \\ 0 & -4A+\mu_{\mathrm{B}}B & 0 & 0 & 0 \\ 0 & 0 & 6A & 0 & 0 \\ 0 & 0 & 0 & -4A-\mu_{\mathrm{B}}B & 0 \\ 5A & 0 & 0 & 0 & A-2\mu_{\mathrm{B}}B \end{pmatrix} \tag{3.47}$$

Calculate the eigenvalues and eigenvectors in this case. A plot of the eigenvalues is shown in Fig. 3.25.
Also show that at low temperatures, $k_{\mathrm{B}}T \ll A$, the magnetic susceptibility in vanishing magnetic field is equal to

$$\chi = \begin{cases} \frac{2\mu_0\mu_{\mathrm{B}}^2 n}{3k_{\mathrm{B}}T} & \text{for} \quad A > 0 \\ \frac{2\mu_0\mu_{\mathrm{B}}^2 n}{5A} & \text{for} \quad A < 0 \end{cases} \tag{3.48}$$

the first case being temperature independent and small, while the second case is equivalent to that expected for a free ion with $l = 1$.

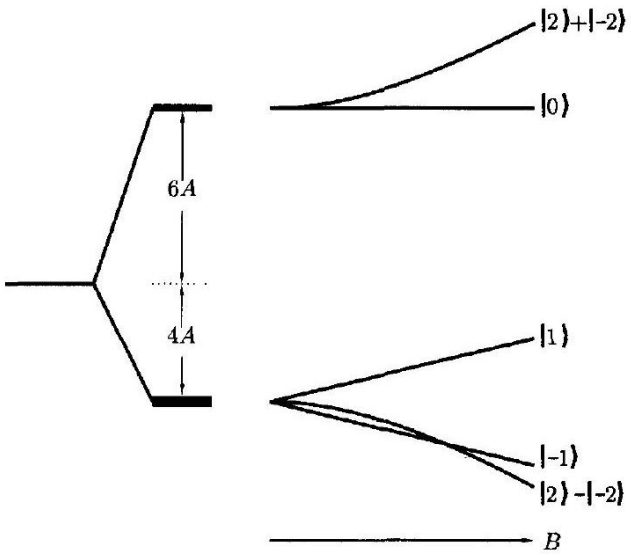


Fig. 3.25 A crystal field splits an ion with $l = 2$ into a triplet and an excited singlet. A magnetic field B splits these states further. The eigenstates as listed refer to vanishing magnetic field (as the field increases, the higher state acquires progressively more $|2\rangle$ character and the lower state acquires progressively more $|-2\rangle$ character).

(3.4) An ion, whose nucleus has zero nuclear spin, has a ground state comprising two degenerate levels corresponding to an effective spin $S = 1/2$. The application of a magnetic field of flux density B produces a separation of the levels which is linear in B. A single electron paramagnetic resonance line is observed for the ion at a frequency of 30 GHz and a magnetic flux density of 0.6 T . An isotope with non-zero nuclear spin I gives rise to a hyperfine structure (described in the spin-Hamiltonian by a term $A\mathbf{I} \cdot \mathbf{S}$) comprising four approximately equally spaced resonance lines, with separation 10^{-2} T, symmetrically disposed about the line due to the isotope with zero nuclear spin. What information does this give about (a) the nuclear spin and (b) the nuclear magnetic moment of the isotope? Calculate the value of the parameter A. Why is it sometimes necessary to perform these measurements at low temperature?
(3.5) Figure 3.26 shows the magnetic moment per ion plotted versus B/T for three paramagnetic salts. Show that eqn 2.9 has the same form as the data at both low and high B/T. The 3d shells of free Cr^{3+} and Fe^{3+} ions contain three and five electrons respectively, and the 4 f shell of Gd^{3+} contains seven. Using Hund's rules, calculate the values of S, L, J and g_J for these ions, and show that the values you obtain account well for the data in the figure for Gd^{3+} and Fe^{3+}, but not for Cr^{3+}. Infer the actual g-factor and ground-state quantum numbers for Cr^{3+} from the experimental data, and describe briefly
the reason for the discrepancy between the data and your initial calculation.

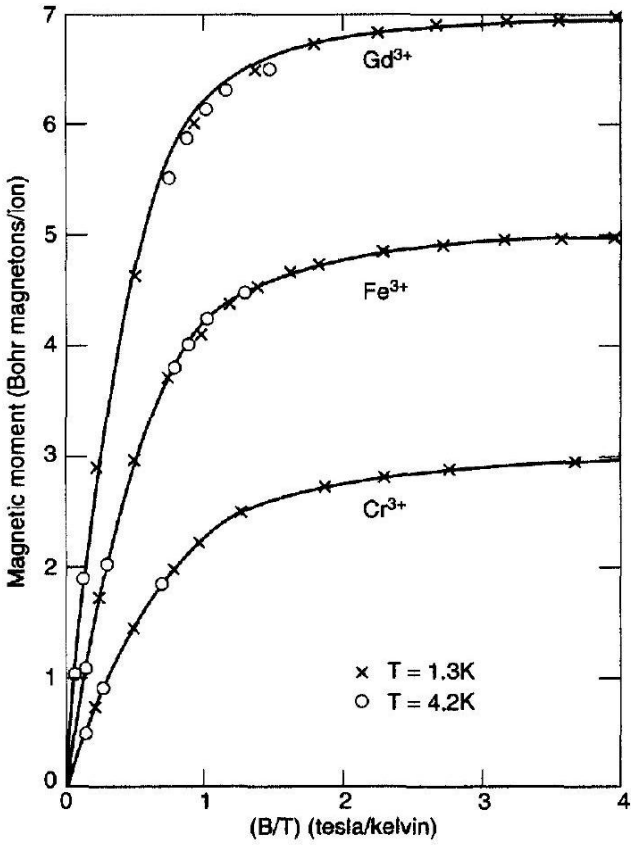


Fig. 3.26 The magnetic moment per ion for three paramagnetic salts: $\mathrm{KCr \left(\mathrm{SO}\right)_{4}\right)_{2} .12 \mathrm{H}_{2} \mathrm{O}, \mathrm{NH}_{4} \mathrm{Fe}\left(\mathrm{SO}_{4}\right)_{2} .12 \mathrm{H}_{2} \mathrm{O}$ and $\mathrm{Gd}_{2}\left(\mathrm{SO}_{4}\right)_{3} .8 \mathrm{H}_{2} \mathrm{O}$ $\left(\mathrm{Gd}^{3+}\right)$. After W. Henry, Phys. Rev. 88, 559 (1952).}

(3.6) An NMR spectrometer operates at a frequency of 60 MHz . At what applied magnetic field would you expect to observe the resonance of $^{1}\mathrm{H}, ^{2}\mathrm{H}, ^{13}\mathrm{C}$ and ^{19} F nuclei? (Use the data in Table 2.3.)

(3.7) An ESR spectrometer operates at a frequency of 9 GHz (known as X-band). What magnetic field would be required to observe a signal from the unpaired electron in DPPH (an organic molecule used to calibrate ESR spectrometers which gives a sharp signal at $g = 2$).

(3.8) A spin Hamiltonian for a system is given by

$$\hat{\mathcal{H}} = g_{\parallel}\mu_{\mathrm{B}} B_z S_z + g_{\perp}\mu_{\mathrm{B}} \left(B_x S_x + B_y S_y\right) + D S_z^2. \tag{3.49}$$

Show that for two electrons in a triplet state ($S = 1$) in this system, the energy eigenstates E can be obtained from

$$\begin{aligned} &E(E-D)^2 - E\left(g_{\parallel}\mu_{\mathrm{B}} B\cos\theta\right)^2 \\ &\quad - (E-D)\left(g_{\perp}\mu_{\mathrm{B}} B\sin\theta\right)^2 = 0 \end{aligned} \tag{3.50}$$

Find and sketch the solutions for $\theta = 0$ and $\theta = \pi/2$. Illustrate the possible photon transitions for $\hbar\omega > D$ and hence show how the parameters $g_{\parallel}, g_{\perp}$ and D could be obtained by measuring ESR for $\theta = 0$ and $\theta = \pi/2$.

(3.9) In the Mössbauer effect a 14.4 keV photon is emitted from an ^{57}Fe atom. Calculate the velocity of recoil for a free ^{57}Fe atom and a ^{57}Fe atom that is rigidly held in the lattice of a 10 mg crystal. What is the Doppler shift of the 14.4 keV photon in each case? A sharp line in a Mössbauer experiment was obtained when the relative velocity between source and sample was 2 mm s^{-1}. Calculate the corresponding frequency shift in Hz and cm^{-1} and also the energy shift in eV .
(3.10) The average muon lifetime is 2.2μ s; assuming that a measurable fraction of muons live for 20μ s, what lower limit does this put on the internal field that can be measured. In a typical experimental run, 10^7 muons are measured. Of these, how many live for 20μ s or longer? Pulses of muons are produced by a synchrotron with a pulse-width of 50 ns . What upper limit does this put on the internal magnetic field in a specimen that can be measured? (Hint: consider the precession signal resulting from the muons at the back and the front of the pulse.) At a cyclotron, a continuous beam of muons can be produced which avoids this problem. An ordered magnetic oxide has an internal field of 0.4 T at the interstitial site which is occupied by an implanted muon. What precession frequency do you expect to be measured in this case?
(3.11) The saturation moment of an Fe^{3+} ion (^{6} S$_{5/2}$) in a crystal is expected to be 5μ_B (see Fig. 3.26). But the effective moment deduced from the susceptibility measurements is expected to be 5.92μ_B (see Table 3.1). Why the difference?

9.4 Interactions

[1] See Fig. 3.8 for a handy conversion from energy to temperature.

In this chapter we consider the different types of magnetic interaction which can be important in allowing the magnetic moments in a solid to communicate with each other and potentially to produce long range order.

9.4.1 4.1 Magnetic dipolar interaction

The first interaction which might be expected to play a rôle is the magnetic dipolar interaction. Two magnetic dipoles μ_1 and μ_2 separated by $\mathbf{r}$ have an energy equal to

$$E = \frac{\mu_0}{4\pi r^3}\left[\mu_1 \cdot \mu_2 - \frac{3}{r^2}\left(\mu_1 \cdot \mathbf{r}\right)\left(\mu_2 \cdot \mathbf{r}\right)\right] \tag{4.1}$$

which therefore depends on their separation and their degree of mutual alignment. We can easily estimate the order of magnitude of this effect for two moments each of $\mu \approx 1\mu_B$ separated by $r \approx 1$Å to be approximately $\mu^2/4\pi r^3 \sim 10^{-23}$ J which is equivalent to about 1 K in temperature. [1] Since many materials order at much higher temperatures (some around 1000 K), the magnetic dipolar interaction must be too weak to account for the ordering of most magnetic materials. Nevertheless, it can be important in the properties of those materials which order at milliKelvin temperatures.

9.4.2 4.2 Exchange interaction

Exchange interactions lie at the heart of the phenomenon of long range magnetic order. The exchange effect is subtle and not a little mysterious, since it seems surprising that one has to go to the bother of thinking about exchange operators and identical particles when all one is dealing with is a bar magnet and a pile of iron filings. But this, as so often with the subject of magnetism, is a demonstration of how quantum mechanics is at the root of many everyday phenomena. Exchange interactions are nothing more than electrostatic interactions, arising

because charges of the same sign cost energy when they are close together and save energy when they are apart.

4.2.1 Origin of exchange

Consider a simple model with just two electrons which have spatial coordinates $\mathbf{r}_1$ and $\mathbf{r}_2$ respectively. The wave function for the joint state can be written as a product of single electron states, so that if the first electron is in state $\psi_a(\mathbf{r}_1)$
and the second electron is in state $\psi_b(\mathbf{r}_2)$, then the joint wave function is $\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)$. However this product state does not obey exchange symmetry, since if we exchange the two electrons we get $\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)$ which is not a multiple of what we started with. Therefore, the only states which we are allowed to make are symmetrized or antisymmetrized product states which behave properly under the operation of particle exchange. This was discussed in Section 1.3.4.

For electrons the overall wave function must be antisymmetric so the spin part of the wave function must either be an antisymmetric singlet state χs ($S = 0$) in the case of a symmetric spatial state or a symmetric triplet state $\chi_{\mathrm{T}}(S = 1)$ in the case of an antisymmetric spatial state. Therefore we can write the wave function for the singlet case Ψ_S and the triplet case Ψ_T as

$$\begin{aligned}\Psi_{\mathrm{S}} &= \frac{1}{\sqrt{2}}\left[\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) + \psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)\right]\chi_{\mathrm{S}} \\ \Psi_{\mathrm{T}} &= \frac{1}{\sqrt{2}}\left[\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) - \psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)\right]\chi_{\mathrm{T}}\end{aligned} \tag{4.2}$$

where both the spatial and spin parts of the wave function are included. The energies of the two possible states are

$$\begin{aligned}E_{\mathrm{S}} &= \int \Psi_{\mathrm{S}}^{*}\hat{\mathcal{H}}\Psi_{\mathrm{S}}\mathrm{d}\mathrm{r}_1\mathrm{d}\mathrm{r}_2 \\ E_{\mathrm{T}} &= \int \Psi_{\mathrm{T}}^{*}\hat{\mathcal{H}}\Psi_{\mathrm{T}}\mathrm{d}\mathrm{r}_1\mathrm{d}\mathrm{r}_2\end{aligned}$$

with the assumption that the spin parts of the wave function χ_S and χ_T are normalized. The difference between the two energies is

$$E_{\mathrm{S}} - E_{\mathrm{T}} = 2\int \psi_a^{*}(\mathbf{r}_1)\psi_b^{*}(\mathbf{r}_2)\hat{\mathcal{H}}\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)\,\mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2 \tag{4.3}$$

Equation 1.70 shows how the difference between singlet and triplet states can be parametrized using $\mathbf{S}_1 \cdot \mathbf{S}_2$. For a singlet state $\mathbf{S}_1 \cdot \mathbf{S}_2 = -\frac{3}{4}$ while for a triplet state $\mathbf{S_1} \cdot \mathbf{S_2} = \frac{1}{4}$. Hence the Hamiltonian can be written in the form of an 'effective Hamiltonian'

$$\hat{\mathcal{H}} = \frac{1}{4}(E_{\mathrm{S}} + 3E_{\mathrm{T}}) - (E_{\mathrm{S}} - E_{\mathrm{T}})\,\mathbf{S}_1 \cdot \mathbf{S}_2. \tag{4.4}$$

This is the sum of a constant term and a term which depends on spin. The constant can be absorbed into other constant energy terms, but the second term is more interesting. The exchange constant (or exchange integral), J is defined by

$$\mathrm{J} = \frac{E_{\mathrm{S}} - E_{\mathrm{T}}}{2} = \int \psi_a^{*}(\mathbf{r}_1)\psi_b^{*}(\mathbf{r}_2)\hat{\mathcal{H}}\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)\,\mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2 \tag{4.5}$$

and hence the spin-dependent term in the effective Hamiltonian can be written

$$\hat{\mathcal{H}}^{\mathrm{spin}} = -2\sqrt{}\mathbf{S_1} \cdot \mathbf{S_2}. \tag{4.6}$$

If $\mathrm{J} > 0, E_\mathrm{S} > E_\mathrm{T}$ and the triplet state $S = 1$ is favoured. If $\mathrm{J} < 0, E_\mathrm{S} < E_\mathrm{T}$ and the singlet state $S = 0$ is favoured. This equation is relatively simple

Werner Heisenberg (1901-1976)

Note that in some books, J is replaced by twice the value used here so that eqns 4.64.8 would become

$$\hat{\mathcal{H}}^{\mathrm{spin}} = -\mathrm{J}\mathbf{S_1} \cdot \mathbf{S_2}, \tag{4.9}$$

$$\hat{\mathcal{H}} = -\frac{1}{2}\sum_{ij} J_{ij}\mathbf{S}_i \cdot \mathbf{S}_j, \tag{4.10}$$

$$\hat{\mathcal{H}} = -\sum_{i>j} J_{ij}\mathbf{S}_i \cdot \mathbf{S}_j. \tag{4.11}$$

to derive for two electrons, but generalizing to a many-body system is far from trivial. Nevertheless, it was recognized in the early days of quantum mechanics that interactions such as that in eqn 4.6 probably apply between all neighbouring atoms. This motivates the Hamiltonian of the Heisenberg model:

$$\hat{\mathcal{H}} = -\sum_{ij} J_{ij}\mathbf{S}_i \cdot \mathbf{S}_j, \tag{4.7}$$

where J_{ij} is the exchange constant between the i^{th} and j^{th} spins. The factor of 2 is omitted because the summation includes each pair of spins twice. Another way of writing eqn 4.7 is

$$\hat{\mathcal{H}} = -2\sum_{i>j} J_{ij}\mathbf{S}_i \cdot \mathbf{S}_j, \tag{4.8}$$

where the $i > j$ avoids the 'double-counting' and hence the factor of two returns. Often it is possible to take J_{ij} to be equal to a constant J for nearest neighbour spins and to be 0 otherwise.

The calculation of the exchange integral can be complicated in general, but we here mention some general features. First, if the two electrons are on the same atom, the exchange integral is usually positive. This stabilizes the triplet state and ensures an antisymmetric spatial state which minimizes the Coulomb repulsion between the two electrons by keeping them apart. This is consistent with Hund's first rule.

When the two electrons are on neighbouring atoms, the situation is very different. Any joint state will be a combination of a state centred on one atom and a state centred on the other. It is worth remembering that the energy of a particle in a one-dimensional box of length L is proportional to L^{-2}; this is a kinetic energy and hence demonstrates that there is a large kinetic energy associated with being squeezed into a small box. The electrons therefore can save kinetic energy by forming bonds because this allows them to wander around both atoms rather than just one (i.e. wander in a 'bigger box'). The correct states to consider are now not atomic orbitals but molecular orbitals (see Fig. 4.1). These can be bonding (spatially symmetric) or 'antibonding' (spatially antisymmetric), with the antibonding orbitals more energetically costly. This is because the antibonding orbital has a greater curvature and hence a larger kinetic energy. This favours singlet (antisymmetric) states and the exchange integral is therefore likely to be negative.

4.2.2 Direct exchange

If the electrons on neighbouring magnetic atoms interact via an exchange interaction, this is known as direct exchange. This is because the exchange interaction proceeds directly without

the need for an intermediary. Though this seems the most obvious route for the exchange interaction to take, the reality in physical situations is rarely that simple.

Very often direct exchange cannot be an important mechanism in controlling the magnetic properties because there is insufficient direct overlap between neighbouring magnetic orbitals. For example, in rare earths the 4 f electrons are strongly localized and lie very close to the nucleus, with little probability density extending significantly further than about a tenth of the interatomic

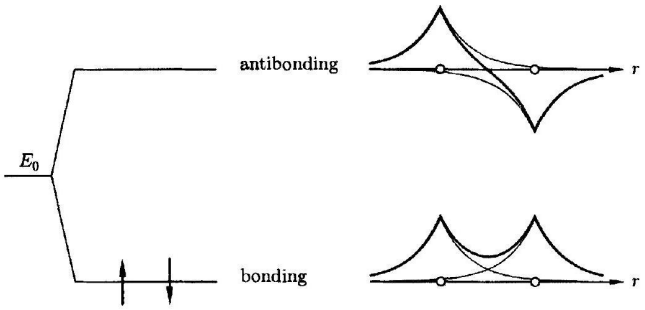


Fig. 4.1 Molecular orbitals for a diatomic molecule. The bonding orbital, which corresponds to the sum of the two atomic orbitals (symmetric under exchange, as far as the spatial part of the wave function is concerned), is of lower energy than the antibonding orbital, which corresponds to the difference of the two atomic orbitals (antisymmetric under exchange). This therefore favours a singlet ground state in which two electrons fill the bonding state and the antibonding state is empty. This diagram is appropriate for the hydrogen molecule $\mathrm{H
_{2}$ which has a lower energy than that of two isolated H atoms (E_0). Note that the diatomic form of helium, He_2, does not form because the four electrons from two He atoms would fill both the bonding and antibonding orbitals, corresponding to no net energy saving in comparison with two isolated He atoms.}

spacing. This means that the direct exchange interaction is unlikely to be very effective in rare earths. Even in transition metals, such as Fe, Co and Ni , where the 3d orbitals extend further from the nucleus, it is extremely difficult to justify why direct exchange should lead to the observed magnetic properties. These materials are metals which means that the role of the conduction electrons should not be neglected, and a correct description needs to take account of both the localized and band character of the electrons.

Thus in many magnetic materials it is necessary to consider some kind of indirect exchange interaction.

4.2.3 Indirect exchange in ionic solids: superexchange

A number of ionic solids, including some oxides and fluorides, have magnetic ground states. For example, MnO (see Fig. 4.2) and MnF_2 are both antiferromagnets, though this observation appears at first sight rather surprising because there is no direct overlap between the electrons on Mn^{2+} ions in each system. The exchange interaction is normally very short-ranged so that the longerranged interaction that is operating in this case must be in some sense 'super'.

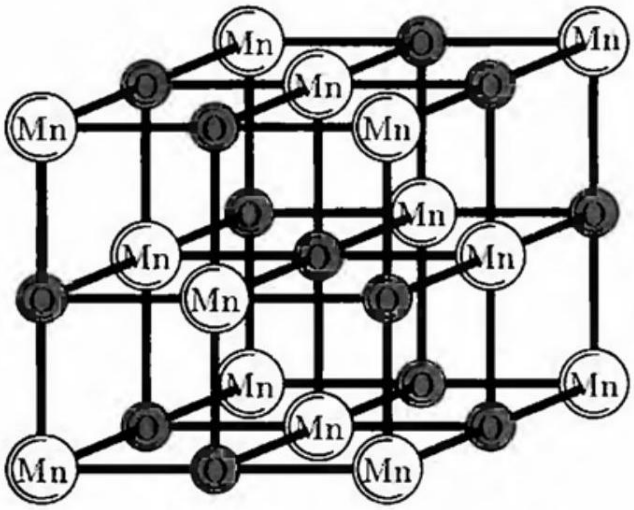
The exchange mechanism which is operative here is in fact known as superexchange. It can be defined as an indirect exchange interaction between non-neighbouring magnetic ions which is mediated by a non-magnetic ion which is placed in between the magnetic ions. It arises because there is a kinetic energy advantage for antiferromagnetism, which can be understood by reference to Fig. 4.3 which shows two transition metal ions separated by an oxygen ion. For simplicity we will assume that the magnetic moment on the transition metal ion is due to a

single unpaired electron (more complicated cases can be dealt with in analogous ways). Hence if this system were perfectly ionic, each metal ion would have a single unpaired electron in a d orbital

Fig. 4.3 Superexchange in a magnetic oxide. The arrows show the spins of the four electrons and how they are distributed over the transition metal (M) and oxygen (O) atoms. *M* is assumed to have a single unpaired electron, making it magnetic. If the moments on the transition metal atoms are coupled antiferromagnetically (a, b, c). the ground state is (a) and this can mix with excited configurations like (b) and (c). The magnetic electrons can thus be delocalized over the M-O-M unit, thus lowering the kinctic energy. If the moments on the metal (M) atoms are coupled ferromagnetically (d.e.f), the ground state (d) cannot mix with excited configurations like (c) and (f) because these configurations are prevented by the exclusion principle. The ferromagnetic configuration therefore costs more energy.

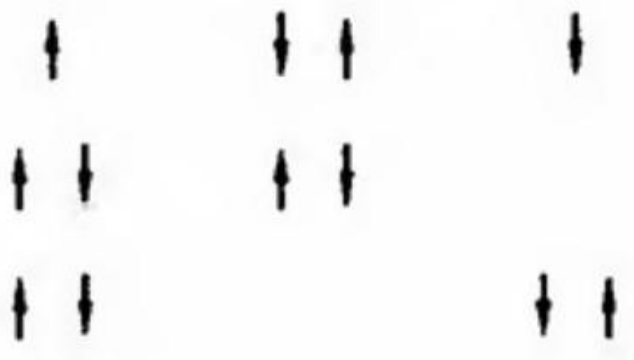
Fig. 4.2 The crystal structure of MnO . Nearest neighbour pairs of $\mathrm{Mn

ˆ{2+}$ (manganese) ions are connected via O^{2-} (oxygen) ions.}



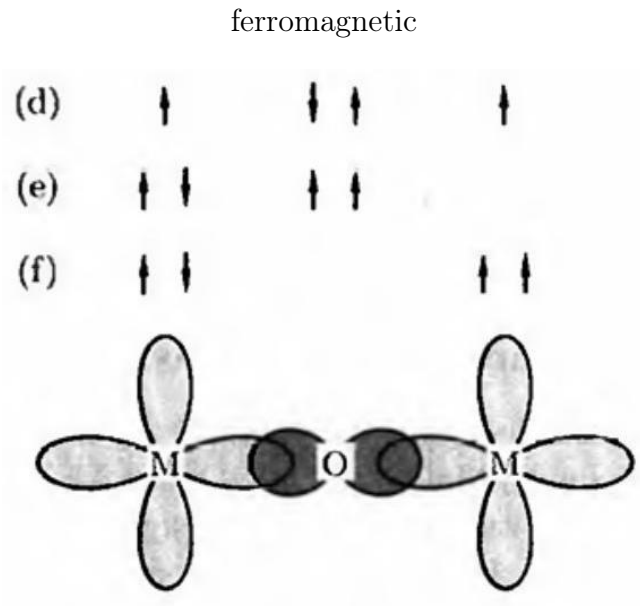
antiferromagnetic antiferromagnetic

(a)



(b)

(c)



and the oxygen would have two p electrons in its outermost occupied states. The figure demonstrates that antiferromagnetic coupling lowers the energy of the system by allowing these electrons to become delocalized over the whole structure, thus lowering the kinetic energy.

Because superexchange involves the oxygen orbitals as well as the metal atom, it is a second-order process and is derived from second-order perturbation theory. A general consequence of second-order perturbation theory is that the energy involved is approximately given by the square of the matrix element of the transition divided by the energy cost of making the excited state. Here the transition matrix element is controlled by a parameter called the hopping integral t, which is proportional to the energy width of the conduction band (i.e. the bandwidth) in a simple tight-binding approach. The energy cost of making an excited state is given by the Coulomb energy U. Thus we have that $\mathrm{J} \sim -t^2/U$. (In fourth-order it is possible to have an interaction of the form $\Delta E \propto -t^4 \left(\mathbf{S}_1 \cdot \mathbf{S}_2\right)^2 / U^3$ which is known as biquadratic exchange.)

The exchange integral consists of two parts. The first is a potential exchange term which represents the electron repulsion and favours ferromagnetic ground states, but is small when the ions are well separated. The second is a kinetic exchange term which dominates here and is the effect discussed above. It depends on the degree of overlap of orbitals and thus superexchange is strongly dependent upon the angle of the $\mathrm{M}-\mathrm{O}-\mathrm{M}$ bond. The figure has been drawn for one type of d orbital only, but the effect of other d orbitals which can overlap with the oxygen orbitals may also need to be added.

In some circumstances, superexchange can actually be ferromagnetic. For example, imagine a situation in which there is a coupling, through an oxygen ion, between an occupied e_g orbital on one magnetic ion and an unoccupied e_g orbital on another magnetic ion. There is an energetic advantage to the e_g electron hopping onto the unoccupied orbital, if when it arrives its spin is aligned with the spin of the $\mathbf{t}_{2g}$ electrons because of the Hund's rule coupling. Thus the superexchange could be ferromagnetic in this case, but this is weaker interaction and less common than the usual antiferromagnetic superexchange.

4.2.4 Indirect exchange in metals

In metals the exchange interaction between magnetic ions can be mediated by the conduction electrons. A localized magnetic moment spin-polarizes the conduction electrons and this polarization in turn couples to a neighbouring localized magnetic moment a distance r away. The exchange interaction is thus indirect because it does not involve direct coupling between

magnetic moments. It is known as the RKKY interaction (or also as itinerant exchange). The name RKKY is used because of the initial letters of the surnames of the discoverers of the effect, Ruderman, Kittel, Kasuya and Yosida. The coupling takes the form of an r-dependent exchange interaction $J_{RKKY}(r)$ given by

$$\mathrm{J_{RKKY}}(r) \propto \frac{\cos\left(2k_{\mathrm{F}}r\right)}{r^3} \tag{4.12}$$

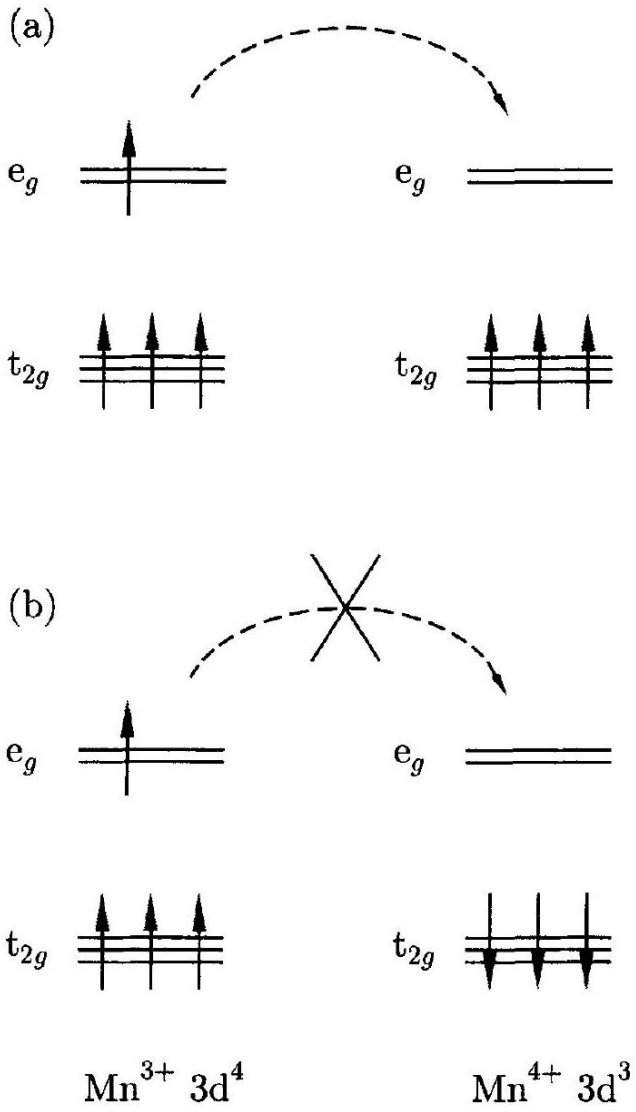
at large r (assuming a spherical Fermi surface of radius k_{F}). The interaction is long range and has an oscillatory dependence on the distance between the magnetic moments. Hence depending on the separation it may be either ferromagnetic or antiferromagnetic. The coupling is oscillatory with wavelength π/k_{F} because of the sharpness of the Fermi surface. The RKKY interaction will be considered in more detail in chapter 7 .

4.2.5 Double exchange

In some oxides, it is possible to have a ferromagnetic exchange interaction which occurs because the magnetic ion can show mixed valency, that is it can exist in more than one oxidation state. Examples of this include compounds containing the Mn ion which can exist in oxidation state 3 or 4 , i.e. as Mn^{3+}

Fig. 4.4 Double exchange mechanism gives ferromagnetic coupling between $\mathrm{Mn

ˆ{3+}$ and Mn^{4+} ions participating in electron transfer. The single-centre exchange interaction favours hopping if (a) neighbouring ions are ferromagnetically aligned and not if (b) neighbouring ions are antiferromagnetically aligned.}



or Mn^{4+}. One such is the material $La_{1-x}Sr_xMnO_3(0 \leq x \leq 1)$ which adopts a perovskite structure. Sr is divalent (it exists as Sr^{2+}) and La is trivalent (it exists as La^{3+}). This implies that a fraction x of the Mn ions are Mn^{4+} and $1-x$ are Mn^{3+}. The end members of the series, with $x = 0$ and $x = 1$, are both antiferromagnetic insulators, as would be expected for an oxide material in which the magnetism is mediated by superexchange through the oxygen. $LaMnO_3$ contains only Mn^{3+} ions and Mn^{3+} is a Jahn-Teller ion. $LaMnO_3$ has A-type antiferromagnetic ordering (see Section 5.2). However when $LaMnO_3$ is doped with Sr up to a level of $x = 0.175$, the Jahn-Teller distortion vanishes and the system becomes ferromagnetic with a Curie temperature around room temperature, below which temperature the material becomes metallic.

The ferromagnetic alignment is due to the double exchange mechanism which can be understood with reference to Fig. 4.4. The e_g electron on a Mn^{3+} ion can hop to a neighbouring site only if there is a vacancy there of the same spin (since hopping proceeds without spin-flip of the hopping electron). If the neighbour is a Mn^{4+} which has no electrons in its e_g shell, this should present no problem. However, there is a strong single-centre (Hund's rule number 1) exchange

interaction between the e_g electron and the three electrons in the
t_{2g} level which wants to keep them all aligned. Thus it is not energetically favourable for an e_g electron to hop to a neighbouring ion in which the t_{2g} spins will be antiparallel to the e_g electron (Fig. 4.4(b)). Ferromagnetic alignment of neighbouring ions is therefore required to maintain the high-spin arrangement on both the donating and receiving ion. Because the ability to hop gives a kinetic energy saving, allowing the hopping process shown in Fig. 4.4(a) reduces the overall energy. Thus the system ferromagnetically aligns to save energy. Moreover, the ferromagnetic alignment then allows the e_g electrons to hop through the crystal and the material becomes metallic. The issue of conductivity in double exchange ferromagnets will be further discussed in Section 8.9.5. Double exchange is, essentially, ferromagnetic superexchange in an extended system. [2]

Double exchange is also found in magnetite (Fe_3O_4) which contains an equal mixture of Fe^{2+} $\left(3\ d^6\right)$ and Fe^{3+} $\left(3\ d^5\right)$ ions on octahedral sites, together with the same number again of Fe^{3+} ions on tetrahedral sites. A double exchange interaction ferromagnetically aligns the Fe^{2+} and Fe^{3+} ions on the octahedral sites. The Fe^{3+} ions on the tetrahedral sites do not participate in this interaction and are coupled to the Fe^{3+} ions on the octahedral sites by an antiferromagnetic superexchange interaction. Thus the two sets of Fe^{3+} ions cancel out, leaving a net moment due to the Fe^{2+} ions alone. The measured magnetic moment per formula unit is very close to the expected $4\mu_B$ due to just the Fe^{2+} ions.

4.2.6 Anisotropic exchange interaction

It is also possible for the spin-orbit interaction to play a rôle in a similar manner to that of the oxygen atom in superexchange. Here the excited state is not connected with oxygen but is produced by the spin-orbit interaction in one of the magnetic ions. There is then an exchange interaction between the excited state of one ion and the ground state of the other ion. This is known as the anisotropic exchange interaction, or also as the Dzyaloshinsky-Moriya interaction. When acting between two spins $\mathbf{S}_1$ and $\mathbf{S}_2$ it leads to a term in the Hamiltonian, $\hat{\mathcal{H}}_{\rm DM}$ equal to

$$\hat{\mathcal{H}}_{\rm DM} = \mathbf{D} \cdot \mathbf{S}_1 \times \mathbf{S}_2. \tag{4.13}$$

The vector $\mathbf{D}$ vanishes when the crystal field has an inversion symmetry with respect to the centre between the two magnetic ions. However, in general D may not vanish and then will lie parallel or perpendicular to the line connecting the two spins, depending on the symmetry. The form of the interaction is such that it tries to force $\mathbf{S}_1$ and $\mathbf{S}_2$ to be at right angles in a plane perpendicular to the vector $\mathbf{D}$ in such an orientation as to ensure that the energy is negative. Its effect is therefore very often to cant (i.e. slightly rotate) the spins by a small angle. It commonly occurs in antiferromagnets and then results in a small ferromagnetic component of the moments which is produced perpendicular to the spin-axis of the antiferromagnet. The effect is known as weak ferromagnetism. It is found in, for example, $\alpha - Fe_2O_3, MnCO_3$ and $CoCO_3$.

[2] Ferromagnetic superexchange is normally applied to two isolated ions. The kinetic energy saved by ferromagnetic alignment corresponds to hopping into an excited state. Double exchange is applied to an extended system so the kinetic energy saved corresponds to a gain in electron bandwidth.

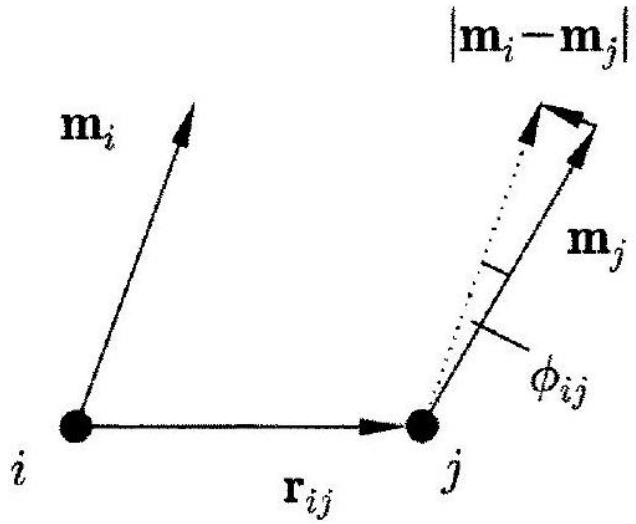


Fig. 4.5 The magnetic moments are represented by the reduced moments $\mathbf{m}_i$ and $\mathbf{m}_j$ at neighbouring sites i and j separated by a vector $\mathbf{r}_{ij}$. The angle between the moments is ϕ_{ij}. The reduced moments $\mathbf{m}_i$ and $\mathbf{m}_j$ are unit vectors by definition.}

4.2.7 Continuum approximation

In this section we return to the Heisenberg model given in eqn 4.7. For what follows later in the book it is useful to find an expression for this interaction in a continuum approximation in which the discrete nature of the lattice is ignored.

Let us first assume that J_{ij} can be taken to be equal to a constant, J , if i and j are nearest neighbours, and to be zero otherwise. Hence we write

$$\hat{\mathcal{H}} = -\sum_{\langle ij \rangle} \mathrm{J}\mathbf{S} \cdot \mathbf{S}_j, \tag{4.14}$$

where the symbol (ij) below the $\sum$ denotes a sum over nearest neighbours only. Let us consider classical spins, and assume that the angle between nearest neighbour spins is ϕ_{ij} and that it is very small, i.e. $\phi_{ij} \ll 1$ for all i and j. What we are essentially doing is assuming that the system shows ferromagnetism (see the next chapter) but the spins are not completely aligned.

With these assumptions, the energy of the system can be written as

$$E = -JS^2 \sum_{<ij>} \cos\phi_{ij} = \text{ constant } + \frac{JS^2}{2} \sum_{<ij>} \phi_{ij}^2 \tag{4.15}$$

where the last equality is obtained using $\cos\phi_{ij} \approx 1 - \phi_{ij}^2/2$ for $\phi_{ij} \ll 1$. We will now ignore the constant term which just refers to the energy of the fully aligned state. We now define the reduced moment by $\mathbf{m} = \mathbf{M}/M_\mathrm{s}$, where $\mathbf{M}$ is the magnetization and M_s is the saturation magnetization. The unit vector $\mathbf{m}$ therefore follows the direction of the spins and m_x, m_y and m_z can be thought of as the direction cosines of the spin at lattice point r_{ij}. Using the notation of Fig. 4.5, we can write

$$|\phi_{ij}| \approx |\mathbf{m}_i - \mathbf{m}_j| \approx |(\mathbf{r}_{ij} \cdot \nabla)\,\mathbf{m}| \tag{4.16}$$

and so the energy can be written

$$E = JS^2 \sum_{\langle ij \rangle} \left[(\mathbf{r}_{ij} \cdot \nabla)\, \mathbf{m} \right]^2 \tag{4.17}$$

In the continuum limit, we ignore the discrete nature of the lattice and therefore write

$$E = A \int_V \left[(\nabla m_x)^2 + (\nabla m_y)^2 + (\nabla m_z)^2 \right] \mathrm{d}^3 r \tag{4.18}$$

where A is given by

$$A = 2JS^2 z/a \tag{4.19}$$

a is the nearest neighbour distance and z is the number of sites in the unit cell ($z = 1$ for simple cubic, $z = 2$ for body-centred cubic (bcc) and $z = 4$ for face-centred cubic (fcc)).

This result can be seen from another perspective. If we assert that exchange arises from a non-uniform magnetization distribution, then if the nonuniformities are relatively smooth we can derive a result based on symmetry. The expression must be invariant with respect to spin rotations and also under change of sign of the magnetization components. Therefore we look for an expression of the lowest even orders of the derivatives of $\mathbf{M}$, consistent with the symmetry of the crystal. Since terms proportional to $\partial M_\alpha / \partial x_\beta$ would change sign if M were reversed, we are left with terms quadratic in the gradient of the magnetization. The most general expression is

$$E = \sum_{\alpha\beta\gamma} C_{\alpha\beta} \frac{\partial M_\gamma}{\partial x_\alpha} \frac{\partial M_\gamma}{\partial x_\beta} \tag{4.20}$$

where $C_{\alpha\beta}$ is a tensor with the symmetry of the crystal. In a cubic crystal, this reduces to

$$E = C \sum_{\alpha\gamma} \frac{\partial M_\gamma}{\partial x_\alpha} \frac{\partial M_\gamma}{\partial x_\alpha} = C \left[(\nabla M_x)^2 + (\nabla M_y)^2 + (\nabla M_z)^2 \right] \tag{4.21}$$

which is equivalent to the result we have in eqn 4.18 above.

9.4.3 Further reading

- Further information on exchange interactions may be found in C. Herring in Magnetism, ed. G. Rado and H. Suhl, vol. 2B, p.1, Academic Press, New York 1966.
- A theoretical account of interactions in magnetic systems may be found in K. Yosida, Theory of magnetism, Springer 1996.

The expression in eqn 4.21 is sometimes written $C|\nabla \mathbf{M}|^2$ but note that you have to be a bit careful when taking the gradient of a vector.

Various exchange interactions in real systems are reviewed in P. A. Cox, Transition metal oxides, OUP 1995.

- A very thorough and helpful reference on exchange and exchange interactions is D. C. Mattis, The theory of magnetism I, Springer 1981.

Exercises

(4.1) Show that two magnetic dipoles μ_1 and μ_2 separated by $\mathbf{r}$ have a dipolar energy equal to

$$E = \frac{\mu_0}{4\pi r^3}\left[\mu_1 \cdot \mu_2 - \frac{3}{r^2}\left(\mu_1 \cdot \mathbf{r}\right)\left(\mu_2 \cdot \mathbf{r}\right)\right] \tag{4.22}$$

(4.2) Calculate the magnitude of the magnetic field 1Å and 10Å from a proton in a direction (a) parallel and (b) perpendicular to the proton spin direction.
(4.3) Estimate the ratio of the exchange and dipolar coupling of two adjacent Fe atoms in metallic Fe . (The exchange constant in Fe can be crudely estimated by setting it equal to $k_{\mathrm{B}}T_{\mathrm{C}}$ where T_{C} is the Curie temperature. For Fe, $T_{\mathrm{C}} = 1043$ K.)
(4.4) Provide a rough estimate of the size of the exchange constant in a magnetic oxide which is coupled by superexchange using the measured value of the electronic bandwidth (determined by inelastic neutron scattering) of
0.05 eV . Take the Coulomb energy to be $\sim$ 1eV. Hence estimate the antiferromagnetic ordering temperature.
(4.5) Consider the case of two interacting spin $-\frac{1}{2}$ electrons. The good quantum numbers are $S = 0$ and 1 so that there is a triplet state and a singlet state which will be separated by an energy gap Δ. We define the sign of Δ so that when $\Delta > 0$ the singlet state ($S = 0$) is the lower state and when $\Delta < 0$ the triplet state is the lower state. These situations are shown in Fig. 4.6(a) and (b). Show that the susceptibility in this model is given by

$$\chi = \frac{2Ng\mu_{\mathrm{B}}^2}{k_{\mathrm{B}}T\left(3 + \mathrm{e}^{\Delta/k_{\mathrm{B}}T}\right)}, \tag{4.23}$$

which is known as the Bleaney-Bowers equation. It is plotted in Fig. 4.6(c).

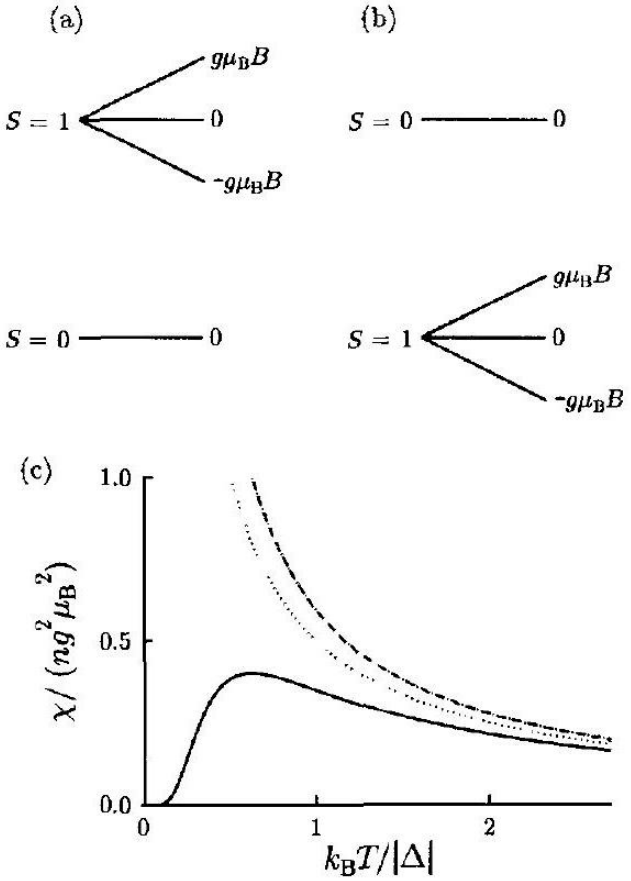


Fig. 4.6 Two spins coupled by Δ give rise to a singlet and triplet state. The energy levels for the case when (a) $\Delta > 0$ and (b) $\Delta < 0$. (c) The susceptibility which is given by the BleaneyBowers equation. It is plotted for $\Delta > 0$ (lower curve), $\Delta = 0$ (middle curve), and $\Delta < 0$ (upper curve).

(4.6) Two nuclei with spins $\mathbf{I}_1$ and $\mathbf{I}_2$, separated by a vector $\mathbf{r}$, are coupled by a dipole-dipole interaction and the Hamiltonian is

$$\hat{\mathcal{H}} = \frac{\mu_0}{4\pi r^3}\left[\hat{\boldsymbol{\mu}}_1 \cdot \hat{\boldsymbol{\mu}}_2 - \frac{3}{r^2}\left(\hat{\boldsymbol{\mu}}_1 \cdot \mathbf{r}\right)\left(\hat{\boldsymbol{\mu}}_2 \cdot \mathbf{r}\right)\right], \tag{4.24}$$

where $\mu_i = g_I \mu_N \mathbf{I}_i, i = 1, 2$. Show that the Hamiltonian can be expressed as

$$\hat{\mathcal{H}} = \frac{\mu_0 g_I^2 \mu_N^2}{4\pi r^3}[A + B + C + D + E + F]. \tag{4.25}$$

where

$$
\begin{aligned}
A &= \left(1-3\cos^2\theta\right) I_{1z}I_{2z} \\
B &= -\frac{1}{4}\left(1-3\cos^2\theta\right)\left(I_1^+I_2^- + I_1^-I_2^+\right) \\
C &= -\frac{3}{2}\sin\theta\cos\theta \mathrm{e}^{-\mathrm{i}\phi}\left(I_{1z}I_2^+ + I_1^+I_{2z}\right) \\
D &= -\frac{3}{2}\sin\theta\cos\theta \mathrm{e}^{\mathrm{i}\phi}\left(I_{1z}I_2^- + I_1^-I_{2z}\right) \\
E &= -\frac{3}{4}\sin\theta \mathrm{e}^{-2\mathrm{i}\phi} I_1^+I_2^+ \\
F &= -\frac{3}{4}\sin\theta \mathrm{e}^{2\mathrm{i}\phi} I_1^-I_2^-,
\end{aligned}
\tag{4.26}
$$

the angles θ and ϕ relate to the vector $\mathbf{r}$ in spherical polars, and the raising and lowering operators are $I_j^{\pm} = I_{jx} \pm \mathrm{i}I_{jy}$. For two protons ($I_1 = I_2 = \frac{1}{2}$) in the triplet
state (i.e. for $\mathbf{I} = \mathbf{I}_1 + \mathbf{I}_2$ with magnitude $I = 1$) show that the terms in the Hamiltonian above induce transitions as follows: $\Delta M = 0$ for A and $B, \Delta M = +1$ for C, $\Delta M = -1$ for $D, \Delta M = +2$ for E and $\Delta M = -2$ for F.
(4.7) A diatomic molecule is formed when two atoms bond together. Two electrons are in a molecule and have coordinates $\mathbf{r}_1$ and $\mathbf{r}_2$. There are two nuclei, one fixed at $\mathbf{R}_1$ and the other fixed at $\mathbf{R}_2$, each with charge Ze. Ignoring the repulsion between the nuclei, the Hamiltonian can be written as $\hat{\mathcal{H}}_0 + \hat{\mathcal{H}}'$ where

$$
\begin{aligned}
\hat{\mathcal{H}}_0 &= -\frac{h^2}{2m}\left(\nabla_1^2 + \nabla_2^2\right) - \frac{Ze^2}{4\pi\epsilon_0\left|\mathbf{r}_1 - \mathbf{R}_1\right|} - \frac{Ze^2}{4\pi\epsilon_0\left|\mathbf{r}_2 - \mathbf{R}_2\right|} \\
&= \hat{\mathcal{H}}_1 + \hat{\mathcal{H}}_2
\end{aligned}
\tag{4.27}
$$

which is a sum of single particle Hamiltonians $\hat{\mathcal{H}}_1$ and $\hat{\mathcal{H}}_2$ (each electron orbiting its own nucleus) and

$$
\hat{\mathcal{H}}' = \frac{e^2}{4\pi\epsilon_0\left|\mathbf{r}_1 - \mathbf{r}_2\right|} - \frac{Ze^2}{4\pi\epsilon_0\left|\mathbf{r}_1 - \mathbf{R}_2\right|} - \frac{Ze^2}{4\pi\epsilon_0\left|\mathbf{r}_2 - \mathbf{R}_1\right|},
\tag{4.28}
$$

which contains terms for the electron repulsion energy and the attraction energy for each electron being attracted by its partner's nucleus. Show that the energy of the molecule is given by

$$
E = E_1 + E_2 + \frac{K \pm J}{1 \pm S}
\tag{4.29}
$$

where E_1 and E_2 are the eigenvalues of $\hat{\mathcal{H}}_1$ and $\hat{\mathcal{H}}_2$ respectively,

$$
\begin{aligned}
K &= \int \psi_a^*\left(\mathbf{r}_1\right)\psi_b^*\left(\mathbf{r}_2\right)\hat{\mathcal{H}}'\psi_a\left(\mathbf{r}_1\right)\psi_b\left(\mathbf{r}_2\right)\mathrm{d}\mathbf{r}_1\ \mathrm{d}\mathbf{r}_2 \\
&= \int \left|\psi_a\left(\mathbf{r}_1\right)\right|^2 \hat{\mathcal{H}}'\left|\psi_b\left(\mathbf{r}_2\right)\right|^2\ \mathrm{d}\mathbf{r}_1\ \mathrm{d}\mathbf{r}_2
\end{aligned}
\tag{4.30}
$$

which is known as the Coulomb integral,

$$
\mathrm{J} = \int \psi_a^*\left(\mathbf{r}_1\right)\psi_b^*\left(\mathbf{r}_2\right)\hat{\mathcal{H}}'\psi_a\left(\mathbf{r}_2\right)\psi_b\left(\mathbf{r}_1\right)\mathrm{d}\mathbf{r}_1\ \mathrm{d}\mathbf{r}_2
\tag{4.31}
$$

which is known as the exchange integral,

$$S = \int \psi_a^* (\mathbf{r}_1)\, \psi_b^* (\mathbf{r}_2)\, \psi_a (\mathbf{r}_2)\, \psi_b (\mathbf{r}_1)\, \mathrm{d}\mathbf{r}_1\ \mathrm{d}\mathbf{r}_2 \tag{4.32}$$

which is known as the overlap integral, and where the eigenfunctions are assumed to be given by

$$\mathbf{\Psi}_{\pm} = \frac{1}{\sqrt{2}} \left[\psi_a (\mathbf{r}_1)\, \psi_b (\mathbf{r}_2) \pm \psi_a (\mathbf{r}_2)\, \psi_b (\mathbf{r}_1)\right]. \tag{4.33}$$

(4.8) Three $S = 1$ atoms are placed on the corners of an equilateral triangle and can be described by a Hamiltonian

$$\hat{\mathcal{H}} = -2J \left(\hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2 + \hat{\mathbf{S}}_2 \cdot \hat{\mathbf{S}}_3 + \hat{\mathbf{S}}_3 \cdot \hat{\mathbf{S}}_1\right). \tag{4.34}$$

Show that the energy eigenvalues of the system are $0, 2$ J and 6 J .

9.5 Order and magnetic structures

In the previous chapter the different types of magnetic interaction which operate between magnetic moments in a solid have been presented. In this chapter we will consider the different types of magnetic ground state which can be produced by these interactions. Some of these ground states are illustrated in Fig. 5.1. The different ground states include ferromagnets in which all the magnetic moments are in parallel alignment, antiferromagnets in which adjacent magnetic moments lie in antiparallel alignment, spiral and helical structures in which the direction of the magnetic moment precesses around a cone or a circle as one moves from one site to the next, and spin glasses in which the magnetic moments lie in frozen random arrangements. This chapter will be concerned with showing how, in broad terms, the interactions discussed in the previous chapter lead to these differing ground states. In the following chapter the phenomenon of order will be examined in a more general context and it will be seen that order is a consequence of broken symmetry.

9.5.1 5.1 Ferromagnetism

A ferromagnet has a spontaneous magnetization even in the absence of an applied field. All the magnetic moments lie along a single unique direction. [1] This effect is generally due to exchange interactions which were described in the previous chapter. For a ferromagnet in an applied magnetic field $\mathbf{B}$, the appropriate Hamiltonian to solve is

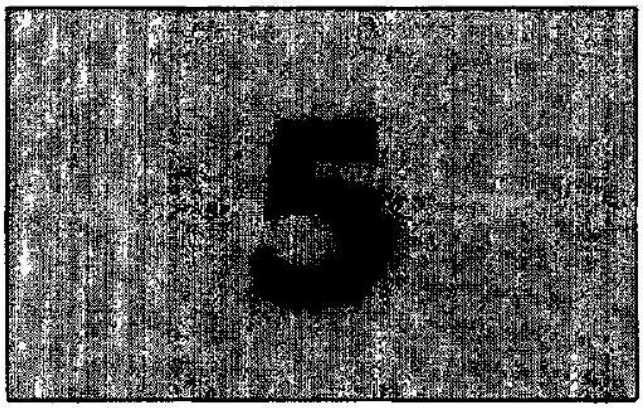
$$\hat{\mathcal{H}} = -\sum_{ij} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j + g\mu_{\mathrm{B}} \sum_j \mathbf{S}_j \cdot \mathbf{B}, \tag{5.1}$$

and the exchange constants for nearest neighbours will be positive in this case, to ensure ferromagnetic alignment. The first term on the right is the Heisenberg exchange energy (see eqn 4.7). The second term on the right is the Zeeman energy (see eqn 1.35). To keep things simple to begin with, let us assume [2] that we are dealing with a system in which there is no orbital angular momentum, so that $L = 0$ and $J = S$.

5.1.1 The Weiss model of a ferromagnet

To make progress with solving eqn 5.1 it is necessary to make an approximation. We define an effective molecular field at the i^{th} site by

$$\mathbf{B}_{\mathrm{mf}} = -\frac{2}{g\mu_{\mathrm{B}}} \sum_j \mathrm{J}_{ij}\mathbf{S}_j. \tag{5.2}$$



5.1 Ferromagnetism 85
5.2 Antiferromagnetism 92
5.3 Ferrimagnetism 97
5.4 Helical order 99
5.5 Spin glasses 100
5.6 Nuclear ordering 101
5.7 Measurement of magnetic order 102

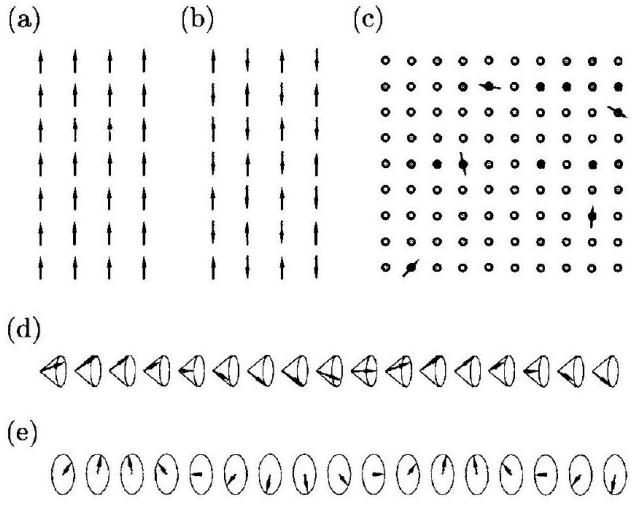
Notation reminder: In this book J refers to the exchange constant, J is total angular momentum.

[1] In fact in many ferromagnetic samples this is not true throughout the sample because of domains. In each domain there is a uniform magnetization, but the magnetization of each domain points in a different direction from its neighbours. See Section 6.7 for more on magnetic domains.

[2] We will relax this assumption later in Section 5.1.4.

[3] See Section 4.2.1.

Fig. 5.1 Various spin arrangements in ordered systems: (a) ferromagnets, (b) antiferromagnets, (c) spin glasses and (d) spiral and (e) helical structures.



Now focus on the i^{th} spin. Its energy is due to a Zeeman part $g\mu_{\mathrm{B}}\mathbf{S}_i \cdot \mathbf{B}$ and an exchange part. The total exchange interaction between the i^{th} spin and its neighbours is $-2\sum_j J_{ij}\mathbf{S}_i \cdot \mathbf{S}_j$,

where the factor of 2 is because of the double counting. [3] This term can be written as

$$-2\mathbf{S}_i \cdot \sum_j \mathrm{J}_{ij}\mathbf{S}_j = -g\mu_\mathrm{B}\mathbf{S}_i \cdot \mathbf{B}_\mathrm{mf} \tag{5.3}$$

Hence the exchange interaction is replaced by an effective molecular field $\mathbf{B}_\mathrm{mf}$ produced by the neighbouring spins. The effective Hamiltonian can now be written as

$$\hat{\mathcal{H}} = g\mu_\mathrm{B}\sum_i \mathbf{S}_i \cdot (\mathbf{B} + \mathbf{B}_\mathrm{mf}) \tag{5.4}$$

which now looks like the Hamiltonian for a paramagnet in a magnetic field $\mathbf{B} + \mathbf{B}_\mathrm{mf}$. The assumption underpinning this approach is that all magnetic ions experience the same molecular field. This may be rather questionable, particularly at temperatures close to a magnetic phase transition, as will be discussed in the following chapter. For a ferromagnet the molecular field will act so as to align neighbouring magnetic moments. This is because the dominant exchange interactions are positive. (For an antiferromagnet, they will be negative.)

Since the molecular field measures the effect of the ordering of the system, one can assume that

$$\mathbf{B}_\mathrm{mf} = \lambda\mathbf{M} \tag{5.5}$$

where λ is a constant which parametrizes the strength of the molecular field as a function of the magnetization. For a ferromagnet, $\lambda > 0$. Because of the large Coulomb energy involved in the exchange interaction, the molecular field is often found to be extremely large in ferromagnets.

We are now able to treat this problem as if the system were a simple paramagnet placed in a magnetic field $\mathbf{B} + \mathbf{B}_\mathrm{mf}$. At low temperature, the moments can be aligned by the internal molecular field, even without any applied field
being present. Notice that the alignment of these magnetic moments gives rise to the internal molecular field that causes the alignment in the first place, so that this is something of a 'chicken-and-egg' scenario. At low temperature the magnetic order is self-sustaining. As the temperature is raised, thermal fluctuations begin to progressively destroy the magnetization and at a critical temperature the order will be destroyed. This model is known as the Weiss model of ferromagnetism.

To find solutions to this model, it is necessary to solve simultaneously the equations

$$\frac{M}{M_\mathrm{s}} = \mathrm{B}_J(y) \tag{5.6}$$

(see eqn 2.38) and

$$y = \frac{g_J\mu_\mathrm{B}J(B + \lambda M)}{k_\mathrm{B}T} \tag{5.7}$$

(see eqn 2.37). Without the λM term due to the molecular field, this would be identical to our treatment of a paramagnet in Section 2.4.3.

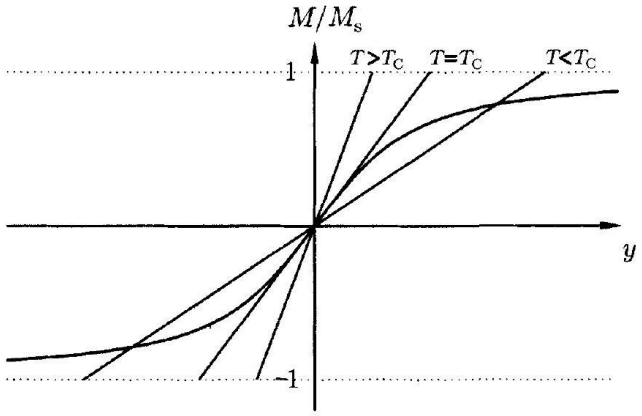


Fig. 5.2 The graphical solution of eqns 5.6 and 5.7 for $B = 0$.

These equations can be solved graphically. First, we restrict our attention to the case of $B = 0$, so that $M = k_{\mathrm{B}}Ty/g_J\mu_{\mathrm{B}}J\lambda$. Hence the straight line produced by plotting M against y has a gradient which is proportional to temperature T as illustrated in Fig. 5.2. For high temperature, there is no simultaneous solution of eqns 5.6 and 5.7 except at the origin where $y = 0$ and $M_{\mathrm{s}} = 0$. This situation changes when the gradient of the line is less than that of the Brillouin function at the origin. At low temperatures there are then three solutions, one at $M_{\mathrm{s}} = 0$ and another two for M_{s} at $\pm$ some non-zero value. It turns out that when the curve is less steep than the Brillouin function at the origin, the non-zero solutions are stable and the zero-solution is unstable. (If the system has $M_{\mathrm{s}} = 0$ for $T < T_{\mathrm{C}}$, any fluctuation, no matter how small, will cause the system to turn into either one of the two stable states.) Thus below a certain temperature, non-zero magnetization occurs and this grows as the material is cooled. The substance thus becomes magnetized, even in the absence of an external field. This spontaneous magnetization is the characteristic of ferromagnetism.

The temperature at which the transition occurs can be obtained by finding when the gradients of the line $M = k_{\mathrm{B}}Ty/g_J\mu_{\mathrm{B}}J\lambda M_{\mathrm{S}}$ and the curve $M = M_{\mathrm{s}}\mathrm{B}_J(y)$ are equal at the origin. For small y, $\mathrm{B}_J(y) = (J+1)y/3J + O\left(y^3\right)$.

Pierre Weiss (1865-1946)

Reminder: we are assuming $J = S$ and $L = 0$ at this stage.

Table 5.1 Properties of some common ferromagnets.

Fig. 5.3 The mean-field magnetization as a function of temperature, deduced for different values of J.

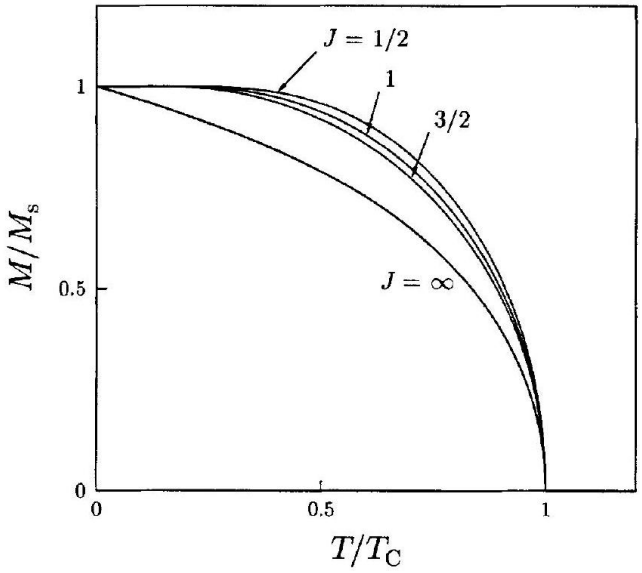
Material	T_{C}	magnetic moment (μ_{B} (K)
Fe	1043	2.22
Co	1394	1.715
Ni	631	0.605
Gd	289	7.5
MnSb	587	3.5
EuO	70	6.9
EuS	16.5	6.9

The transition temperature, known as the Curie temperature T_C, is then defined by

$$T_{\mathrm{C}} = \frac{g_J \mu_{\mathrm{B}} (J+1) \lambda M_{\mathrm{s}}}{3k_{\mathrm{B}}} = \frac{n\lambda\mu_{\mathrm{eff}}^2}{3k_{\mathrm{B}}}. \tag{5.8}$$

The molecular field $B_{\mathrm{mf}} = \lambda M_{\mathrm{s}}$ is thus $3k_{\mathrm{B}}T_{\mathrm{C}}/g_J\mu_{\mathrm{B}}(J+1)$ and so for a ferromagnet with $J = \frac{1}{2}$ and $T_{\mathrm{C}} \sim 10^3$ K, $B_{\mathrm{mf}} = k_{\mathrm{B}}T_{\mathrm{C}}/\mu_{\mathrm{B}} \sim 1500$ T. This is an enormous effective magnetic field and reflects the strength of the exchange interaction.

Fig. 5.3 The mean-field magnetization as a function of temperature, deduced for different values of J.



The solutions of these equations as a function of temperature are shown in Fig. 5.3 for a range of values of J. Although the form of the curves is slightly different in each case, some general features persist. The magnetization is zero for temperatures $T \geq T_{\mathrm{C}}$ and is non-zero for $T < T_{\mathrm{C}}$. The magnetization is continuous at $T = T_{\mathrm{C}}$, but its gradient is not. This classifies the phase transition between the non-magnetic and ferromagnetic phases in this molecular field model as a second-order phase transition. The order of a phase transition is the order of the lowest differential of the free energy which shows a discontinuity at the transition. A first-order phase transition would have a discontinuous jump in the first derivative of the free energy, i.e. in quantities like the volume, entropy or the magnetization. The jump in the entropy gives a latent heat. A second-order phase transition has a discontinuity in the second derivative of the free energy, i.e. in quantities like the compressibility or the heat capacity. In the present case the discontinuity is in the gradient of the magnetization, i.e. in the second derivative of the free energy, so the transition is second order. Phase transitions and critical exponents will be considered in more detail in Section 6.4. The properties of some common ferromagnets are listed in Table 5.1.

5.1.2 Magnetic susceptibility

Applying a small B field at $T \geq T_{\mathrm{C}}$ will lead to a small magnetization, so that the $y \ll 1$ approximation for the Brillouin function can be used. Thus

$$\frac{M}{M_{\mathrm{s}}} \approx \frac{g_J \mu_{\mathrm{B}} (J+1)}{3k_{\mathrm{B}}} \left(\frac{B + \lambda M}{T} \right) \tag{5.9}$$

so that

$$\frac{M}{M_s} \approx \frac{T_C}{\lambda M_s}\left(\frac{B+\lambda M}{T}\right). \tag{5.10}$$

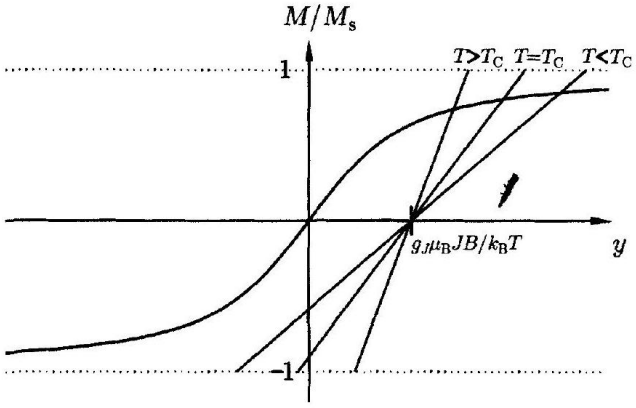
This can be rearranged to give

$$\frac{M}{M_s}\left(1-\frac{T_C}{T}\right) \approx \frac{T_C B}{\lambda M_s} \tag{5.11}$$

so that

$$\chi = \lim_{B\to 0}\frac{\mu_0 M}{B} \propto \frac{1}{T-T_C} \tag{5.12}$$

which is known as the Curie Weiss law.



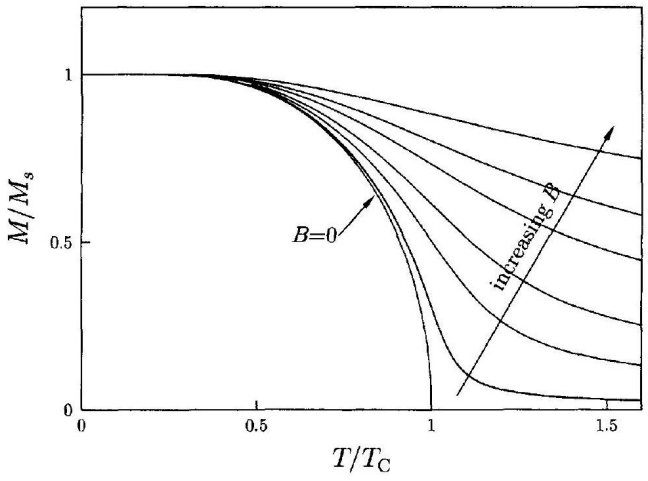
5.1.3 The effect of a magnetic field

The effect of adding a magnetic field is to shift to the right the straight line in the graphical solution of the equations (see Fig. 5.4). This results in a solution with $M \neq 0$ for all temperatures and so the phase transition is removed. For ferromagnets in a non-zero magnetic field there is always an energetic advantage to have a non-zero magnetization with the moments lining up along the magnetic field. This removal of the phase transition can be seen in Fig. 5.5 which shows graphical solutions to eqns 5.6 and 5.7 for a range of magnetic fields. In this model it is not necessary to consider the effect of applying a magnetic field in different directions. Whichever direction the magnetic field is applied, the magnetization will rotate round to follow it. The model contains no special direction associated with the ferromagnet itself. In a real ferromagnet this is not the case, and the effect of magnetic anisotropy associated with the material will need to be considered (see the following chapter).

Fig. 5.4 The graphical solution of eqns 5.6 and 5.7 for $B \neq 0$.

[04] The molecular field is a convenient fiction and one shouldn't think that magnetic fields $\sim 10^3$ T are experienced by electrons in ferromagnets. The exchange interaction is purely an electrostatic effect. When we talk about the molecular field we are pretending that the exchange interaction is actually internal magnetic field. The point is that, if it really existed and Heisenberg exchange didn't, the molecular field would have to be $\sim 10^3$ T to explain the effects we see.

Fig. 5.5 The mean-field magnetization as a function of temperature for $J=\frac{1}{2}$, calculated for different values of the applied field B. The phase transition is only present when $B = 0$.}



At $T = T_{\rm C}$, the effect of the magnetic field is simple to work out analytically. At this temperature, the magnetization is given by $M \propto B^{1/3}$ for small magnetic fields. To prove this, it is necessary to take the next term in the Taylor expansion of $\mathrm{B}_J(y)$. Writing $\mathrm{B}_J(y) = (J+1)y/3J - \zeta y^3 + O\,(y^5)$, where ζ is a constant, we must solve simultaneously $M = M_{\rm s}\mathbf{B}_J(y)$ and

$$y = \frac{g_J \mu_{\rm B} J(B + \lambda M)}{k_{\rm B} T_{\rm C}} = \frac{(B + \lambda M)}{\zeta M_{\rm s}} \tag{5.13}$$

which yields

$$M = M_{\rm s}\zeta \frac{(B + \lambda M)}{\lambda} - \zeta M_{\rm s} \left(\frac{3J(B + \lambda M)}{\lambda (J+1) M_{\rm s}} \right)^3 \tag{5.14}$$

and hence

$$B \propto (B + \lambda M)^3 \tag{5.15}$$

and given that $\lambda M \gg B$, the right-hand side is dominated by the M^3 term so that $M \propto B^{1/3}$.

5.1.4 Origin of the molecular field

When Weiss proposed his molecular field model in 1907 he was disappointed that the constant λ needed to be very large to agree with the large values of $T_{\rm C}$ found in nature. Considering only dipole fields, it was not possible to account for an internal field which, as discussed above, needs to be $\sim 10^3$ T to account for the Curie temperature of Fe . Thirty years later, Heisenberg showed that it was the exchange interaction, which involves large Coulomb energies, which is responsible for the large molecular field. [4]

The molecular field, parametrized by λ, can be related to the size of the exchange interaction, parametrized by J_{ij}. Assuming that the exchange interaction is effective only over the z nearest neighbours of an ion, where it takes
a value J, then using eqns 5.1, 5.4 and 5.5, one can easily find that

$$\lambda = \frac{2z\ \mathrm{J}}{n g^2 \mu_{\rm B}^2}. \tag{5.16}$$

Using eqn 5.8 , the Curie temperature can then be written

$$T_{\mathrm{C}} = \frac{2zJJ(J+1)}{3k_{\mathrm{B}}}. \tag{5.17}$$

In our discussion so far we have assumed that $L = 0$ and $J = S$. This works for most 3d ions. Exchange is between spin degrees of freedom and therefore depends on S. The magnetic moment of an ion depends on J, the total (spin + orbital) angular momentum. For 3d ions they are the same thing, because L is quenched.

For 4 f ions however, **S** is not a good quantum number, but **J** is. It follows that the component of **S** which is perpendicular to **J** must average to zero. The component of **S** which is parallel to **J** is conserved. Thus one must project **S** onto **J**. Now $\mathbf{J} = \mathbf{L} + \mathbf{S}$ and $\mathbf{L} + 2\mathbf{S}$ is equal to $g_J\mathbf{J}$ plus a component perpendicular to **J**. Hence the component of **S** that is a good quantum number is $(g_J - 1)\,\mathbf{J}$. Values of $(g_J - 1)$ for various 4 f ions are listed in Table 5.2. From this, it is clear that **S** and **J** are parallel for the so-called 'heavy rare earths' (Gd to Yb), but antiparallel for the so-called 'light rare earths' (Ce to Sm).

Table 5.2 The g-factors for 4 f ions using Hund's rules.

ion	shell	S	L	J	g_J	$g_J - 1$	$(g_J-1)^2 J(J+1)$
Ce^{3+}	$\mathbf{4}f^1$	$\frac{1}{2}$	3	$\frac{5}{2}$	$\frac{6}{7}$	$-\frac{1}{7}$	0.18
Pr^{3+}	$4f^2$	1	5	4	$\frac{4}{5}$	$-\frac{1}{5}$	0.80
Nd^{3+}	$4f^3$	$\frac{3}{2}$	6	$\frac{9}{2}$	$\frac{72}{99}$	$-\frac{27}{99}$	1.84
Pm^{3+}	$4f^4$	2	6	4	$\frac{3}{5}$	$-\frac{2}{5}$	3.20
Sm^{3+}	$4f^5$	$\frac{5}{2}$	5	$\frac{5}{2}$	$\frac{2}{7}$	$-\frac{5}{7}$	4.46
Eu^{3+}	$4f^6$	3	3	0	-	-	-
Gd^{3+}	$4f^7$	$\frac{7}{2}$	0	$\frac{7}{2}$	2	1	15.75
Tb^{3+}	$4f^8$	3	3	6	$\frac{3}{2}$	$\frac{1}{2}$	10.50
Dy^{3+}	$4f^9$	$\frac{5}{2}$	5	$\frac{15}{2}$	$\frac{4}{3}$	$\frac{1}{3}$	7.08
Ho^{3+}	$4f^{10}$	2	6	8	$\frac{5}{4}$	$\frac{1}{4}$	4.50
Er^{3+}	$\mathbf{4}f^{11}$	$\frac{3}{2}$	6	$\frac{15}{2}$	$\frac{6}{5}$	$\frac{1}{5}$	2.55
Tm^{3+}	$4f^{12}$	1	5	6	$\frac{7}{6}$	$\frac{1}{6}$	1.17
Yb^{3+}	$4f^{13}$	$\frac{1}{2}$	3	$\frac{7}{2}$	8 7	$\frac{1}{7}$	0.32
Lu^{3+}	$4f^{14}$	0	0	0	-	-	-

Using $(g_J - 1)\,\mathbf{J}$ for the conserved part of **S**, the expression $-\sum_{ij} J_{ij}\mathbf{S}_i \cdot \mathbf{S}_j$ can be replaced with $-\sum_{ij} (g_J - 1)^2 J_{ij}\mathbf{J}_i \cdot \mathbf{J}_j$. The magnetic moment $\mu = -g_J\mu_{\mathrm{B}}J$ can then be used to repeat the calculation leading to eqn 5.16 , resulting in

$$\lambda = \frac{2zJ\,(g_J-1)^2}{ng_J^2\mu_{\mathrm{B}}^2}. \tag{5.18}$$

Pierre-Gilles de Gennes (1932-)
(For the special case of transition metal ions with orbital quenching, $g_J = 2$ and eqn 5.18 reduces to eqn 5.16 .) The Curie temperature can then be written

$$T_{\mathrm{C}} = \frac{2z\,(g_J-1)^2\ \mathrm{J}}{3k_{\mathrm{B}}} J(J+1). \tag{5.19}$$

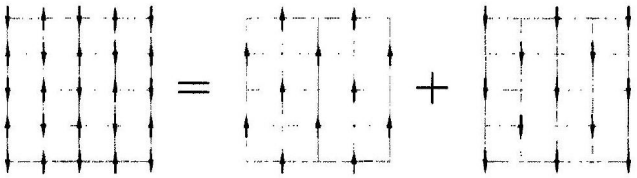
The critical temperature is therefore expected to be proportional to the de Gennes factor $(g_J - 1)^2 J(J+1)$. Values of the de Gennes factor are also listed in Table 5.2. Gd, with the

largest de Gennes factor, is a ferromagnet but the rare earth metals show a variety of different ground states, including antiferromagnetism and helimagnetism which will be considered in the next sections.

9.5.2 5.2 Antiferromagnetism

If the exchange interaction is negative, J < 0, the molecular field is oriented such that it is favourable for nearest neighbour magnetic moments to lie antiparallel to one another. This is antiferromagnetism. Very often this occurs in systems which can be considered as two interpenetrating sublattices (see Fig. 5.6), on one of which the magnetic moments point up and on the other of which they point down. The nearest neighbours of each magnetic moment in Fig. 5.6 will then be entirely on the other sublattice. Initially we will therefore assume that the molecular field on one sublattice is proportional to the magnetization of the other sublattice. We will also assume that there is no applied magnetic field.

Fig. 5.6 An antiferromagnet can be decomposed into two interpenetrating sublattices.



5.2.1 Weiss model of an antiferromagnet

If we label the 'up' sublattice + and the 'down' sublattice - then the molecular field on each sublattice is

$$\begin{aligned} B_+ &= -|\lambda| M_- \\ B_- &= -|\lambda| M_+ \end{aligned} \tag{5.20}$$

where λ is the molecular field constant which is now negative. On each sublattice, the molecular field is therefore given by

$$M_\pm = M_\mathrm{s} \mathbf{B}_J \left(-\frac{g_J \mu_\mathrm{B} J |\lambda| M_\mp}{k_\mathrm{B} T} \right). \tag{5.21}$$

The two sublattices are equivalent in everything except the direction of the moments so that

$$|M_+| = |M_-| \equiv M, \tag{5.22}$$

and hence

$$M = M_\mathrm{s} \mathbf{B}_J \left(\frac{g_J \mu_\mathrm{B} J |\lambda| M}{k_\mathrm{B} T} \right). \tag{5.23}$$

This is almost identical to the corresponding equation for ferromagnetism (eqns 5.6 and 5.7), and so the molecular field on each sublattice will follow exactly the form shown in Fig. 5.3 and will disappear for temperatures above a transition temperature, known as the Néel temperature T_N, which is then defined by

$$T_{\rm N} = \frac{g_J \mu_{\rm B}(J+1)|\lambda| M_{\rm s}}{3k_{\rm B}} = \frac{n|\lambda|\mu_{\rm eff}^2}{3k_{\rm B}}. \tag{5.24}$$

Although the magnetization on each sublattice will follow the form shown in Fig. 5.3, the two magnetizations will be in oppositedirections so that the net magnetization $M_+ + M_-$of the antiferromagnet will be zero. One can define a quantity known as the staggered magnetization as the difference of the magnetization on each sublattice, $M_+ - M_-$; this is then non-zero for temperatures below T_N and hence can be used as an order parameter [5] for antiferromagnets.

v5.2.2 Magnetic susceptibility For temperatures above $T_{\rm N}$ the effect of a small applied magnetic field can be calculated in the same way as for the ferromagnet, by expanding the Brillouin function $\mathbf{B}_J(y) = (J+1)y/3J + O\left(y^3\right)$, and results in the magnetic susceptibility χ being given by

$$\chi = \lim_{B\to 0} \frac{\mu_0 M}{B} \propto \frac{1}{T + T_N}, \tag{5.25}$$

which is the Curie Weiss law again but with the term $-T_{\rm C}$ replaced by $+T_{\rm N}$.
This result gives a ready means of interpreting susceptibility data in the paramagnetic state (i.e. for temperatures above the transition to magnetic order). The magnetic susceptibility can be fitted to a Curie Weiss dependence

$$\chi \propto \frac{1}{T - \theta}, \tag{5.26}$$

where θ is the Weiss temperature. If $\theta = 0$, the material is a paramagnet (see eqn 2.44). If $\theta > 0$ the material is a ferromagnet and we expect $\theta = T_{\rm C}$ (see eqn. 5.12). If $\theta < 0$ the material is a antiferromagnet and we expect $\theta = -T_{\rm N}$ (see eqn. 5.25). These possibilities are shown in Fig. 5.7.
Experimentally determined Weiss temperatures in antiferromagnets are often a long way from $-T_{\rm N}$ (see Table 5.3 which contains data for some common antiferromagnets). This discrepancy is largely due to the assumption we have made that the molecular field on one sublattice depends only on the magnetization of the other sublattice. A more realistic calculation is considered in Exercise 5.3.

Applying a magnetic field to an antiferromagnet at temperatures below $T_{\rm N}$ is more complicated than the case of a ferromagnet below $T_{\rm C}$ because the direction in which the magnetic field is applied is crucial. There is no longer an energetic advantage for the moments to line up along the field because any energy saving on one sublattice will be cancelled by the energy cost for

Louis E. F. Néel (1904-2000)

[5] An order parameter will be defined in Section 6.1.

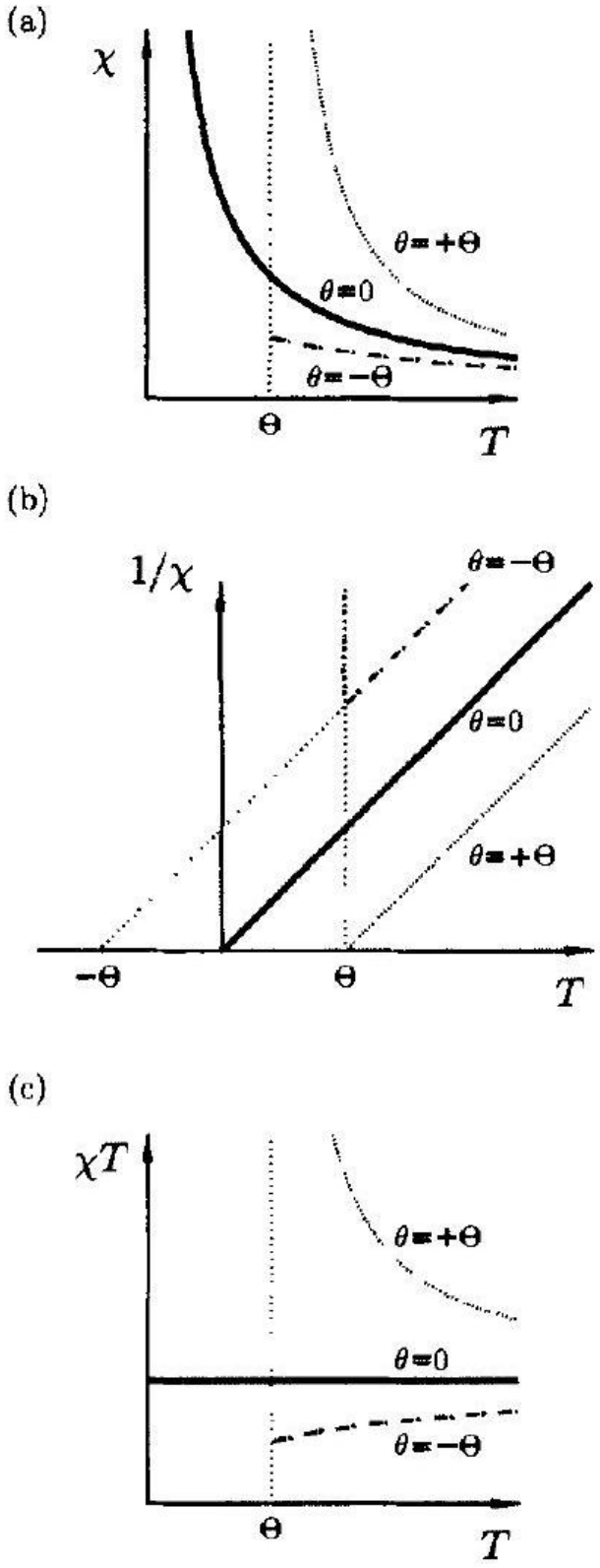


Fig. 5.7 The Curie Weiss law states that $\chi \propto 1/(T-\theta)$ for $T > \theta$. This is shown in (a) for three cases: $\theta = 0$ (paramagnet), $\theta = \Theta > 0$ (ferromagnet) and $\theta = -\Theta < 0$ (antiferromagnet). Straight-line graphs are obtained by plotting $1/\chi$ against T as shown in (b) with the intercept on the temperature axis yielding θ. A graph of χT against T can be constant ($\theta = 0$), increasing for decreasing $T(\theta > 0)$ or decreasing for decreasing $T(\theta < 0)$, as shown in (c).

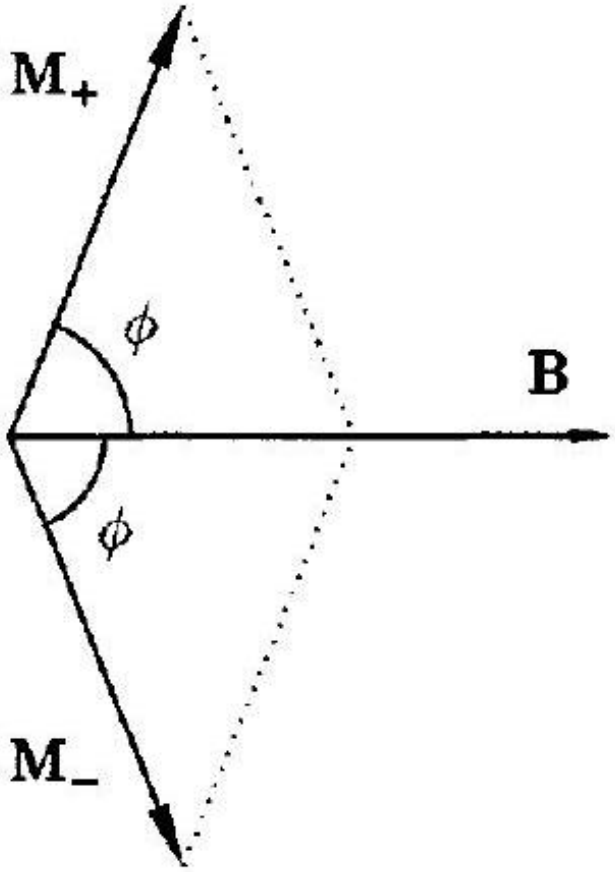


Fig. 5.8 The origin of $\chi_{\perp}$. A small magnetic field B is applied perpendicular to the magnetization direction of the sublattices, causing the magnetization of both sublattices to tilt slightly so that a component of magnetization is produced along the applied magnetic field.}

Table 5.3 Properties of some common antiferromagnets.

Material	T_N (K)	θ (K)	J
MnF_2	67	-80	5/2
MnO	116	-510	5/2
CoO	292	-330	3/2
FeO	116	-610	2
Cr_2O_3	307	-485	3/2
$\alpha - Fe_2O_3$	950	-2000	5/2

the other sublattice, if the magnetization on the two sublattices is equal and opposite.

Consider first the case of absolute zero ($T = 0$), so that thermal agitation effects can be ignored. $|M_+|$and $|M_-|$are both equal to M_s. If a small magnetic field is applied parallel to the magnetization direction of one of the sublattices (and hence antiparallel to the magnetization direction of the other sublattice), a small term is added or subtracted to the local field of each sublattice. Since both sublattices are already saturated, this has no effect and the net magnetization induced in the material is zero so that $\chi_{\parallel} = 0$. If instead the small magnetic field is applied perpendicular to the magnetization direction of one of the sublattices, this causes

the magnetization of both sublattices to tilt slightly so that a component of magnetization is produced along the applied magnetic field (see Fig. 5.8). Thus $\chi_\perp \neq 0$.

If the temperature is now increased, but still kept below T_N, the thermal fluctuations decrease the molecular field at each sublattice. This greatly affects the case of applying a small magnetic field parallel to the magnetization direction of one of the sublattices, since the field enhances the magnetization of one sublattice and reduces it on the other. In the perpendicular case, raising the temperature has little effect since the $M_+$and $M_-$are reduced equally and are also symmetrically affected by the small magnetic field. $\chi_\perp$ is independent of temperature, whereas $\chi\|$ rises from 0 up to $\chi \perp$ as $T \to T_N$. These characteristics are shown in Fig. 5.9.

5.2.3 The effect of a strong magnetic field

Let us first consider the effect of a strong magnetic field on an antiferromagnet with $T = 0$ to avoid any complications from thermal fluctuations. If the magnetic field is large enough, it must eventually dominate over any internal molecular field and force all the magnetic moments to lie parallel to each other. But as the field is increased, although the final end result is clear, the route to that destination depends strongly on the direction of the applied field with respect to the initial direction of sublattice magnetization.

If the applied magnetic field is perpendicular to the sublattice magnetizations, all that happens is that as the field increases the magnetic moments bend round more and more (ϕ gets progressively smaller, see Fig. 5.8) until the moments line up with the applied magnetic field.

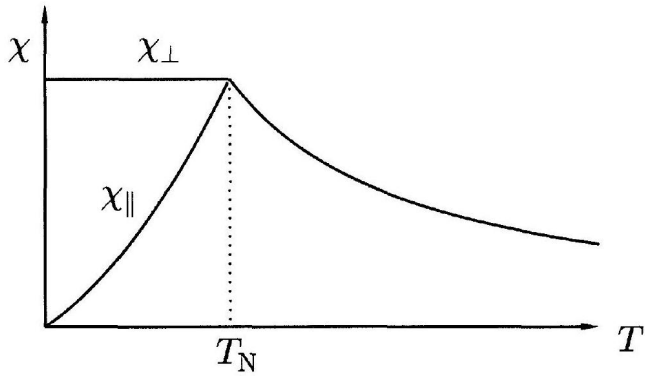


Fig. 5.9 The effect of temperature on $\chi_{|}$ $and $\chi_\perp$.}

If the applied magnetic field is parallel to the sublattice magnetizations, the case is more interesting. At small magnetic fields the moments don't rotate round but stay in line (Fig. 5.10(a)). However, at a critical field the system suddenly snaps into a different configuration (Fig. 5.10(b)); this is called a spin-flop transition. For further increases of magnetic field the angle θ gets progressively smaller until eventually the magnetic moments line up with the applied magnetic field.

These effects can be calculated quantitatively. Let $\mathbf{M}_+$lie at an angle of θ to the magnetic field (measured counterclockwise) and let $\mathbf{M}_-$lie at an angle of ϕ to the magnetic field (measured clockwise). We will apply the magnetic field along the crystallographic z axis. The antiferromagnetic phase corresponds to $\theta = 0$ and $\phi = \pi$ (see Fig. 5.10(a)) and the spin-flop phase corresponds to $\theta = \phi$. It is necessary to determine which phase has lower energy.

We assume that the total energy E is due to the sum of the Zeeman energies of the individual sublattices and a term representing the exchange coupling which will depend on the relative orientation between the two sublattice moments. This leads to

$$E = -MB\cos\theta - MB\cos\phi + AM^2\cos(\theta+\phi), \tag{5.27}$$

where A is a constant connected with the exchange coupling. To model the magnetic anisotropy, it is necessary to add on a term of the form

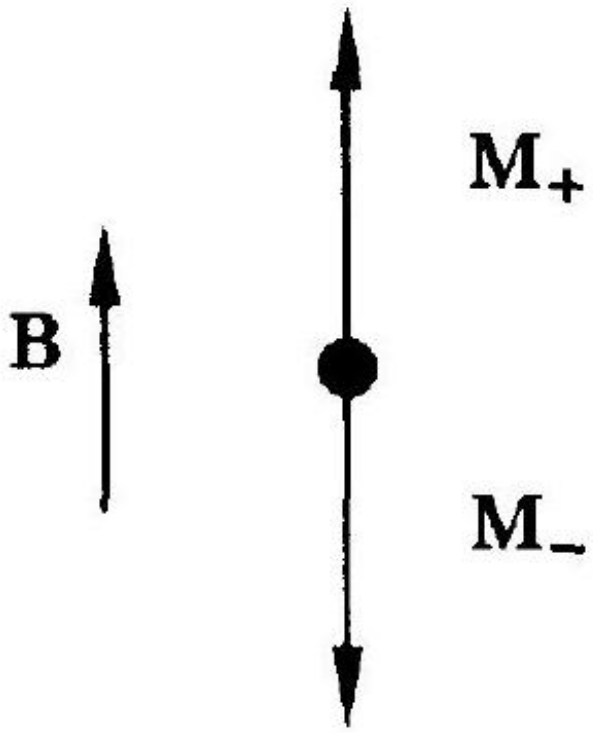
$$-\frac{1}{2}\Delta\left(\cos^2\theta + \cos^2\phi\right) \tag{5.28}$$

where Δ is a small constant. This accounts for the fact that the magnetizations actually do prefer to lie along a certain crystallographic axis (in this case the z axis) so that that θ and ϕ prefer to be 0 or π but not somewhere in between. In the antiferromagnetic case ($\theta = 0, \phi = \pi$) we have $E = -AM^2 - \Delta$ which is independent of field. In the spin-flop case ($\phi = \theta$) we have

$$E = -2MB\cos\theta + AM^2\cos 2\theta - \Delta\cos^2\theta. \tag{5.29}$$

The condition $\partial E/\partial\theta = 0$ shows that there is a minimum energy when $\theta = \cos^{-1}[B/2AM]$, ignoring the anisotropy term. Substituting this back into E

(a)



(b)

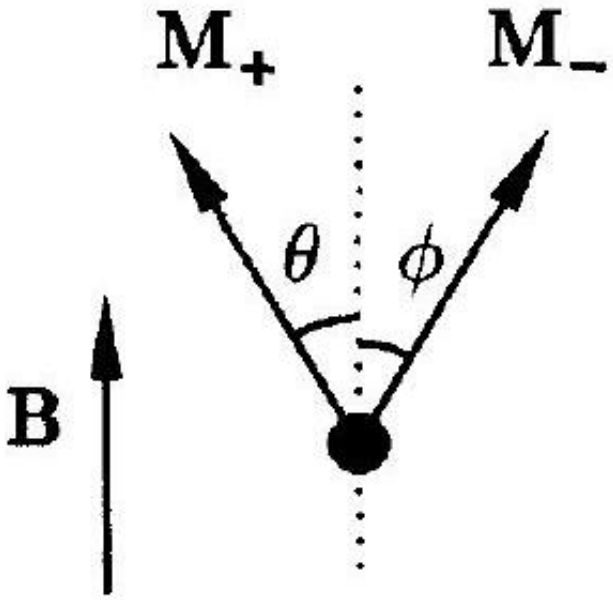


Fig. 5.10 A magnetic field is applied parallel to the sublattice magnetizations. (a) For small fields nothing happens and the system remains in the antiferromagnetic phase. (b) Above a critical field the system undergoes a spin-flop transition into a spin-flop phase.

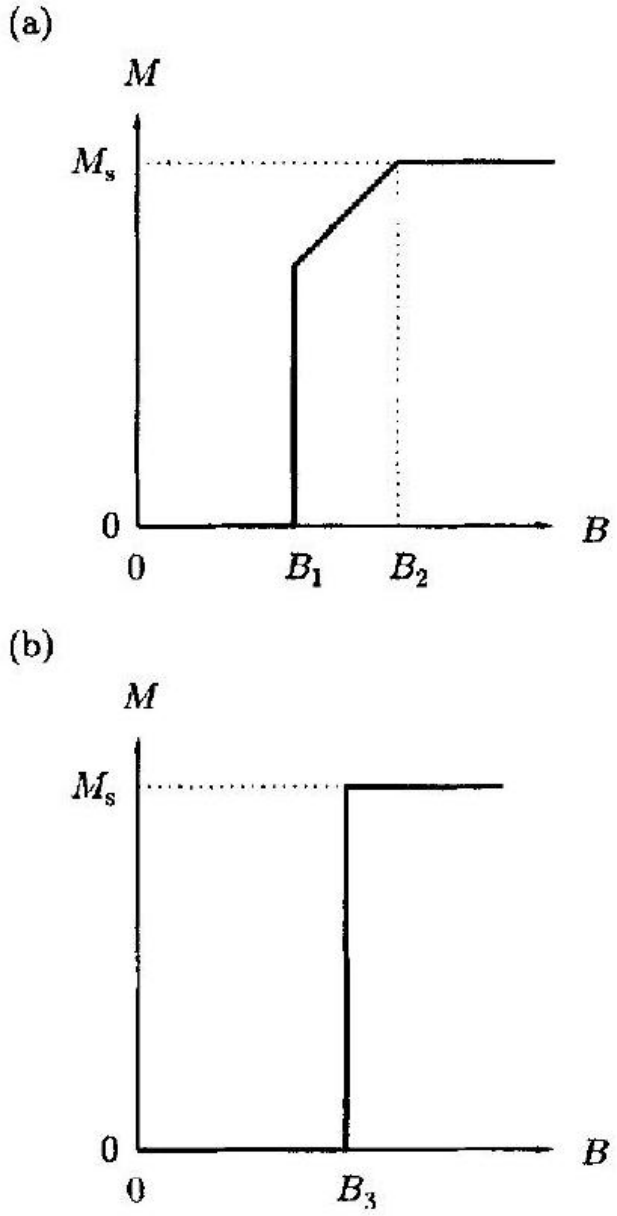
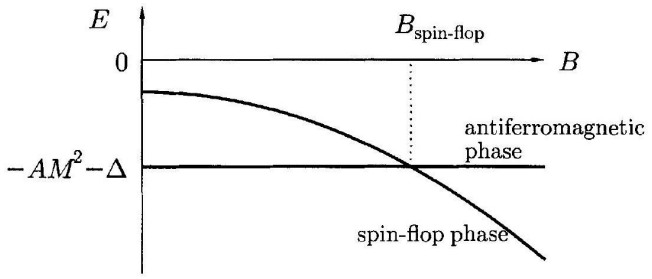


Fig. 5.12 (a) Magnetization for applying a parallel magnetic field to an antiferromagnet. Initially nothing happens but then there is a spin-flop transition to a spin-flop phase at $B_{1 $. The magnetic field then rotates the moments until saturation is achieved at the field B_2, (b) If there is a strong preference for the spins to lie along the parallel direction, no spinflop occurs. Instead there is a spin-flip transition at B_3. Both figures show the expected curves for absolute zero. Finite temperature will round off the sharp corners. This is also known as a metamagnetic transition.}

Fig. 5.11 The energy of the antiferromagnetic phase and the spin-flop phase as a function of B.



and plotting the result leads to the graph shown in Fig. 5.11. Below the critical field $B_{\text{spin-flop}}$ the antiferromagnetic case has the lowest energy. At the critical field $B_{\text{spin-flop}}$ the system switches from one state to the other and we have a spin-flop transition. Above this field the spin-flop phase has the lowest energy.

The magnetization for the antiferromagnet in a large parallel magnetic field is shown in Fig. 5.12(a). There is no effect until the spin-flop transition, above which the magnetization increases steadily until saturation is reached. If the anisotropy effect is very strong (Δ is large), another effect can occur. In this case, if the external field is along z, no spin-flop occurs. Instead we get a spin-flip transition, i.e. the magnetization of one sublattice suddenly reverses when B reaches a critical value, and the system moves in a single step to the ferromagnetic state. This is illustrated in Fig. 5.12(b).

5.2.4 Types of antiferromagnetic order

Another complication with antiferromagnetism is that there is a large number of ways of arranging an equal number of up and down spins on a lattice. The different possible arrangements also depend on the kind of crystal lattice on which the spins are to be arranged. A selection of possible arrangements is shown in Figures 5.13 and 5.14.

In cubic perovskites, which have the magnetic atoms arranged on a simple cubic lattice, G-type ordering (see Fig. 5.13(d)) is very common because superexchange interactions through oxygen atoms force all nearest-neighbour magnetic atoms to be antiferromagnetically aligned. This is the case for G-type ordering only and is, for example, found in $LaFeO_3$ and $LaCrO_3$.$LaMnO_3$ is also a cubic perovskite but shows A-type ordering (see Fig. 5.13(a)), with alternately aligned ferromagnetic (100) planes. This occurs because of the Jahn-Teller distortion of the Mn^{3+} ions which gives alternate long and short $Mn-O$ bonds within the (100) planes. The orbitals on adjacent Mn^{3+} ions are differently oriented, and the superexchange leads to an interaction between an occupied orbital on one atom with an unoccupied orbital on its neighbour. The in-plane interaction is thus ferromagnetic while the out-of-plane interaction is antiferromagnetic because of the conventional operation of the superexchange interaction.

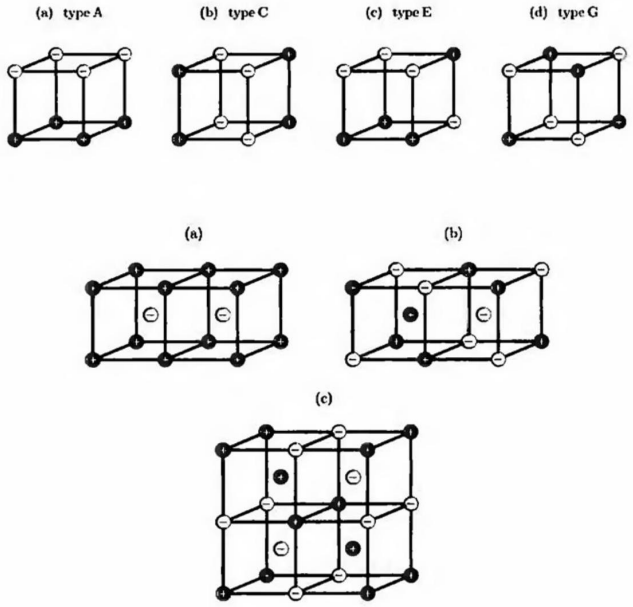


Fig. 5.13 Four types of antiferromagnetic order which can occur on simple cubic lattices. The two pussible spin states are marked and -.

Fig. 5.14 Three types of antiferromagnetic order which can occur on body-centred cubic lattices.

9.5.3 5.3 Ferrimagnetism

The above treatment of antiferromagnetism assumed that the two sublattices were equivalent. But what if there is some crystallographic reason for them not to be equivalent? In this case the magnetization of the two sublattices may not be equal and opposite and therefore will not cancel out. The material will then have a net magnetization. This phenomenon is known as ferrimagnetism. Because the molecular field on each sublattice is different, the spontancous magnetizations of the sublattices will in general have quite different temperature dependences. The net magnetization itself can therefore have a complicated temperature dependence. Sometimes one sublattice can dominate the magnetization at low temperature but another dominates at higher temperature; in this case the net magnetization can be reduced to zero and change sign at a temperature known as the compensation temperature. The magnetic susceptibilities of ferrimagnets therefore do not follow the Curie Weiss law.

Ferrites are a family of ferrimagnets. They are a group of compounds with the chemical formula $MO \cdot Fe_2O_3$ where M is a divalent cation such as Zn^{2+}. $Co^{2+}, Fe^{2+}.Ni^{2+}, Cu^{2+}$ or Mn^{2+}. The crystal structure is the spinel structure which contains two types of lattice sites, tetrahedral sites (with four oxygen

Table 5.4 Properties of some common ferrimagnets.

Material	T_C (K)	magnetic moment (μ_B /formula unit)	Compensation temperature (K)
Fe_3O_4	858	4.1	-
$CoFe_2O_4$	793	3.7	-
$NiFe_2O_4$	858	2.3	-
$CuFe_2O_4$	728	1.3	-
$Y_3Fe_5O_{12}$	560	5.0	-
$Gd_3Fe_5O_{12}$	564	16.0	290
$Dy_3Fe_5O_{12}$	563	18.2	220
$Ho_3Fe_5O_{12}$	567	15.2	137

neighbours, these are known as A sites) and octahedral sites (with six oxygen neighbours, these are known as B sites). There are twice as many B sites as A sites. The two sublattices are non-equivalent because there are two types of crystallographic site and they contain two types of different ion. In normal spinels, the M^{2+} cations sit at the A sites and the Fe^{3+} ($^6S_{5/2}$ and therefore a moment of $5\mu_B$) cations sit at the B sites. In inverse spinels the M^{2+} cations sit at half of the B sites, while the Fe^{3+} cations occupy the other half of the B sites and all the A sites. In inverse spinels, the moments of the Fe^{3+} cations on the A and the B sites are antiparallel, so that the total moment of the sample is due to the M^{2+} ions only.

The case of M = Fe, i.e. Fe_3O_4 (which is a semiconductor, in contrast to the other ferrites which are insulators), has already been discussed in Section 4.2.5.

Another family of ferrimagnets is the garnets which have the chemical formula $R_3Fe_5O_{12}$ where R is a trivalent rare earth atom. The crystal structure is cubic, but the unit cell is quite complex. Three of the Fe^{3+} ions are on tetrahedral sites, two are on octahedral sites and the R^{3+} ions are on sites of dodecahedral symmetry. In yttrium iron garnet (YIG), $Y_3Fe_5O_{12}$, the Y^{3+} has no magnetic moment (it is $4d^0$) and the moments of the Fe^{3+} ions on the tetrahedral sites are antiparallel to those on the octahedral sites, so that the net moment is $5\mu_B$.

Barium ferrite ($BaFe_{12}O_{19} = BaO \cdot 6Fe_2O_3$) has a hexagonal structure. Eight of the Fe^{3+} ions are antiparallel to the other four, so that the net moment is equivalent to four Fe^{3+} ions, i.e. $20\mu_B$. In powder form, it is used in magnetic recording since it has a high coercivity (see Section 6.7.9). The properties of some common ferrimagnets are listed in Table 5.4.

Most ferrimagnets are electrical insulators and this fact is responsible for many of their practical applications. Ferromagnets are often metallic and thus are unsuitable in applications in which an oscillating magnetic field is involved; a rapidly changing magnetic field induces a voltage and causes currents (known as eddy currents) to flow in conductors. These currents cause resistive heating in a metal (eddy current losses). Many ferrimagnets therefore can be used when a material with a spontaneous magnetization is required to operate at high frequencies, since the induced voltage will not be able to cause any significant eddy currents to flow in an insulator. Solid ferrite cores are used in
many high frequency applications including aerials and transformers requiring high permeability and low energy loss, as well as applications in microwave components. Also many ferrimagnets are more corrosion resistant than metallic ferromagnets since they are already oxides.

9.5.4 5.4 Helical order

In many rare earth metals, the crystal structure is such that the atoms lie in layers. Consider first the case (relevant for dysprosium) in which there is ferromagnetic alignment of atomic moments within the layers and that the interaction between the layers can be described by a

nearest-neighbour exchange constant J_1 and a next-nearesi-neighbour exchange constant J_2. If the angle between the magnetic moments in successive basal planes (i.e. the planes corresponding to the layers) is θ (see Fig. S.15(a)), then the energy of the system can be written

$$E = -2NS^2 (J_1 \cos\theta + J_2 \cos 2\theta) . \tag{5.30}$$

where N is the number of atoms in each plane. The energy is minimized when $\partial E/\partial\theta = 0$ which yields

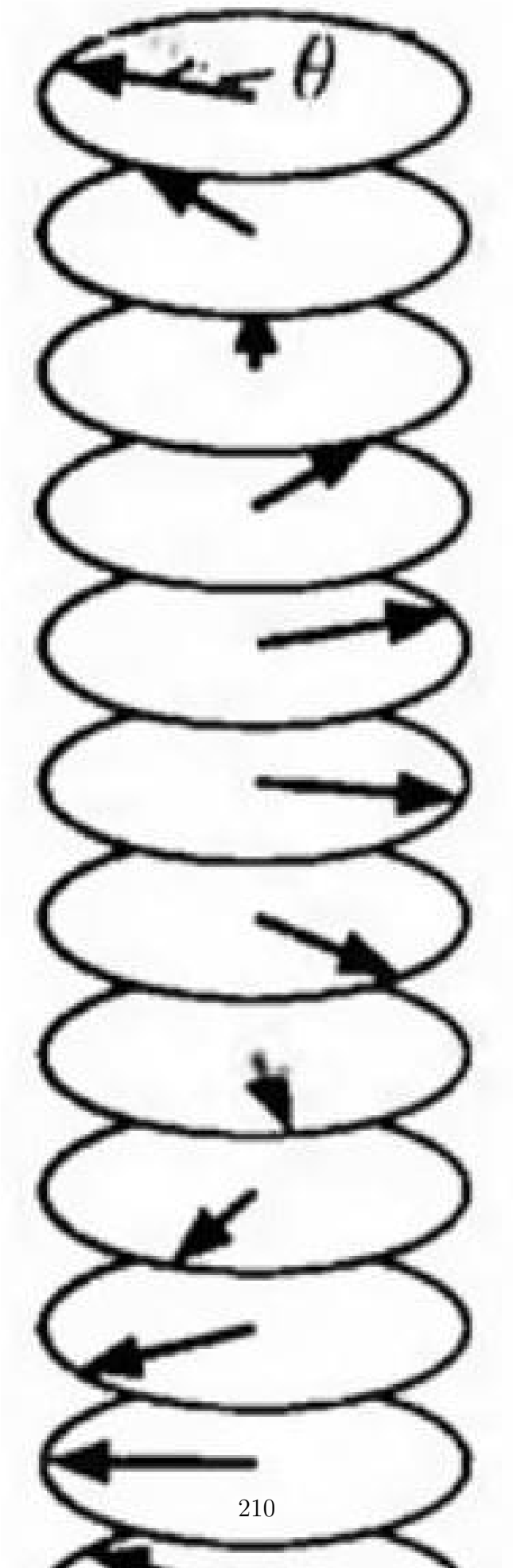
$$(J_1 + 4J_2 \cos\theta) \sin\theta = 0. \tag{5.31}$$

Solutions to this are either $\sin\theta = 0$. which implies $\theta = 0$ or $\theta = \pi$ (ferromagnetism or antiferromagnetism), or

$$\cos\theta = -\frac{J_1}{4J_2}. \tag{5.32}$$

This last solution corresponds to helical order (also known as helimagnetism) and is favoured over either ferromagnetism or antiferromagnetism when $J_2 < 0$ and $|J_1| < 4J_2$ | (sec Fig. 5.15(b)). The pitch of the spiral will not in general be commensurate with the lattice parameter and so no two layers in a crystal will have exactly the same spin directions.

(a)



(b)

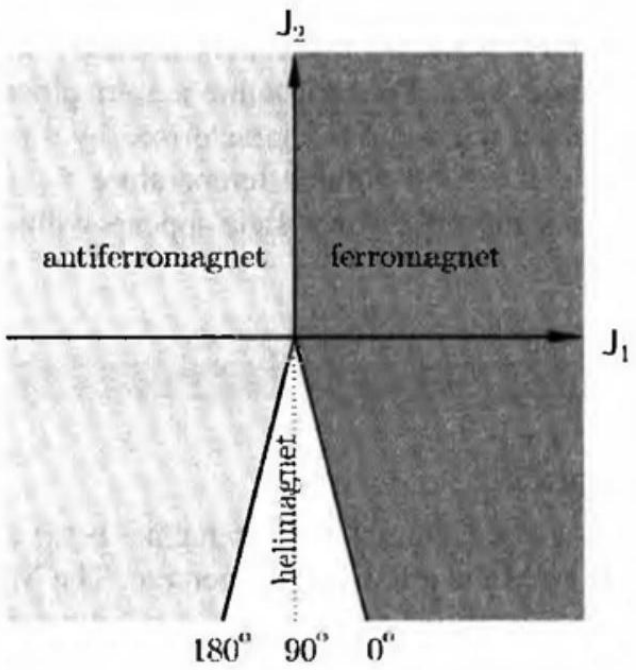


Fig. 5.15 (a) Helimagnetic ordering. (b) The phase diagram for the model of planes coupled by a nearest-neighbour exchange constant J_1 and a next-neares-neighbour exchange constant J_2.

Helical structures are found in many magnetic systems, most famously in rare earth metals. Many of these rare earth metals have hexagonally close packed crystal structures. The axis of the helix is perpendicular to the hexagonally close packed planes, along what is usually defined as the c axis. The plane in which the spins rotate in Tb, Dy and Ho is the hexagonally close packed plane, but in Er and Tm the easy axis for spins is the c axis so that the c component of the spins is modulated sinusoidally over a certain temperature range. Spiral structures (Fig. 5.1(d)) are also possible in which the spins have a component along the easy axis, perpendicular to the plane, but also a component in the plane which precesses around a helix.

The exchange interaction in rare earth metals is an indirect RKKY interaction mediated by the conduction electrons. It has much longer range than superexchange interactions and changes its sign as a function of distance. The model presented above, which contains just nearest and next-nearest neighbour interactions, is therefore an over-simplification. Details of the Fermi surface of each rare earth metal are needed to compute the wave vector-dependent exchange interaction $J(\mathbf{q})$. If this takes a maximum at a certain wave vector $\mathbf{q}$, then helimagnetism can be induced with wave vector q. A further feature is the large effect of the crystal field in rare earth metals. The crystal field splits the electronic states and usually one finds that not only the ground state but also some excited states of the system are thermally populated, further complicating the analysis that needs to be performed to understand these systems.

9.5.5 5.5 Spin glasses

So far we have considered materials in which there is one or more magnetic moments in each unit cell. What happens if we start with a non-magnetic lattice and sparsely populate it with a dilute, random distribution of magnetic atoms? One's intuition might suggest that the end result would be something which is entirely random which would not be likely to exhibit a phase transition from a high temperature disordered state to a low temperature ordered state.

This is only partly right because such systems, although inherently random, do show something approximating to a phase transition at a particular temperature to a state which, while not ordered, is distinctly different from the high temperature disordered state. We can define a spin glass as a random, magnetic system with mixed interactions characterized by a random, yet cooperative, freezing of spins at a well defined temperature T_f (the freezing temperature) below which a metastable frozen state appears without the usual magnetic long-range ordering.

Example 5.1

A well studied example of a spin glass is the alloy CuMn in which the concentration of Mn is a few atomic percent. The Mn ions are therefore present only in dilute quantities and their magnetic moments interact with each other via a RKKY (see sections 4.2.4 and 7.7.3) interaction mediated by the conduction
electrons in the Cu. The RKKY interaction oscillates in sign so that the interactions can be either antiferromagnetic or ferromagnetic. The net result is a system with a great deal of in built frustration with no well defined ground state, but a large number of alternative possible ground states. At high temperature the Mn moments thermally fluctuate, but as the temperature is reduced, locally correlated clusters build up as the moments slow down. At T_f the moments get stuck in one of these degenerate ground states and become frozen. In CuMn site-randomness implies a distribution of distances between spins and thus provides the frustration; spin glasses can also be found in systems which have bond-randomness, for which the nearest-neighbour interactions vary between +J and -J . Spin glasses and frustration will be discussed in more detail in Section 8.2.

9.5.6 5.6 Nuclear ordering

So far only electronic moments have been considered. But some nuclei possess magnetic moments; can they order? The answer is yes, but the effect is very small and only important at low temperature in materials with no electronic moments. Nuclear spins are well localized inside the atoms and do not couple strongly with each other; there can be no direct exchange interaction between them. Nuclear moments are also about three orders of magnitude smaller than electronic moments and since the dipolar interaction is proportional to the square of the magnetic moment, any dipolar interaction will be six orders of magnitude lower than in the electronic case. If electronic moments are present, they will completely dwarf any nuclear effect. The small size of the nuclear magnetic moments means that nuclear spin systems can only show magnetic order at temperatures below $\sim 1\mu$ K. There is however also the possibility of RKKY coupling between nuclear moments in metallic samples in which exchange can be mediated by the conduction electrons. However this effect is about the same order of size as the nuclear dipolar interaction.

Some remarkable experiments have nevertheless been performed on a number of materials which have no electronic moments and nuclear spin ordering has actually been observed. In copper there is a first order phase transition to an antiferromagnetically ordered nuclear ground state at 58 nK . In silver, the Néel temperature for the nuclear spins is found to be 560 pK . How may these extremely low temperatures be obtained? The experiments rely on the fact that the lattice and the nuclei can have different temperatures, even though they are in the same sample. Using adiabatic demagnetization (see Section 2.6) the nuclei can be cooled and reach thermal equilibrium among themselves in a time characterized by the spin-spin relaxation time T_2 which is typically a few milliseconds. The nuclear spins only relax back up to the lattice temperature in a time characterized by T_1 which can be as long as a few hours.

There is a further cunning trick which may be used: if the nuclei are placed in a magnetic field, the nuclear levels are split and populated according to their Boltzmann probability. If the magnetic field is then suddenly reversed (in a time which is short in comparison to T_2) the nuclear spins do not immediately move relative to their levels and one obtains a population inversion. Moreover,
one has the correct Boltzmann distribution between the levels appropriate for a negative temperature. The negative temperature state is a distinct thermodynamic state with a distinct magnetic phase diagram. Experiments on silver show that at low negative temperatures the silver nuclei order ferromagnetically at -1.9 nK .

9.5.7 5.7 Measurement of magnetic order

In the preceding sections, the various types of magnetic order have been described. But how do you measure them? In this section, various techniques for measuring magnetic order are discussed.

5.7.1 Magnetization and magnetic susceptibility

The most obvious experimental technique to try is a conventional measurement of magnetization. This is essentially a measurement of the sample's net magnetic moment μ (dividing this by the sample volume yields the magnetization).

In an extraction magnetometer, the sample is placed at the centre of a coil and then removed to a large distance, inducing a voltage V in the coil. The magnetic flux Φ produced by the sample is equal to $\int V \, \mathrm{d}t$. Usually the magnetometer is constructed so that there are two counterwound coils, arranged so that voltages due to changes in the applied field are cancelled out and only the signal from the sample remains. A vibrating sample magnetometer (commonly known as a VSM) vibrates a sample sinusoidally up and down, and then an electrical signal can be induced in a stationary pick-up coil by the movement of the magnetic moment of the sample. The signal is proportional to the magnetic moment, as well as to the amplitude and frequency of the vibration. In the alternating gradient magnetometer, the sample is mounted on a piezoelectric strip and an alternating field is provided by counterwound coils, so that at the sample there is an alternating field gradient. This results in a force on the sample equal to $(\mu \cdot \nabla)\mathbf{B}$. The piezoelectric strip therefore deflects in time with the driving frequency and by an amount that can be measured by the piezoelectric signal, allowing μ to be deduced. In a torque magnetometer, the sample is suspended on a torsion fibre. The application of a magnetic field $\mathbf{B}$ on the sample produces a torque equal to $\mu \times \mathbf{B}$. One of the most sensitive techniques uses a SQUID (superconducting quantum interference device). This is a superconducting ring which contains a 'weak-link', so that there is a Josephson junction in it, and the ring is therefore able to act like a very sensitive quantum interferometer. If a sample is passed through the ring, the persistent current induced is proportional to the magnetization of the sample. These techniques are of course useful only if the sample has a non-zero magnetization. For an antiferromagnet the magnetization is zero unless the applied magnetic field is large.

An alternative strategy is to measure the magnetic susceptibility which gives a good indication of the type of magnetic order present. This can be measured and data fitted to the Curie Weiss law. The magnetic susceptibility can be obtained by a number of methods. It is equal to the magnetization induced in a sample for a small applied magnetic field, and so the techniques described
above which measure magnetization can be used to measure magnetic susceptibility. In addition, a method which is mainly used for extracting small magnetic susceptibilities is the Gouy method. In this method a long cylindrical sample is suspended between the pole pieces of a magnet so that one end of the sample is between the pole pieces and the other end is well out

of the magnet in a region of zero field. When the magnet is turned on, the apparent weight of the sample changes. In the related Faraday method, a smaller sample can be used but this time the force is produced using a magnetic field gradient obtained with appropriately shaped pole pieces.

5.7.2 Neutron scattering

Neutrons have been found to be extremely useful in studying magnetism in condensed matter, and neutron scattering has significant advantages over other experimental techniques in the study of magnetic structure and dynamics. They give the most direct information on the arrangement of magnetic moments in a specimen.

The neutron is a spin- $\frac{1}{2}$ particle (see Table 2.3) and has a non-zero magnetic moment. Neutrons are produced in great quantities by fission reactions inside the fuel elements of nuclear reactors. The beam of neutrons which emerges from a reactor has a spectrum of energies determined by the temperature T of the moderator, usually room temperature but T can be higher or lower if the moderator is heated or cooled. Neutrons emerging from the reactor therefore have a distribution of velocities given by

$$n(v) \propto v^3 \exp\left(-\frac{\frac{1}{2}m_{\mathrm{n}}v^2}{k_{\mathrm{B}}T}\right) \tag{5.33}$$

where $n(v)\mathrm{d}v$ is the number of neutrons through unit area per second and m_{n} is the mass of the neutron. The distribution contains a Boltzmann factor and a v^3 term which is the product of a phase space factor v^2 (as for the Maxwellian distribution of velocities in kinetic theory) and a factor of v due to effusion through the hole. This function is plotted in Fig. 5.16(a). The maximum of this distribution is at

$$v = \sqrt{\frac{3k_{\mathrm{B}}T}{m_{\mathrm{n}}}} \tag{5.34}$$

which corresponds to the condition $\frac{1}{2}m_{\mathrm{n}}v^2 = \frac{3}{2}k_{\mathrm{B}}T$. The de Broglie wavelength λ of a neutron with velocity v is

$$\lambda = \frac{h}{m_{\mathrm{n}}v}. \tag{5.35}$$

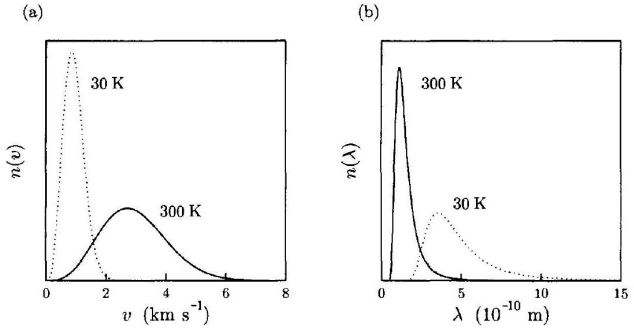
The distribution function is plotted as a function of wavelength in Fig. 5.16(b). This demonstrates that thermal neutrons have wavelengths similar to atomic spacings, thus permitting diffraction measurements to be performed. The ability to tune the wavelength means that diffraction experiments range in lengthscale from directly probing the wave function of the small atoms to the low-resolution study of macromolecules.

Louis George Gouy (1854-1926)

Michael Faraday (1791-1867)

Fig. 5.16 (a) The distribution function of neutrons as a function of velocity v for temperatures 30 K and 300 K . (b) The same distribution functions plotted as a function of de Broglie wavelength $\lambda=h / m_{\mathrm{n

} v$. The curves are normalized to have the same area.}



Neutrons scatter from the nucleus via the strong nuclear force and also from variations in magnetic field within a crystal via the electromagnetic interaction. The latter interaction enables the neutron to probe magnetic properties of materials since the neutron couples to the magnetic moment of the electrons.

We first consider the nuclear scattering. The nuclear forces which cause scattering of neutrons are very short range ($10^{-15} - 10^{-14}$ m), much smaller than the wavelength of typical thermal neutrons. Hence an incident neutron wave $\psi_i = \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathrm{r}}$ produces a spherically symmetric elastically scattered wave $\psi_f = -(b/r)\mathrm{e}^{\mathrm{i}kr}$. The quantity b is known as the scattering length and the minus sign in our expression for ψ_f ensures that $b > 0$ for a repulsive nuclear potential. The scattering length depends on the particular nucleus and the spin state of the nucleus-neutron system (which can be $I + \frac{1}{2}$ or $I - \frac{1}{2}$ for a nucleus of spin I).

In an elastic neutron scattering experiment, elastically scattered neutrons produce strong Bragg reflections when the scattering vector is equal to a reciprocal lattice vector (Fig. 5.17(a, b)). These reflections can be studied in samples by a number of experimental methods, including (i) rotating a crystal and measuring the scattering from a monochromatic beam, looking for strong reflections, (ii) using a range of incident wavelengths (the Laue method, see Fig. 5.17(c)) and (iii) measuring the diffraction of monochromatic neutrons from a powder sample (see Fig. 5.17(d)). A possible experimental arrangement at a reactor source is shown schematically in Fig. 5.17(e). Neutrons can also be produced at a spallation source. High energy protons from a synchrotron source strike a heavy metal target (e.g. ^{238}U) resulting in the production of neutrons. The proton beam from the synchrotron is pulsed, so that the neutron beam is also pulsed. The spread of energies in the pulse of neutrons can be exploited by building spectrometers that use the time-of-flight technique. By measuring the time dependence of the scattered neutrons after the start of the pulse, the scattering of a range of incident neutron wave vectors can be extracted from a single pulse (see Fig. 5.17(f)).

The nuclear scattering of neutrons results from interaction with the nucleus, rather than with the electron cloud, and so the scattering power of an atom is not strongly related to its atomic number. This is in sharp contrast with the cases of both X-ray and electron scattering. This has some advantages:

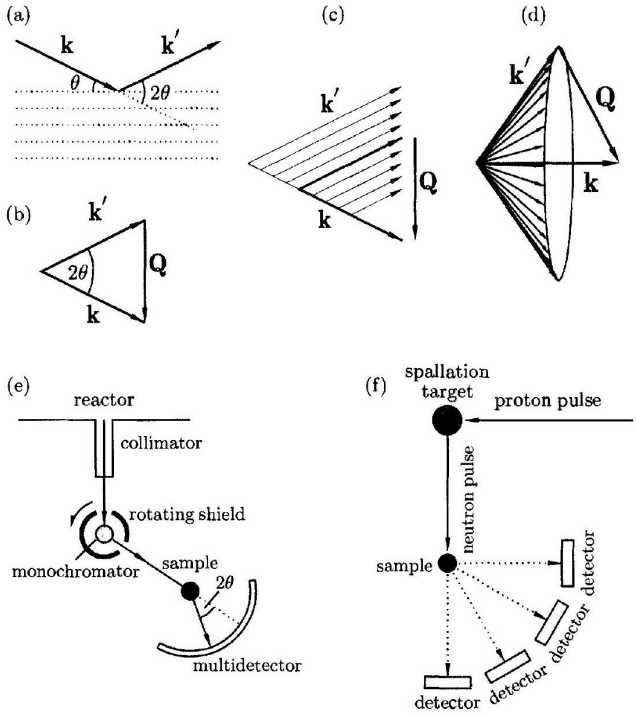


Fig. 5.17 (a) Incident neutrons with wave vector $\mathbf{k
$ are scattered to $\mathbf{k}'$ through a scattering angle of 2θ by the crystal planes. Elastic scattering implies that $|\mathbf{k}| = |\mathbf{k}'|$. The incident neutron wavelength is $\lambda = 2\pi/|\mathbf{k}|$. (b) Constructive interference of the scattered neutrons from crystal planes occurs when the scattering vector $\mathbf{Q} = \mathbf{k} - \mathbf{k}'$ equals a reciprocal lattice vector. (c) The Laue method in which a beam of neutrons with a fixed incident direction but with a range of wavelengths gives rise to a Bragg reflection when Q is a reciprocal lattice vector. (d) In a powder diffraction method, a monochromatic beam of incident neutrons is incident upon a powder sample. The scattered neutrons lie on a cone (the Debye-Scherrer cone) of semiangle 2θ. (e) Neutron diffraction at a reactor source. The monochromator and rotating shield are used to select the wavelength of the incident neutrons and the scattered neutrons from the sample are measured by a multidetector. (f) At a spallation source, a pulse of neutrons is produced by spallation from a target. The detector signals are recorded as a function of time, data from the faster (shorter wavelength) neutrons arriving earlier.}

it is easier to sense light atoms, such as hydrogen, in the presence of heavier ones, and neighbouring elements in the periodic table generally have substantially different scattering cross sections and can be distinguished. The nuclear dependence of scattering allows isotopes of the same element to have substantially different scattering lengths for neutrons, so that isotopic substitution can thus be employed. Neutrons are a highly penetrating probe, allowing the investigation, again non-destructively, of the interior of materials, rather than just the surface layers probed by techniques such as X-ray scattering, electron microscopy or optical methods.

Neutrons have no charge, which is why the electron clouds are invisible to them. However, if an atom has a net magnetic moment, the neutron's magnetic moment can directly couple to it. This is the second type of scattering mechanism which is important for neutrons: magnetic scattering. The result is that spin-up and spin-down electronic moments 'look' different to neutrons. Hence, when a sample becomes magnetic, new peaks can appear in the neutron diffraction pattern. An example of this is shown in Figures 5.18 and 5.19. which show the magnetic structure of the antiferromagnet MnO . Above $T_{\mathrm{N}} = 116$ K there are neutron diffraction

peaks resulting from the face-centred cubic

Fig. 5.18 The magnetic structure of MnO . The Mn ions are on a face-centred cubic lattice. The $\mathrm{O

ˆ{2-}$ ions are not shown (see Fig. 4.2). The Mn ions are shaded black or white depending on their spin state.}

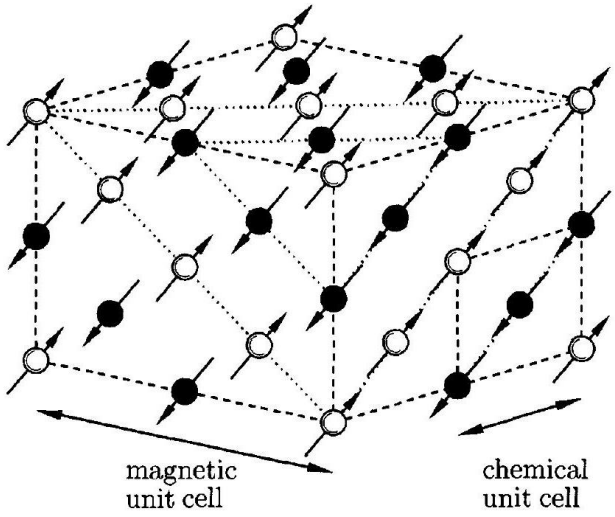
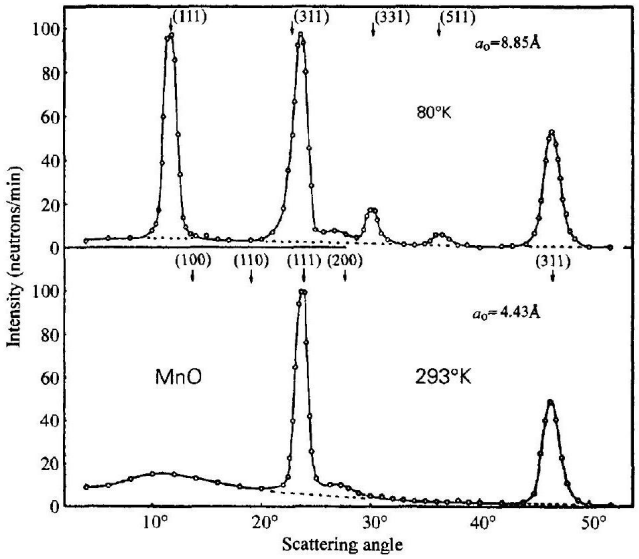


Fig. 5.19 Neutron diffraction patterns for MnO below and above $T_{\mathrm{N

}$. After C. G. Shull, W. A. Strauser and E. O. Wollan, Phys. Rev., 83, 333 (1951).}



(fcc) lattice of Mn ions, all of which are equivalent (see Fig. 5.19). The fcc lattice gives rise to systematic absences in the diffraction pattern. Below T_N, the magnetic structure is as shown in Fig. 5.18 with spins all being parallel within each (111) plane, but the moments in any two adjacent (111) planes are antiparallel. The magnetic unit cell is thus double the size of the chemical unit cell and so a few peaks suddenly appear in the neutron diffraction pattern below
T_{N}. The amplitude of the magnetic Bragg peaks can be used as a measure of the strength of the magnetic order, and hence the magnetic order can be followed as a function of temperature.

The amplitude of the scattering from a magnetic moment depends on the direction of alignment of the moment. Hence neutron diffraction can be used to determine the arrangement of the atomic magnetic moments in a magnetically ordered crystal.

In a periodic system, sharp Bragg reflections occur when the scattering vector $\mathbf{Q}$ equals a reciprocal lattice vector $\mathbf{G}$. In a helimagnet, there is an additional periodicity in the system because of the helix which can be described by the wave vector of the helix $\mathbf{q}$. The helix hence has its axis along the vector $\mathbf{q}$ and the pitch of the helix is equal to $2\pi/|\mathbf{q}|$. Bragg reflections are then observed also at $\mathbf{G} \pm \mathbf{q}$. Since the helix pitch is larger than the lattice spacing, $|\mathbf{q}| < |\mathbf{G}|$ so that these Bragg peaks observed above the magnetic transition temperature are still observed but together with smaller satellite peaks on either side of each Bragg peak below the transition temperature.

Neutron scattering is the most direct method available for determining the details of the magnetic state of a sample. Neutrons are however expensive and because neutrons interact relatively weakly with matter, large sample sizes are needed (although this drawback is most serious for inelastic experiments, as described in Section 6.6.4).

5.7.3 Other techniques

The internal field can be measured by the techniques described in chapter 3. NMR, Mössbauer and μSR experiments can all measure the temperature dependences of the magnetization of ferromagnets. Because they are local probes, i.e. they do not measure an average magnetization but rather the magnetization at a particular crystallographic site, they can be used to measure the magnetization of individual sublattices in antiferromagnets or in ferrimagnets. They can also give information about the spontaneous magnetization from within a magnetic domain, even if the sample as a whole is unmagnetized.

A further technique to be considered is X-ray scattering. Magnetic Xray scattering is typically a rather weak effect (in diffraction experiments the magnetic scattering is nominally five orders of magnitude smaller than the non-magnetic scattering) so X-rays were, until recently, mainly used for structural rather than magnetic investigations. This has changed because of the development of X-ray resonant exchange scaltering (XRES) in which the magnetic scattering is enhanced as the energy of the X-rays is tuned through an absorption edge. Another major factor has been the development of high brilliance X-ray synchrotron sources and the provision of good beams of polarized photons. These factors counter the intrinsic weakness of the effect and have allowed X-rays to be used in the study of magnetism. X-rays have a number of advantages. They can be used to study small samples (because the diameter of the beam can be very small), they are element specific (because you tune the energy to the absorption edge of a particular atom in your sample and therefore measure a signal which is a signature of a particular element in the sample) and can give better resolution than neutrons. They can also be used to separately measure the spin and orbital contributions to the magnetization density.

Dichroism is the polarization dependence of the absorption of light (a Polaroid sheet which consists of oriented long-chain molecules is a well known example of linear dichroism at optical frequencies). Magnetic dichroism works in a similar way, and the effect is driven by asymmetries in the magnetization distribution around an atom caused by unpaired electrons. This is essentially an orbital phenomenon since a photon transfers helicity ± 1 to an absorber, but the spin-orbit interaction can lead to a dependence of the dichroism on the spin angular momentum. In the technique of magnetic X-ray circular dichroism (MXCD for short) the difference in attenuation between right-handed and left-handed circularly polarized X-rays is measured and the results used to separately infer the orbital and spin magnetic moments.

9.5.8 Further reading

- A review of experiments on nuclear magnetic ordering may be found in A. S. Oja and O. V. Lounsmaa, Reviews of Modern Physics 67, 1 (1997).
- Neutron scattering is described in M. Dove, Structure and dynamics, OUP, forthcoming, G. L. Squires, Introduction to the theory of thermal neutron scattering, CUP 1978, and W. Marshall and S. W. Lovesey, Theory of thermal neutron scattering, OUP 1971.
- A more advanced treatment may be found in Theory of neutron scattering from condensed matter by S. W. Lovesey, OUP 1984.
- X-ray scattering and absorption by magnetic materials by S. W. Lovesey and S. P. Collins, OUP 1996, provides useful information on magnetic X-ray scattering.

Exercises

(5.1) Estimate the size of the molecular field, $B_{\rm mf}$ in iron in units of Tesla. Compare this with $\mu_0 M$, the contribution to the B-field due to the magnetization. Hence explain why the concept of exchange is necessary to explain the ferromagnetism of iron. The density of Fe is 7873 kg m^{-3}, the relative atomic mass is 55.847 , the Curie temperature $T_{\rm C}$ is 1043 K and each atom carries a moment of approximately $2.2\mu_{\rm B}$.
(5.2) Generalize the Weiss model for spins $S > \frac{1}{2}$. Show that the magnetization M just below $T_{\rm C}$ is given by

$$\frac{M}{M_{\rm s}} = \left(\frac{10(S+1)^2\,(T_{\rm C}-T)}{3\,[(S+1)^2+S^2]\,T_{\rm C}}\right)^{1/2} \tag{5.36}$$

and that there is a discontinuity in the heat capacity at $T_{\rm C}$ equal to

$$\Delta C = \frac{5}{2} n k_{\rm B} \frac{(2S+1)^2-1}{(2S+1)^2+1} \tag{5.37}$$

You will need the relation

$$\begin{aligned}\mathbf{B}_S(y) =& \left[\frac{(2S+1)^2-1}{(2S)^2}\right]\frac{y}{3} \\ &- \left[\frac{(2S+1)^4-1}{(2S)^4}\right]\frac{y^3}{45} + O\left(y^5\right) y \ll 1.\end{aligned} \tag{5.38}$$

(5.3) The molecular field on each sublattice of an antiferromagnet is given by

$$\begin{aligned}B_+ &= -\Gamma M_+ - |\lambda| M_- \\ B_- &= -\Gamma M_- - |\lambda| M_+\end{aligned} \tag{5.39}$$

where Γ is a constant which expresses the contribution to the molecular field from the same sublattice. Generalize the treatment given in Section 5.2 to show that on each sublattice the magnetization is

$$M_\pm = M_{\rm s}\ {\rm B}_J\left(-\frac{g_J \mu_{\rm B} J\,(\Gamma M_\pm + |\lambda| M_\mp)}{k_{\rm B} T}\right) \tag{5.40}$$

Hence show that the Néel temperature, the temperature at which the spontaneous magnetization on each sublattice disappears, is

$$T_{\mathrm{N}} = \frac{n(|\lambda| - \Gamma)\mu_{\mathrm{eff}}^2}{3k_{\mathrm{B}}} \tag{5.41}$$

Also show that the susceptibility is $\chi \propto (T - \theta)^{-1}$ where θ is given by

$$\theta = -\frac{n(|\lambda| + \Gamma)\mu_{\mathrm{eff}}^2}{3k_{\mathrm{B}}} \tag{5.42}$$

so that $|\theta| = T_N$ only if $\Gamma = 0$.

(5.4) For the helimagnet described in Section 5.4, the energy is given by eqn 5.30. Show that the energies for ferromagnetism, antiferromagnetism and helimagnetism are given by $E_{\mathrm{FM}}, E_{\mathrm{AFM}}$ and E_{HM} respectively where

$$E_{\mathrm{FM}} = -2NS^2\,(\ \mathrm{J}_1 + \mathrm{J}_2) \tag{5.43}$$

$$E_{\mathrm{AFM}} = -2NS^2\,(-\mathrm{J}_1 + \mathrm{J}_2) \tag{5.44}$$

$$E_{\mathrm{HM}} = -2NS^2\left(\frac{-\mathrm{J}_1^2}{8\ \mathrm{J}_2} - \mathrm{J}_2\right) \tag{5.45}$$

so that helimagnetism is favoured over either ferromagnetism or antiferromagnetism when $J_2 < 0$ and $|J_1| < 4\,|J_2|$.

(5.5) Show that for $T \ll T_{\mathrm{C}}$ the Weiss model of a ferromagnet leads to a magnetization M given by

$$M = M_{\mathrm{s}}\left(1 - 2\mathrm{e}^{-2T_{\mathrm{C}}/T}\right) \tag{5.46}$$

Note that this is in disagreement with the experimentally determined power law behaviour for $T \ll T_{\mathrm{C}}$ which follows

$$M = M_{\mathrm{s}}\left(1 - AT^{3/2}\right) \tag{5.47}$$

where A is a constant. This low temperature behaviour can be explained by the effect of spin waves (see Section 6.6).

(5.6) The susceptibility χ of a ferrimagnet can be treated using mean-field theory. Consider a ferrimagnet with two inequivalent sublattices such that the molecular fields B_1 and B_2 on sublattice 1 and 2 are given by

$$B_1 = \mu_0 H - \lambda M_1 \tag{5.48}$$

$$B_2 = \mu_0 H - \lambda M_2, \tag{5.49}$$

where M_1 and M_2 are the magnetization on each sublattice and H is the applied field. However, the two sublattices have different Curie constants C_1 and C_2 (i.e. different values of $g_J J$). Show that the magnetic susceptibility is given by

$$x = \frac{1}{T^2 - \theta^2}\left[(C_1 + C_2)\,T - \frac{2\lambda C_1 C_2}{\mu_0}\right], \tag{5.50}$$

and find a value for the transition temperature θ. Show that if $C_1 = C_2 = C$ the susceptibility reduces to that for an antiferromagnet with $\theta = \lambda C/\mu_0$.

(5.7) Show that the Hamiltonian for the Heisenberg model

$$\hat{\mathcal{H}} = -\sum_{\langle ij \rangle} \mathrm{J}\mathbf{S}_i \cdot \mathbf{S}_j, \tag{5.51}$$

can be rewritten as

$$\hat{\mathcal{H}} = -\sum_{\langle ij \rangle} \mathrm{J}\left[S_i^z S_j^z + \frac{1}{2}\left(S_i^+ S_j^- + S_i^- S_j^+\right)\right] \tag{5.52}$$

Hence show that for a Heisenberg ferromagnet ($J > 0$) the state $|\Phi\rangle$, which consists of spins on every site pointing up (say), is an eigenstate of the Hamiltonian and has energy E_0 given by

$$E_0 = -NS^2\ \mathrm{J} \tag{5.53}$$

Consider the Heisenberg antiferromagnet ($J < 0$) with the spins residing on two sublattices, each spin interacting only with those on the other sublattice. Show that the 'obvious' ground state, namely one with each sublattice ferromagnetically aligned but with oppositely directed sublattice magnetizations, is not an eigenstate of the Hamiltonian. This emphasizes that the Heisenberg antiferromagnet is a complex and difficult problem.

(5.8) MnF_2 has a tetragonal crystal structure in which the Mn ions are situated at the corners of the tetragonal unit cell $a = b = 0.5$ nm and $c = 0.3$ nm, and at the body-centred position in the unit cell. Below 70 K the spins of the Mn ions become antiferromagnetically aligned along the c axis with the spins of an ion at the centre of the unit cell aligned opposite to those at the corners. The neutron scattering from a powdered sample of MnF_2 is measured using an incident neutron wavelength of 0.3 nm and an angle of scattering between 0° and 90°. Sketch the results you would expect to observe at (a) 100 K and (b) 10 K . You can neglect the scattering from the F ions.

(5.9) It is possible to find exact solutions to the Weiss model. With no applied field the problem reduces to solving simultaneously the equations

$$\tilde{m} = \frac{M}{M_\mathrm{s}} = \mathrm{B}_J(y) \tag{5.54}$$

and

$$y = \frac{g_J \mu_\mathrm{B} J \lambda M}{k_\mathrm{B} T} \tag{5.55}$$

where $\tilde{m}$ is the reduced magnetization and $M_\mathrm{s} = ngJ\mu_\mathrm{B}J$. It is helpful to define $x = e^y$.

(a) For $J = \frac{1}{2}, g_J = 2$ and $\mathrm{B}_J(y) = \tanh y$. Show that this implies that

$$\tilde{m} = \frac{x - \frac{1}{2}}{x + \frac{1}{2}} \tag{5.56}$$

and hence that

$$\log x = \frac{1}{2}[\log(1 + \tilde{m}) - \log(1 - \tilde{m})] \tag{5.57}$$

Writing $T_\mathrm{C} = n\mu_\mathrm{B}^2 \lambda / k_\mathrm{B}$, show that

$$\frac{T}{T_\mathrm{C}} = \frac{2\bar{m}}{\log(1 + \tilde{m}) - \log(1 - \tilde{m})} \tag{5.58}$$

(b) For $J = 1$, show that

$$\mathbf{B}_1(y) = \frac{2\sinh y}{2\coth y + 1} \tag{5.59}$$

and hence

$$\tilde{m} = \frac{x - \frac{1}{x}}{x + 1 + \frac{1}{x}} \tag{5.60}$$

with

$$x = \exp\left(2\lambda M \mu_{\mathrm{B}}/k_{\mathrm{B}}T\right) \tag{5.61}$$

Show that this implies that

$$x = \frac{\tilde{m} \pm \sqrt{4 - 3\tilde{m}^2}}{2(1 - \tilde{m})} \tag{5.62}$$

and hence (think carefully about the $\pm$ sign) that this implies that

$$\frac{T}{T_{\mathrm{C}}} = \frac{\frac{3}{2}\tilde{m}}{\log\left(\tilde{m} + \sqrt{4 - 3\tilde{m}^2}\right) - \log(2(1 - \tilde{m}))} \tag{5.63}$$

with $T_{\mathrm{C}} = 8n\mu_{\mathrm{B}}^2\lambda/3k_{\mathrm{B}}$. The approach given in this question is detailed in M.A.B. Whittaker, American Journal of Physics 57, 45 (1989).

9.6 Order and broken symmetry

The appearance of spontaneous order at low temperature is a fundamental phenomenon of condensed matter physics. Ferromagnets, antiferromagnets, liquid crystals and superconductors are all ordered phases, as is the solid state itself. All these phenomena share some fundamental properties and characteristics. For example, they are all characterized by a temperature dependence in which some relevant physical property shows a marked difference above and below a critical temperature T_{c}. For each phase one can define an order parameter which is zero for $T > T_{\mathrm{c}}$ and non-zero for $T < T_{\mathrm{c}}$. This quantity therefore acts as an indicator of whether or not the system is ordered. In the case of ferromagnetism, the order parameter is simply the magnetization. In this chapter we will show that each type of ordered phase is associated with a broken symmetry, a concept which is explained in the following section.

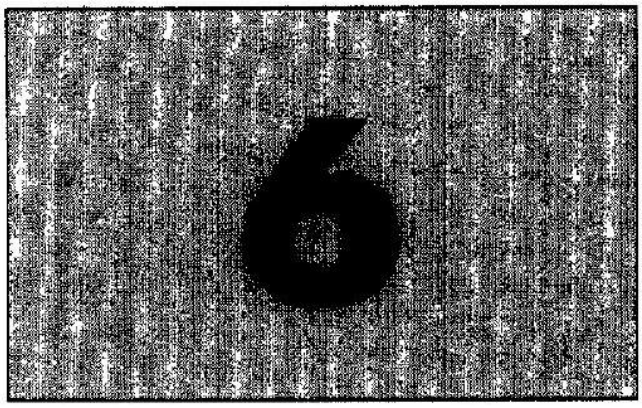
9.6.1 6.1 Broken symmetry

In ferromagnets a single unique direction has been chosen along which all the atomic magnetic moments have lined up. They have all chosen 'up' rather than 'down'. This is actually a rather surprising effect since the underlying physical equations do not distinguish between 'up' and 'down'. Thus the microscopic physics has a symmetry not possessed by the experimentally observed ground state. To delve further into this mystery, we must consider symmetry in greater detail.

A useful example to consider is the Euler strut, shown in Fig. 6.1. This consists of a vertical rod which is clamped into the ground at the bottom and uniformly loaded from the top. As the weight of the load increases, the rod compresses, and above a critical weight it buckles. Whether it buckles to the left or the right depends on the precise details of exactly how symmetrically one places the load on the top. But even in the idealization that it is placed perfectly centrally on the top, the rod still has to choose which way to buckle. Balanced precariously, as if on a knife-edge, a random thermal fluctuation might be enough to send it buckling one way or the

other. The left-right symmetry, inherent in the lightly loaded rod, is said to be broken by the buckling.

Many similar features are observed in condensed matter systems. The parameter which drives the symmetry-breaking transition can be a force, such as an applied pressure, but it is very often temperature. Figure 6.2 shows atoms in a liquid and in a solid. As a liquid cools there is a very slight contraction of the system but it retains a very high degree of symmetry. However, below a critical temperature, the melting temperature, the liquid becomes a solid and that symmetry is broken. This may at first sight seem surprising because the picture



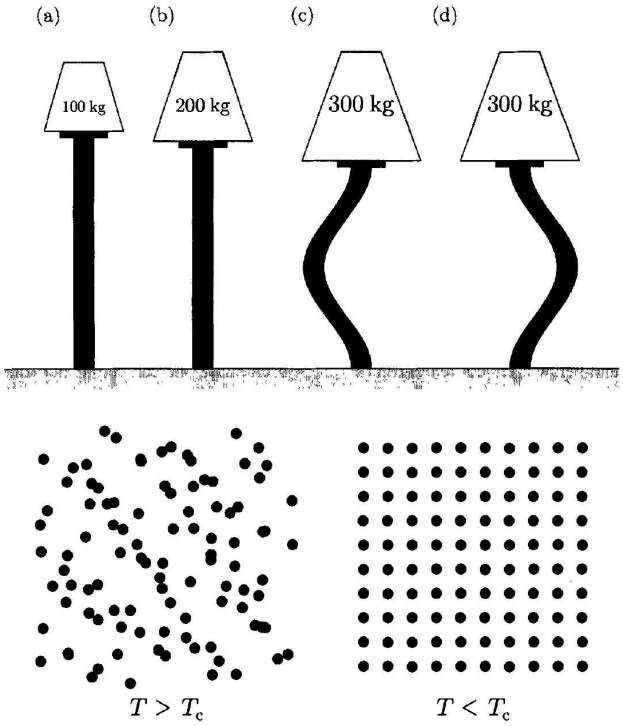
6.1 Broken symmetry 111
6.2 Models 115
6.3 Consequences of broken symmetry 117
6.4 Phase transitions 119
6.5 Rigidity 121
6.6 Excitations 121
6.7 Domains 127

Terminology concerning symmetry

A system possesses a particular symmetry if the Hamiltonian describing it is invariant with respect to the transformations associated with elements of a symmetry group. A group is a collection of elements and an operation which combines them, that follow a particular set of rules: the group is 'closed', so that combinations of elements are also members of the group; the group combination is associative; an inverse and identity are well defined.

A discrete symmetry refers to a symmetry group with countable elements, such as the rotational symmetry group of a cube. A continuous symmetry has an uncountable continuum of elements, such as the rotational symmetry group of a sphere. A system possesses a global symmetry if it is invariant under the symmetry elements of the group being applied globally to the entire system. A local symmetry applies to the Hamiltonian if it is unchanged after the symmetry operations are applied differently to different points in space. The gauge symmetry of superconductivity is such a local symmetry.

Fig. 6.1 (a) The Euler strut. (b) As an increasing vertical force is applied to the strut, the strut is compressed but retains left-right symmetry. (c) Once the force exceeds a critical value, the strut buckles, breaking the leftright symmetry. (d) The strut could also break the symmetry in the opposite sense.



of the solid 'looks' more symmetrical than that of the liquid. The atoms in the solid are all symmetrically lined up while in the liquid they are all over the place. The crucial observation is that any point in a liquid is exactly the same as any other. If you average the system over time, each position is visited by atoms as often as any other. There are no unique directions or axes along which atoms line up. In short, the system possesses complete translational and rotational symmetry. In the solid however this high degree of symmetry is nearly all lost. The solid drawn in Fig. 6.2 still possesses some residual symmetry: rather than being invariant under arbitrary rotations, it is invariant under four-fold rotations ($\pi/2, \pi, 3\pi/2, 2\pi$); rather than being invariant under arbitrary translations, it is now invariant under a translation of an integer combination of lattice basis vectors. Therefore not all symmetry has been lost but the high symmetry of the liquid state has been broken.

The situation is similar for a ferromagnet (Fig. 6.3). A ferromagnet above the Curie temperature T_C possesses complete rotational symmetry. All directions are equivalent and each magnetic moment can point in any direction. Below T_C the system 'chooses' a unique direction for all the spins to point. The

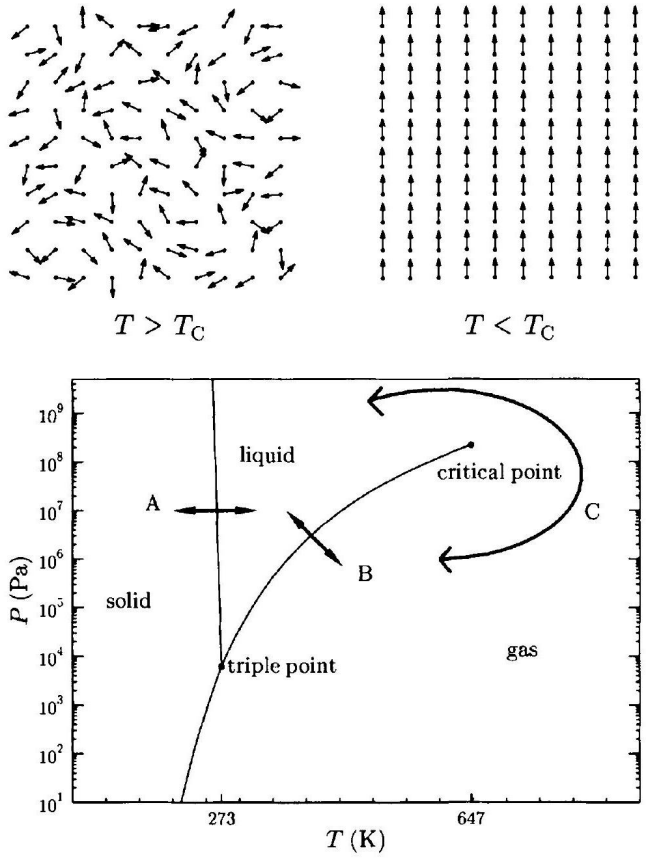


Fig. 6.3 Paramagnet-ferromagnet phase transition. Left: The high temperature state (statistically averaged) has rotational symmetry. Right: This symmetry is broken as the system becomes a ferromagnet below the Curie temperature T_{C}.

higher rotational symmetry of the high temperature state is broken. The lower temperature state has a reduced rotational symmetry (only rotations about the magnetization axis, the up-direction in Fig. 6.3, are allowed).

An important point to notice is that it is impossible to change symmetry gradually. Either a particular symmetry is present or it is not. Hence, phase transitions are sharp and there is a clear delineation between the ordered and disordered states. The appearance of order at low temperatures can then be understood from quite general thermodynamic considerations. The free energy F is related to the energy E and the entropy S by $F = E - TS$. In order to minimize the free energy, at low temperature a system will choose its lowest energy ground state which is usually ordered. Thus it will minimize E. As the temperature increases, it becomes more important in minimizing F to find a state which maximizes S and so a disordered state is favoured.

Not all phase transitions involve a change of symmetry. Figure 6.4 shows the phase diagram of water. The boundary line between the liquid and gas regions is terminated by a critical point. Hence it is possible to 'cheat' the

Fig. 6.4 The phase diagram of water. Paths A and B in the phase diagram cross a phase boundary, while C does not. Path A involves a change of symmetry, while B and C do not.

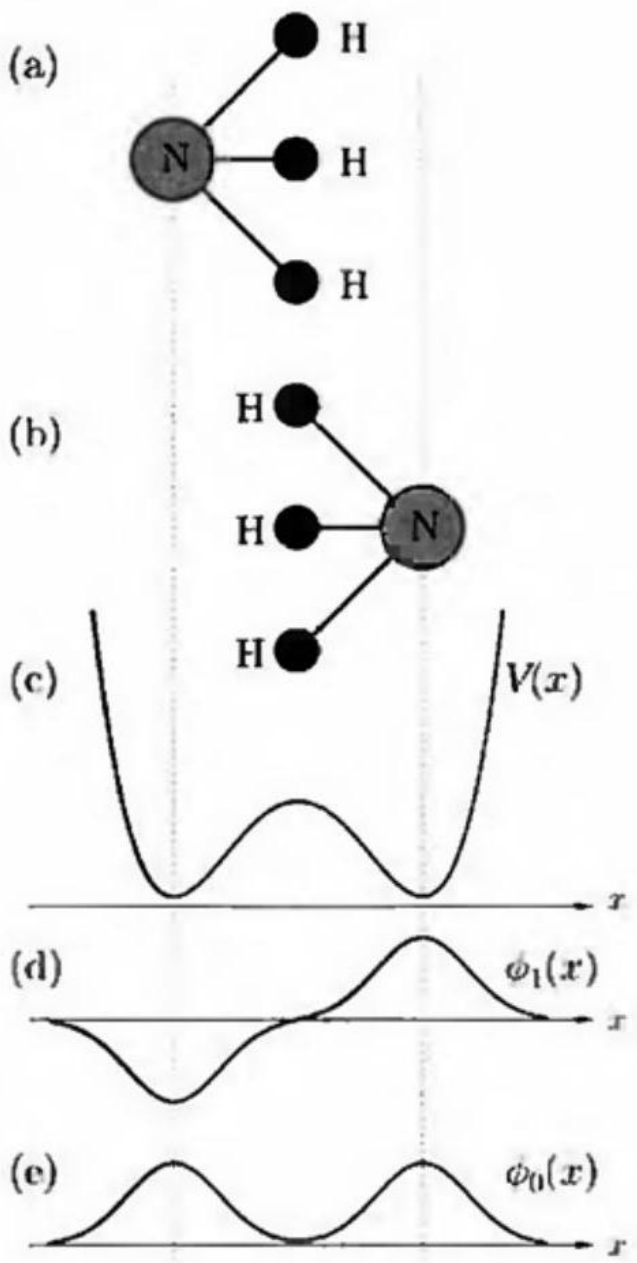


Fig. 6.5 (a) Ammonia. $\mathrm{NH
_{3}$, is a tetrahedral molecule which can exist in one of two forms. The other form is shown in (b). There are two stable positions of the N atom. and we can thus detine a porential energy Vial exsociated with tho system. 10) The tirs excited state wave function $\varphi_I(x)$. (e) The ground state wave function $\phi_{11}(x)$.}

sharp phase transition by taking a path through the phase diagram which avoids a discontinuous change (path C in Fig. 6.4). For temperatures above the critical temperature (647 K) a gas and a liquid are distinguished only by their density. The transition between a gas and a liquid involves no change of symmetry. In contrast, the solid-liquid transition involves a change of symmetry and consequently there is no critical point for the melting curve.

A puzzle about the nature of symmetry-breaking transitions is that the broken-symmetry ground state does not possess the symmetry of the Hamiltonian. On cooling a ferromagnet through the transition temperature, the system has to choose a particular direction along which all the epins will point, even though no particular direction is singled out in the underlying physics. How can this come about?

To understand this, it is useful to consider a non-magnetic example, that of the molecule ammonia (NH_3) shown in Fig. 6.5. The three hydrogen atoms form three corners of an equilateral triangle (viewed edge-on in Figures 6.5(a) and 6.5(b)) and the nitrogen atom can be just above the plane of the triangle or just below it. This pyramidal shape is due to the lone pair of electrons on the nitrogen which jockeys for space with the three nitrogen-hydrogen bonds. This gives the molecule a dipole moment, but alon defines a particular direction in space and breaks the symmetry of the underlying Hamiltonian. Therefore this pyramid cannot be a stable state of the system.

In fact we can imagine transforming between two pyramidal states shown in Figures 6.5(a) and (b) by moving the nitrogen through the middle of the plane, thus inverting the tetrahedron much in the same way as an umbrella is inverted when it is blown out by a strong wind. If we denote the position of the nitrogen atom by x, the distance from the plane, we could imagine a potential energy curve $V(x)$ which has (wo) minima associated with the two stable configurations. This is shown in Fig. 6.5(c). The two lowest energy eigenstates of this potential are shown in Figures 6.5(d) and (e) for the first excited state $\phi_1(x)$ and ground state $\phi_0(x)$ respectively.

The configurations represented in Fig. 6.5(a) and (b) do not therefore represent stable configurations of the system. but are superpositions of energy cigenstates. They therefore represent metastable states. Thus the state in Fig. 6.5(a) is represented by $(\phi_1(x) - \phi_1(x))/\sqrt{2}$ and the state in Fig. 6.5(b) is represented by $\phi_0(x) + \phi_1(x))/\sqrt{2}$. If you prepared the system in the metastable state in Fig. 6.5(a). you would tind that it would oscilate between that and the metastable state in Fig. 6.5(b) at a frequency given by the difference between the energy of $\phi_0(x)$ and $\phi_1(x)$. This frequency is 24 GH , in ammonia and this transition is the one used in the ammonia maser. If you replace the nitrogen for the more massive phosphorous to make PH_2. the frequency goes down by about a factor of ten. Thus as we make the system heavier we increase the time constant of the metastable state in Fig. 6.5(a).

In a ferromagnet we have a more massive state yet, involving not just heavier atoms, but nore atoms, maybe 10^{23} of them. The ground state of the system with all spins pointing in a single direction is actually not a true stationary state of the system. However it is a metastable state but with a stability time which is greater than the age of the universe. Te access the other state, the one with all spins reversed, one would have to flip every spin in the system simultancously. which is extremely unlikely.

9.6.2 6.2 Models

In order to discuss some of the consequences of symmetry breaking it is convenient to think about some simple models of magnetism and try to solve them.

6.2.1 Landau theory of ferromagnetism

A convenient model which simply produces a phase transition was provided by the Russian physicist Lev Landau and arises from some very general considerations. We write down the free energy for a ferromagnet with magnetization M as a power series in M. Because there is no energetic difference between 'up' or 'down', this power series cannot contain any odd power of M. Therefore we can write for the free energy $F(M)$ the expression

$$F(M) = F_0 + a(T)M^2 + bM^4 \tag{6.1}$$

where F_0 and b are constants (we assume $b > 0$) and $a(T)$ is temperature dependent. We can show that this system yields an appropriate phase transition if we allow $a(T)$ to change sign at the transition temperature T_C. Thus in the region of interest, near the transition, we write

$a(T) = a_0 (T - T_C)$ where a_0 is a positive constant. To find the ground state of the system, it is necessary to minimize the free energy so we look for solutions of $\partial F/\partial M = 0$. This condition implies

$$2M \left[a_0 (T - T_C) + 2bM^2 \right] = 0 \tag{6.2}$$

The left-hand side of this equation is a product of two terms, so either of them could be zero. This means

$$M = 0 \text{ or } M = \pm \left[\frac{a_0 (T_C - T)}{2b} \right]^{1/2} \tag{6.3}$$

The second condition is only valid when $T < T_C$, otherwise one is trying to take the square root of a negative number. The first condition applies above or below T_C but below T_C it only produces a position of unstable equilibrium (which can be deduced by evaluating $\partial^2 F/\partial M^2$). Thus the magnetization follows the curve shown in Fig. 6.6(b); it is zero for temperatures $T \geq T_C$ and is non-zero and proportional to $(T_C - T)^{1/2}$ for $T < T_C$.

Landau's approach to studying phase transitions is called a mean-field theory which means that it assumes that all spins 'feel' an identical average exchange field produced by all their neighbours. This field is proportional to the magnetization. This approach is identical to the Weiss model outlined in Section 5.1.1. Mean-field theories are the simplest type of theory that can be constructed to describe many different types of phase transition and give similar results in each case. They go under different names in different cases: e.g. Bragg-Williams theory for order-disorder transitions in alloys. Mean-field theories fail to explain the critical region accurately because the assumption that all regions of the sample are the same then becomes particularly misplaced.

Mean-field theories ignore correlations and fluctuations which become very important near T_C. Very near to the critical temperature, large fluctuations are seen in the order parameter. The critical region is characterized by fluctuations,

Lev D. Landau (1908-1968)

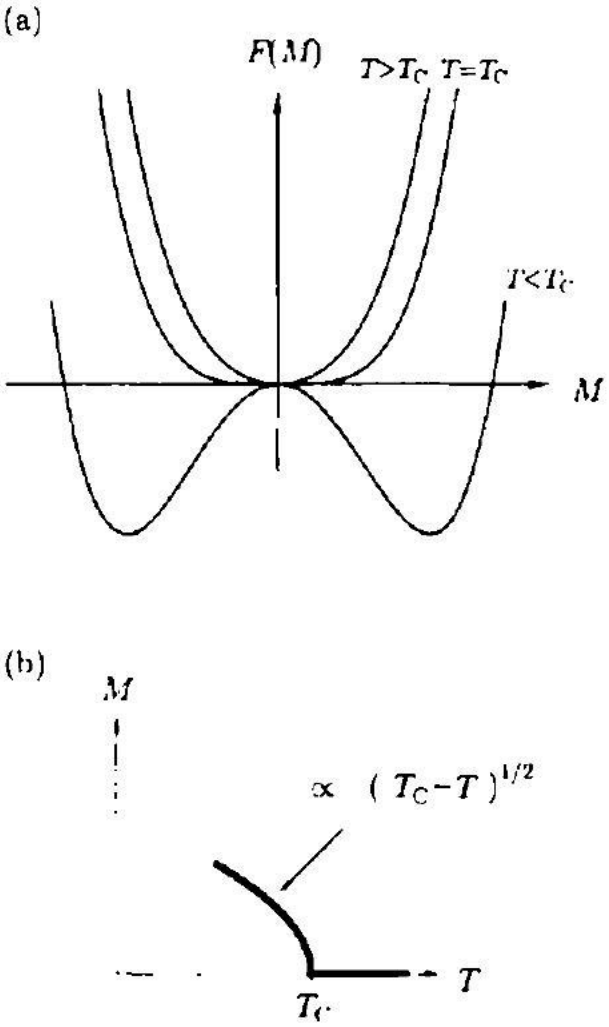


Fig. 6.6 (a) The free energy $F(M)$ of a ferromagnet. (b) The corresponding magnetization as a function of temperature.

[1] A visual demonstration of this phenomenon is critical opalescence, the blurring and clouding of images seen through a volume of gas near its critical point. This occurs because density fluctuations are strong near the critical point and give rise to large variations in refractive index.

Ernst Ising (1900-1998)

In Chapter 1, the operator for the z component of the spin was written as $\hat{S}_z$. Because we now also need to label the site of the spin, we will write the z as a superscript and put the site label as a subscript. Hence $\hat{S}_i^z$ is the operator for the z component of the spin at site i.

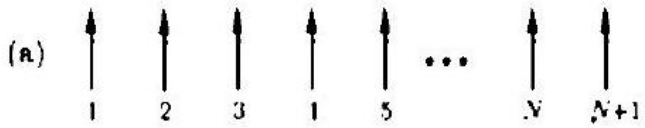


Fig. 6.7 The one-dimensional Ising model. (a) The ground state contains all $N+1$ spins aligned ferromagnetically. (b) A single defect is added.

Fig. 6.7 The one-dimensional Ising model. (a) The ground state contains all $N+1$ spins aligned ferromagnetically. (b) A single defect is added.

at all lengthscales. When a saucepan of water is heated, the water warms quite quietly and unobtrusively until near the boiling point when it makes a great deal of noise and it bubbles and foams. [1] The same type of effect occurs near T_C in a ferromagnet. The dominant length scale which characterizes the fluctuations is the correlation length ξ. As $T \to T_C, \xi \to \infty$ i.e., the correlation length becomes infinite at the critical point. The importance of fluctuations is notable also from the form of $F(M)$. The curvature at $M = 0, (\partial^2 F/\partial M^2)_{M=0}$, tends to zero at $T = T_C$ so that the minimum at $M = 0$ becomes broad and flatbottomed (see Fig. 6.6). Hence predictions about the critical region obtained from mean-field theories, though the theories are straightforward to solve, must be regarded with some caution.

6.2.2 Heisenberg and Ising models

An alternative approach to understanding the magnetic behaviour of solids is to consider particular microscopic models of the magnetic interaction. A commonly studied model is the nearest neighbour Heisenberg model which has a Hamiltonian

$$\hat{\mathcal{H}} = -\sum_{(ij)} \mathrm{J}\mathbf{S}_i \cdot \mathbf{S}_j. \tag{6.4}$$

where the constant J is the exchange integral and the symbol $\langle ij \rangle$ below the $\sum$ denotes a sum over nearest neighbours. The spins $\mathbf{S}_i$ are treated as threedimensional vectors because we allow them to point in any direction in threedimensional space. However the sum can be taken over a lattice of 1,2 or 3 dimensions. Here it is important to distinguish between the dimensionality d of the lattice on which the spins sit and the dimensionality D of the spins themselves (in general D is known as the dimensionality of the order parameter). For the Heisenberg model $D = 3$ (because the spins are threedimensional vectors). However we could be considering a lattice of these spins in 1,2 or 3 dimensions (or 4 dimensions if we wanted to!) so d can be $1, 2, 3$,

A related model is the lsing model in which the spins are only allowed to point up or down, i.e. we only consider the z component of the spin. The Hamiltonian of this model is

$$\hat{\mathcal{H}} = -\sum_{\langle ij \rangle} J S_i^z S_j^z. \tag{6.5}$$

Here the dimensionality of the order parameter D is equal to 1 (the spins are only allowed to point along $\pm\mathbf{z}$). Nevertheless we could arrange these onedimensional spins on a lattice with $d = 1.2, \ldots$

6.2.3 The one-dimensional Ising model ($D = 1, d = 1$)

If the Ising spins are placed on a one-dimensional lattice, we will show that there is no phase transition. First, consider a chain with $N + 1$ spins (and hence we need to consider N 'bonds' between each neighbour). The Hamiltonian is

$$\hat{\mathcal{H}} = -2\rfloor \sum_{i=1}^{N} S_i^z S_{i+1}^z \tag{6.6}$$

and assume J > 0 so that the ground state is obtained by having all adjacent spins lined up ferromagnetically (this is illustrated in Fig. 6.7(a)). The ground state thus has energy $-N$ J/2 because $S_i^z = +\frac{1}{2}$ for all values of z. Now consider adding one 'mistake', a single defect (see Fig. 6.7(b)). This costs an extra energy $E =$ J because we have to turn one favourable interaction (energy saving J/2) into an unfavourable one (energy cost J/2, so the change in energy is J). However, there is an entropy gain equal to $S = k_{\mathrm{B}} \ln N$ because we can put the defect in any one of N places. As we let the chain get very large ($N \to \infty$) the energy cost of a defect remains the same (J) but the entropy gain becomes infinite. The properties are determined by the free energy $F = E - TS$ so that as long as the temperature is not zero, the entropy consideration means that the presence of the defect causes the free energy to plummet to $-\infty$. This means that defects can spontaneously form and in fact no long range order occurs for $T > 0$. Another way of saying this is that the critical temperature is zero. This consideration is valid for all models on onedimensional lattices (because entropy always wins in one dimension) and we conclude that long range order is not possible in one-dimension.

6.2.4 The two-dimensional Ising model ($D = 1, d = 2$)

If the Ising spins are placed on a two-dimensional lattice, a phase transition to a magnetically ordered state below a non-zero critical temperature will result. This is because the energy cost and the entropy gain of making a defect both scale with the perimeter size of the defect (see Fig. 6.8). Energy and entropy therefore can have a fair fight with neither having an overwhelming advantage. The detailed solution of this model was one of the outstanding achievements of twentieth century statistical physics (Lars Onsager solved it in 1944) and his solution is beyond the scope of this book. This illustrates that even problems which are simple to state are by no means easy to solve.

9.6.3 6.3 Consequences of broken symmetry

When you break symmetry there are various consequences:

- Phase transitions: The system will have a sharp change of behaviour at a temperature T_{c}. We describe this by saying the system has changed phase (e.g. liquid → solid, paramagnet → ferromagnet, etc). The region near the phase transition is called the critical region (see Section 6.4).

- Rigidity: Having broken the symmetry, the system will have a strong energetic preference for staying in that broken-symmetry state, and attempts by us to change the way the system has broken the symmetry meet with resistance. Thus crystals don't bend easily and ferromagnets show permanent magnetism (see Section 6.5).

- Excitations. At $T = 0$ the system is perfectly ordered. At finite temperature this order is weakened by excitations in the order parameter. In crystals these excitations are called lattice waves, quantized into phonons. In ferromagnets the analogous modes are called spin waves, quantized into magnons (see Section 6.6).

+ + + + + + + + +

+ + + + + + + + +

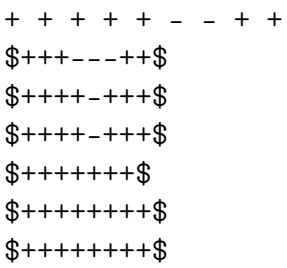


Fig. 6.8 The two-dimensional Ising model. Up spins are labelled + and down spins are labelled -. The energy and entropy of the down spin defect scale with the perimeter of the defect boundary.
Lars Onsager (1903-1976)

- Defects: If symmetry is broken differently in two adjacent parts of a macroscopic sample, the boundary will contain a defect: e.g. a dislocation in a crystal or a domain wall in a ferromagnet (see Section 6.7).
  A summary of the properties of a number of different broken symmetry phases is contained in Table 6.1. Each is characterized by a high temperature disordered state and a low temperature ordered state. Each has an order parameter. a rigidity phenomenon, a set of excitations representing wave-like departures from the ordered state, and defects associated with breaking the symmetry differently in different spatial regions.

A crystalline solid gives rise to Bragg peaks in a diffraction experiment. Thus we can use as an order parameter $\rho_{\mathbf{G}}$, the Fourier component of the charge density corresponding to a spatial frequency equal to a reciprocal lattice vector $\mathbf{G}$. The case of an antiferromagnet is similar to that of a ferromagnet, except that the order parameter is the magnetization on one sublattice of the antiferromagnet. A superconductor has a complex order parameter given by $\psi = |\psi|\mathrm{e}^{\mathrm{i}\phi}$. Here $|\psi|^2$ represents the density of superconducting electrons (for a superfluid or Bose condensate, it represents the condensate fraction). The phase ϕ of the wave function is the broken symmetry; it is free to take any value at any point in the normal state, but in the superconducting state it becomes uniform over the whole sample. For a nematic liquid crystal (see Fig. 6.9), the order parameter is given by $S = \left\langle \frac{1}{2}\left(3\cos^2\theta - 1\right)\right\rangle$, which is the average value of the function $\frac{1}{2}\left(3\cos^2\theta - 1\right)$ where θ is the angle between the long axis of the molecule and the director, $\hat{\mathbf{n}}$, a unit vector which points along the mean orientation of the axis of the liquid crystal molecule. An isotropic (liquid) state corresponds to $S = 0$. Full alignment corresponds to $S = 1$.

Table 6.1 The properties of broken symmetry phases. Here $\rho_{\mathbf{G}}$ is the Fourier component of the charge density corresponding to a spatial frequency equal to a reciprocal lattice vector $\mathbf{G}$. The complex wave function in a superconductor is $\psi = |\psi|\mathbf{e}^{\mathrm{i}\phi}$. The electric polarization $\mathbf{P}$ is the electric dipole moment per unit volume.}

Phenomenon	High T Phase	Low T Phase	Order parameter	Excitations
crystal	liquid	solid	ρ_{G}	phonons
ferromagnet	paramagnet	ferromagnet	M	magnons
antiferromagnet	paramagnet	antiferromagnet	M (on sublattice)	magnons
nematic (liquid crystal)	liquid	oriented liquid	$S = \left\langle \frac{1}{2}\left(3\cos^2\theta - 1\right)\right\rangle$	director fluctuati
ferroelectric	non-polar crystal	polar crystal	P	soft modes
superconductor	normal metal	superconductor	$\|\psi\|e^{i\phi}$	-

\section*{(a) IIIII IIIII

(c)

Fig. 6.9 A nematic liquid crystal contains rod-shaped molecules. (a) At low temperatures these order as a solid. (c) At high temperatures they are in a liquid state. (b) There is an intermediate temperature region where they are in the nematic state, in which the long-axis of the molecules become aligned. even though the positions of the centre of mass of the molecules are not ordered.

Table 6.2 Critical exponents for various models.

Model	Mean-field	Ising	Ising	Helsenberg
D	any	1	1	3
d	any	2	3	3
β	$\frac{1}{2}$	$\frac{1}{8}$	0.326	0.367
γ	1	$\frac{7}{4}$	1.2378(6)	1.388(3)
δ	3	15	4.78	4.78

9.6.4 6.4 Phase transitions

In Section 6.2 the Landau theory of ferromagnetism was presented. This is a mean-field theory in which an identical exchange field is felt by all spins. This leads to the magnetization behaving as $(T_\mathrm{C} - T)^{1/2}$ below the transition. In real systems it is found that the magnetization does behave as $(T_\mathrm{C} - T)^{\beta}$ close to the transition, but the exponent β is not necessarily equal to $\frac{1}{2}$. The exponent therefore gives important information about the nature of the phase transition. A number of other similar exponents, known as critical exponents, can be defined. Thus near the phase transition temperature T_C it is found experimentally that

$$\begin{array}{lll} \chi & \propto (T - T_\mathrm{C})^{-\gamma} & T > T_\mathrm{C} \\ M & \propto (T_\mathrm{C} - T)^{\beta} & T < T_\mathrm{C} \\ M & \propto H^{1/\delta} & T = T_\mathrm{C} \end{array}$$

where β, γ, and δ are the critical exponents. The values taken by these critical exponents for various models are shown in Table 6.2.

As stated earlier, correlations and fluctuations are ignored in mean-field theories. This means that they cannot hope to give a correct description close to T_{C}, precisely where one might wish to calculate critical exponents. It turns out however that mean-field theory is a correct description of systems with four or more dimensions (i.e. $d \geq 4$) and predicts the correct critical exponents.

Despite the failure of mean-field theory to account successfully for critical behaviour in systems with dimensionality below four, it is only necessary
to consider a small representative set of ideal statistical models to calculate critical exponents of any physical system if the hypothesis of universality is accepted. This hypothesis is based on the observation that critical exponents do seem to be surprisingly independent of the type of phase transition, whether liquid-gas, ferromagnetic-paramagnetic, superconducting-nonsuperconducting, or any other. It is supposed that for a continuous phase transition, the critical exponents depend only on
(1) The dimensionality of the system, d.
(2) The dimensionality of the order parameter, D. (Actually the symmetry of the order parameter.)
(3) Whether the forces are short or long range.

Hence, it is only important to look at particular universality classes, i.e. particular values of D and d for both short and long range forces, and pick the simplest models in each class and calculate the exponents for those. (Actually the assumptions stated above apply to static critical exponents, rather than the dynamic critical exponents which characterize time-dependent properties.)

We now need to list those models for which solutions are known. There are a number of exactly solved models which include the following:
(1) All cases for $d = 1$. As we have seen, such systems do not exhibit continuous phase transitions.
(2) All cases for $d \geq 4$, which give mean-field solutions.
(3) All cases for long range interactions, which give mean-field solutions.
(4) The case $d = 2, D = 1$. This is the 2D Ising model (see Table 6.2).
(5) The case $D = \infty$ for any d. This is known as the spherical model.

Unfortunately, most real situations correspond to $d = 3$ and short-range interactions, which have not been solved exactly.

A method of calculating critical behaviour, even in models which cannot be exactly solved, was discovered in the early 1970s, in which a block of spins with volume L^d is considered, where L is much less than the correlation length ξ. If we increase (scale) the linear dimensions of this block by a factor n, the Hamiltonian will be suitably scaled, and it is possible to examine how the order parameter is thus changed (renormalized) by this transformation. If such a transformation is denoted by τ, then the Hamiltonian $\mathcal{H}$ transforms as $\mathcal{H}' = \tau(\mathcal{H})$. At the critical point, the correlation length $\xi \to \infty$, so that there must exist a limiting function $\mathcal{H}^*$ such that $\mathcal{H}^*$ is a fixed point of τ, i.e. $\tau(\mathcal{H}^*) = \mathcal{H}^*$. This approach, namely that of looking for fixed points of scaling transformations, is known as the renormalization group method, since the set of scaling transformations τ form a group. It has proved to be particularly successful in condensed matter physics, and has deep connections with renormalization in quantum field theory. A more detailed treatment of the renormalization group is however beyond the scope of this book.

Cluster models can be useful in describing the paramagnetic regime of ferromagnets, i.e. the region $T > T_{\mathrm{C}}$. They retain some of the simplicity (and hence predictive power) of mean-field approaches but make some attempt to model the fluctuations. The idea is to evaluate the

configuration of a single finite-size cluster of spins exactly, and then to couple that to the mean field from the rest of the sample. Of course the correlation length diverges at T_C and so the size of the cluster needs to increase as you get closer to the transition. Very good agreement with data can be obtained if a cluster model is chosen in which the size of the clusters is unrestricted, and is allowed to vary. [2]

9.6.5 6.5 Rigidity

Breaking symmetry involves a choice of ground state: should the spins all point up, or down, or left or right? The energy of a macroscopic sample is minimized when the symmetry is broken the same way throughout its volume. If you try to make different parts of a macroscopic sample break symmetry differently, forces will appear reflecting an additional energy cost. This gives rise to a generalized rigidity or stiffness. This rigidity is intimately connected to the order, as can be deduced from the following argument for a crystal.

Consider a cubic crystal with lattice parameter a and imagine stretching one plane of bonds by increasing a to $a + u$ inducing a strain $\epsilon = u/a$ (see Fig. 6.10). This produces a stress equal to $G\epsilon$ (where G is an elastic modulus) so that the force on an individual bond is equal to $G\epsilon a^2 = Gau$. The energy stored in this bond is then $\frac{1}{2}Gau^2$. At high temperatures (large compared with the Debye temperature [3]) the mean stored energy in each bond is equal to $\frac{1}{2}k_BT$ (by the equipartition theorem, see Appendix E) so that

$$\langle u^2 \rangle = \frac{k_BT}{Ga} \tag{6,8}$$

and hence $\langle u^2 \rangle \propto 1/G$. Thus if G vanishes, the fluctuations diverge and vice versa. Hence the existence of a non-zero elastic modulus (i.e. a rigidity) is connected with the finiteness of the fluctuations (i.e. the stability of the crystal).

This means that because we have an ordered crystal, we must have a rigid crystal. Crystals do not deform easily (in contrast to liquids) because there is an elastic energy proportional to the elastic modulus and to $(\nabla u)^2$ where u is the lattice displacement, and the lattice transmits this force to the other end of a macroscopic sample. Solids are therefore rigid.

These ideas are also applicable to ferromagnets. The spins are aligned in a ferromagnet and it costs energy to turn them with respect to each other. Hence we have the phenomenon of permanent magnetism. If the magnetization is non-uniform there is an exchange cost proportional to $(\nabla \mathbf{M})^2$ (see Section 4.2.7, and eqn 4.18 in particular).

As a final example, consider a superconductor in which the phase of the wave function, ϕ, is the order parameter. In a superconductor, ϕ is uniform across a sample and there is a 'phase rigidity' energy proportional to $(\nabla\phi)^2$. (In fact twisting this phase across a sample produces a supercurrent $\propto \nabla\phi$.)

9.6.6 6.6 Excitations

A solid is ordered at $T = 0$, although zero-point fluctuations mean that, even then, atoms are not purely static. At non-zero temperature, the order is disrupted by thermally excited lattice vibrations, which are quantized as phonons. The behaviour of the phonons is characterized by a dispersion relation, i.e.

[2] See R. V. Chamberlin, Nature, 408, 337 (2000).

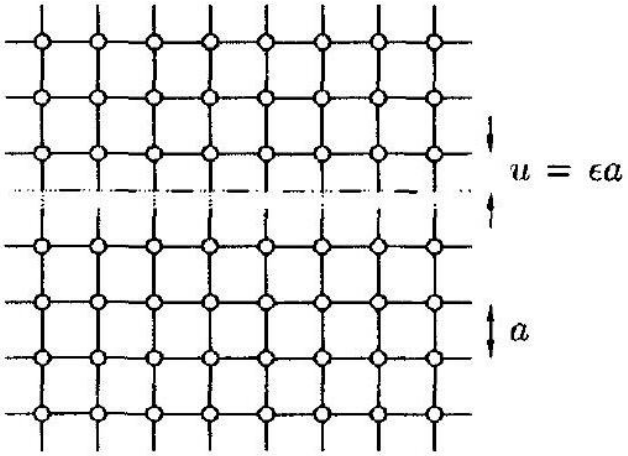


Fig. 6.10 A cubic crystal of lattice parameter a is stretched producing a strain $\epsilon = u/a$.

[3] The Debye temperature Θ_D gives the energy of the highest occupied phonon mode, divided by k_{B}, under the assumption that the phonons are non-dispersive. Of course there is always some dispersion, but Θ_D nevertheless characterizes a typical phonon energy of the system. Most Debye temperatures are in the range $100 - 1000$ K. When $T \gg \Theta_{\mathrm{D}}$, the phonon modes can be treated classically.

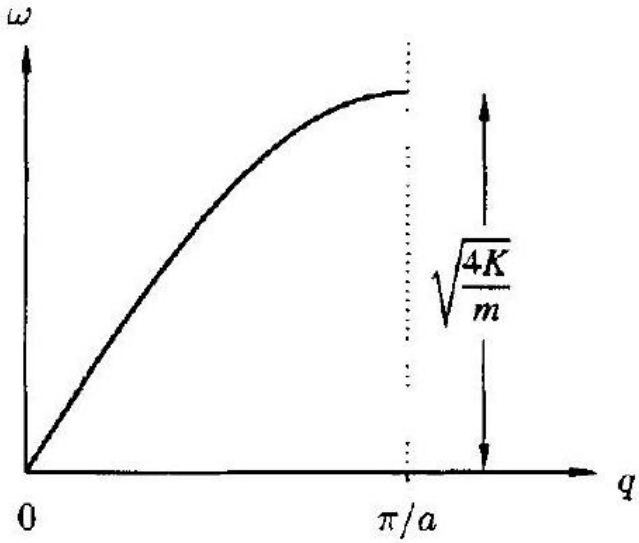


Fig. 6.11 The phonon dispersion for a onedimensional monatomic chain of atoms of mass m connected by bonds with spring constant K. These phonons are acoustic. There is no gap ($\omega = 0$ at $q = 0$) and at small $q, $\omega \approx v_{\mathrm{s} } q$ where v_{s} is the speed of sound. The dispersion is massless because $\omega = 0$ at $q = 0$ (i.e. energy $E = h\omega = 0$ at momentum $p = hq = 0$). A relativistic particle has energy $E^2 = p^2c^2 + m^2c^4$, so that the energy at zero momentum is equal to mc^2, i.e. proportional to the mass.}

a relationship between angular frequency ω and wave vector q (equivalently between energy $\hbar\omega$ and momentum $\hbar q$). An example for a one-dimensional monatomic chain is shown in Fig. 6.11. The crucial feature of this is that $\omega = 0$ at $q = 0$ for acoustic phonons, so that it costs a vanishingly small energy to produce a phonon of wave vector q, provided its wavelength $\lambda = 2\pi/q$ is long enough. The fact that an acoustic phonon can be thermally generated, as long

as the temperature is non-zero, is because there is no energy gap to leap across from the ground state of the system ($q = 0, \omega = 0$) to the lowest acoustic phonon level. There is no energy gap in the phonon dispersion relation at $q = 0$ (in contrast to the case for optic phonons). Acoustic phonons give rise to a T^3 heat capacity at low temperature in a three-dimensional crystal.

Whenever you have broken a continuous global symmetry (as you do when you make a solid from a liquid or a ferromagnet from a paramagnet) it is possible to produce long-wavelength excitations in the order parameter for vanishingly small energy cost. Such excitations are called Goldstone modes (or sometimes Goldstone bosons). Because they cost no energy they are 'massless'. In particle physics an example of this is the photon which is a Goldstone boson. For a superconductor, the situation is rather different because it turns out that you are breaking a continuous local symmetry and you don't get Goldstone modes. The reason for this is rather subtle and is connected with the Higgs mechanism.

A ferromagnet is perfectly ordered at $T = 0$ but at non-zero temperature the order is disrupted by spin waves, quantized as magnons. The crucial feature of these is that it costs a vanishingly small energy to produce a magnon, provided its wavelength is long enough. Thus the magnons play the same rôle in ferromagnets as phonons do in solids and are the Goldstone modes of the system. We will show below that, for an isotropic ferromagnet, there is no energy gap in the magnon dispersion relation at $q = 0$.

6.6.1 Magnons

In this section we will derive the magnon dispersion relation for an isotropic ferromagnet. This is an important problem so two alternative derivations will be presented, the first using a semiclassical approach and the second using a quantum mechanical approach.
(1) We begin with a semiclassical derivation of the spin wave dispersion. First, recall the Hamiltonian for the Heisenberg model,

$$\hat{\mathcal{H}} = -\sum_{\langle ij \rangle} \mathrm{J}\hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j \tag{6.9}$$

(which is eqn 6.4). In a one-dimensional chain each spin has two neighbours, so the Hamiltonian reduces to

$$\hat{\mathcal{H}} = -2J \sum_i \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_{i+1}. \tag{6.10}$$

We can calculate the time dependence of $\left\langle \hat{\mathbf{S}}_j \right\rangle$ using eqn C.7, with the result that

$$\frac{\mathrm{d}\left\langle \hat{\mathbf{S}}_j \right\rangle}{\mathrm{d}t} = \frac{1}{\mathrm{i}\hbar}\left\langle \left[\hat{\mathbf{S}}_j, \hat{\mathcal{H}}\right] \right\rangle \tag{6.11}$$

$$= -\frac{2\rfloor}{\mathrm{i}\hbar}\left\langle \left[\hat{\mathbf{S}}_j, \cdots + \hat{\mathbf{S}}_{j-1} \cdot \hat{\mathbf{S}}_j + \hat{\mathbf{S}}_j \cdot \hat{\mathbf{S}}_{j+1} + \cdots \right] \right\rangle \tag{6.12}$$

$$= -\frac{2\rfloor}{\mathrm{i}\hbar}\left\langle \left[\hat{\mathbf{S}}_j, \hat{\mathbf{S}}_{j-1} \cdot \hat{\mathbf{S}}_j\right] + \left[\hat{\mathbf{S}}_j, \hat{\mathbf{S}}_j \cdot \hat{\mathbf{S}}_{j+1}\right] \right\rangle \tag{6.13}$$

$$= \frac{2\rfloor}{\hbar}\left\langle \hat{\mathbf{S}}_j \times \left(\hat{\mathbf{S}}_{j-1} + \hat{\mathbf{S}}_{j+1}\right) \right\rangle. \tag{6.14}$$

We now treat the spins at each site as classical vectors. The ground state of the system has all the spins aligned, say along the z axis, so that $S_j^z = S, S_j^x = S_j^y = 0$. Consider a state which is a small departure from this state with $S_j^z \approx S, S_j^x, S_j^y \ll S$, so that

$$\frac{\mathrm{d}S_j^x}{\mathrm{d}t} \approx \frac{2\sqrt{S}}{\hbar}\left(2S_j^y - S_{j-1}^y - S_{j+1}^y\right) \tag{6.15}$$

$$\frac{\mathrm{d}S_j^y}{\mathrm{d}t} \approx -\frac{2\sqrt{S}}{\hbar}\left(2S_j^x - S_{j-1}^x - S_{j+1}^x\right) \tag{6.16}$$

$$\frac{\mathrm{d}S_j^z}{\mathrm{d}t} \approx 0. \tag{6.17}$$

We now look for normal mode solutions, so put

$$S_j^x = A\mathrm{e}^{\mathrm{i}(qja-\omega t)} \tag{6.18}$$

$$S_j^y = B\mathrm{e}^{\mathrm{i}(qja-\omega t)} \tag{6.19}$$

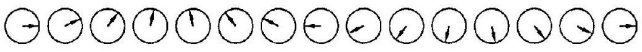
where q is a wave vector. A little algebra leads to $A = \mathrm{i}B$ (showing that the x and y motion are $\pi/2$ out of phase) and hence that

$$\hbar\omega = 4JS(1 - \cos qa), \tag{6.20}$$

which is the dispersion relation for the spin waves and which is plotted in Fig. 6.12. Thus the spin waves are as depicted in Fig. 6.13.

(a)

(b)



(2) It is perhaps more satisfying to derive eqn 6.20 using a quantum mechanical method, and so we now repeat the derivation. [4] We consider the ground state of the system, $|\Phi\rangle$, which consists of all the spins lying along the $+z$ direction. In one dimension the Hamiltonian for the Heisenberg model can be written

$$\hat{\mathcal{H}} = -2\rfloor \sum_i \left[\hat{S}_i^z \hat{S}_{i+1}^z + \frac{1}{2}\left(\hat{S}_i^+ \hat{S}_{i+1}^- + \hat{S}_i^- \hat{S}_{i+1}^+\right)\right], \tag{6.21}$$

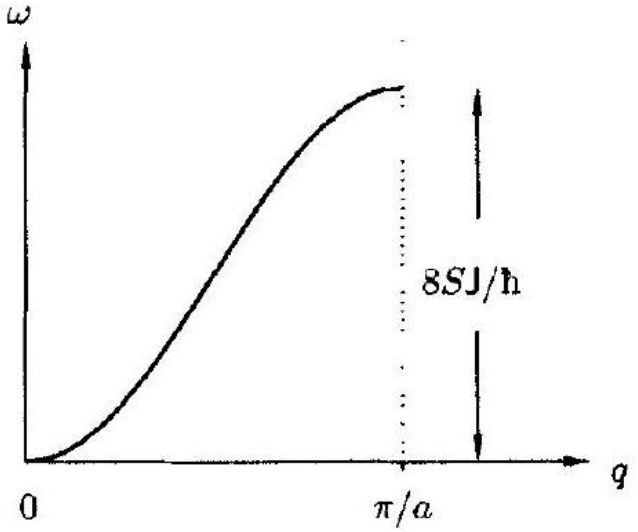


Fig. 6.12 The spin wave dispersion relation for a one-dimensional chain of spins.

Fig. 6.13 A spin wave on a line of spins. (a) perspective view. (b) view from above.
[4] We take $S = \frac{1}{2}$ for now, but it can be easily generalized.

```
$|\Phi\rangle=\cdots \uparrow \uparrow \uparrow \uparrow \uparrow \uparrow \uparrow \upa
    $j$
$|j\rangle=\cdots \uparrow \uparrow \downarrow \uparrow \uparrow \uparrow \uparrow \upar
```

Fig. 6.14 The state $|\Phi\rangle$ consists of all the spins lying along the $+z$ direction. The state $|j\rangle$ is the ground state with the spin at site j flipped.
so that $\hat{\mathcal{H}}|\Phi\rangle = -NS^2\ \mathrm{J}|\Phi\rangle$. Now to create an excitation, flip a spin at site j, so let us consider a state $|j\rangle = \hat{S}_j^-|\Phi\rangle$ which is the ground state with the spin at site j flipped (see Fig. 6.14). By flipping a spin, we have changed the total spin of the system by $\frac{1}{2} - \left(-\frac{1}{2}\right) = 1$. This excitation therefore has integer spin and is a boson. If we apply the Hamiltonian to this new state, we get

$$\hat{\mathcal{H}}|j\rangle = 2\left[\left(-NS^2J + 2SJ\right)|j\rangle - SJ|j+1\rangle - SJ|j-1\rangle\right], \tag{6.22}$$

which is not a constant multiplied by $|j\rangle$, so this state is not an eigenstate of the Hamiltonian. Nevertheless, we can diagonalize the Hamiltonian by looking for plane wave solutions of the form

$$|q\rangle = \frac{1}{\sqrt{N}}\sum_j \mathrm{e}^{\mathrm{i}qR_j}|j\rangle. \tag{6.23}$$

The state $|q\rangle$ is essentially a flipped spin delocalized (smeared out) across all the sites. Since it is composed of states representing a single flipped spin, the total spin of $|q\rangle$ itself has the value $NS - 1$. It is then straightforward to show that

$$\hat{\mathcal{H}}|q\rangle = E(q)|q\rangle, \tag{6.24}$$

where

$$E(q) = -2NS^2\ \mathrm{J} + 4JS(1 - \cos qa). \tag{6.25}$$

The energy of the excitation is then $\hbar\omega = 4JS(1 - \cos qa)$ which is the same as the result we had above in eqn 6.20.

6.6.2 The Bloch $T^{3/2}$ law

At small q, eqn 6.20 yields

$$\hbar\omega \approx 2 \lfloor\ Sq^2a^2, \tag{6.26}$$

so that $\omega \propto q^2$. In three dimensions, the density of states is given by

$$g(q)\mathrm{d}q \propto q^2\ \mathrm{d}q \tag{6.27}$$

which leads to

$$g(\omega)d\omega \propto \omega^{1/2}\ \mathrm{d}\omega \tag{6.28}$$

at low temperature where only small q and small ω are important. The spin waves are quantized in the same way as lattice waves. The latter are termed phonons, and so in the same way the former are termed magnons. As shown in Section 6.6.1, magnons have a spin of one and are bosons.

The number of magnon modes excited at temperature $T, n_{\rm magnon}$, is calculated by integrating the magnon density of states over all frequencies after multiplying by the Bose factor, $(\exp(\hbar\omega/k_{\rm B}T)-1)^{-1}$, which must be included because magnons are bosons. Thus the result is given by

$$n_{\rm magnon} = \int_0^\infty \frac{g(\omega){\rm d}\omega}{\exp(\hbar\omega/k_{\rm B}T)-1} \tag{6.29}$$

which can be evaluated using the substitution $x = \hbar\omega/k_{\rm B}T$. At low temperature, where $g(\omega) \propto \omega^{1/2}$ in three dimensions, this yields the result

$$n_{\rm magnon} = \left(\frac{k_{\rm B}T}{\hbar}\right)^{3/2} \int_0^\infty \frac{x^{1/2}\,{\rm d}x}{{\rm e}^x - 1} \propto T^{3/2} \tag{6.30}$$

Since each magnon mode which is thermally excited reduces the total magnetization [5] by $S = 1$, then at low temperature the reduction in the spontaneous magnetization from the $T = 0$ value is given by

$$\frac{M(0) - M(T)}{M(0)} \propto T^{3/2}. \tag{6.31}$$

This result is known as the Bloch $T^{3/2}$ law and it fits experimental data in the low temperature regime (see Fig. 6.15). [6] The energy of the magnon modes is given by

$$E_{\rm magnon} = \int_0^\infty \frac{\hbar\omega g(\omega){\rm d}\omega}{\exp(\hbar\omega/k_{\rm B}T)-1} \propto T^{5/2} \tag{6.32}$$

so that the heat capacity $C = \partial E_{\rm magnon}/\partial T$ is also proportional to $T^{3/2}$.

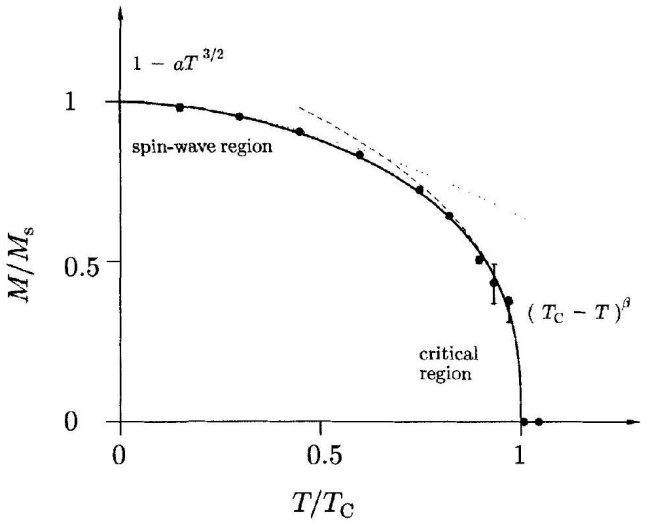


Fig. 6.15 The spontaneous magnetization in a ferromagnet. At low temperatures this can be fitted using the spin-wave model and follows the Bloch $\mathrm{T
ˆ{3 / 2}$ law. Near the critical temperature, the magnetization is proportional to $(T - T_{\rm C})^\beta$ where β is a critical exponent. Neither behaviour fits the real data across the whole temperature range. The data are for an organic ferromagnet which has $T_{\rm C} \approx 0.67$ K for which $\beta \approx 0.36$, appropriate for the three-dimensional Heisenberg model.}

6.6.3 The Mermin-Wagner-Berezinskii theorem

In one and two dimensions, however, the integral in eqn 6.29 diverges, so that $M \to 0$ for all $T > 0$ for the isotropic 1-D and 2-D Heisenberg models. [7] The number of spin waves which are generated at finite temperature diverges, and hence spontaneous ferromagnetism is not possible. This result was first proved by Mermin and Wagner in 1966 and independently by Berezinskii. The absence of long range order in two-dimensional systems (with a continuous symmetry) is often referred to as the Mermin-Wagner-Berezinskii theorem.

[5] Because each magnon mode is a delocalized single reversed spin, see the previous section.

[6] The Bloch $T^{3/2}$ law is only really correct for the spontaneous magnetization within a domain. The data shown in Fig. 6.15 actually measure this because they were obtained using μ SR which measures the local internal field with zero applied field.

[8] If the wavelength of the spin wave is sufficiently long, the exchange energy cost $\alpha \int \left(|\nabla M_x|^2 + |\nabla M_y|^2\right) \mathrm{d}x\ \mathrm{d}y$ (in two dimensions) is minimized. See Exercise 6.7.

[9] This idea is explored in Exercise 6.7

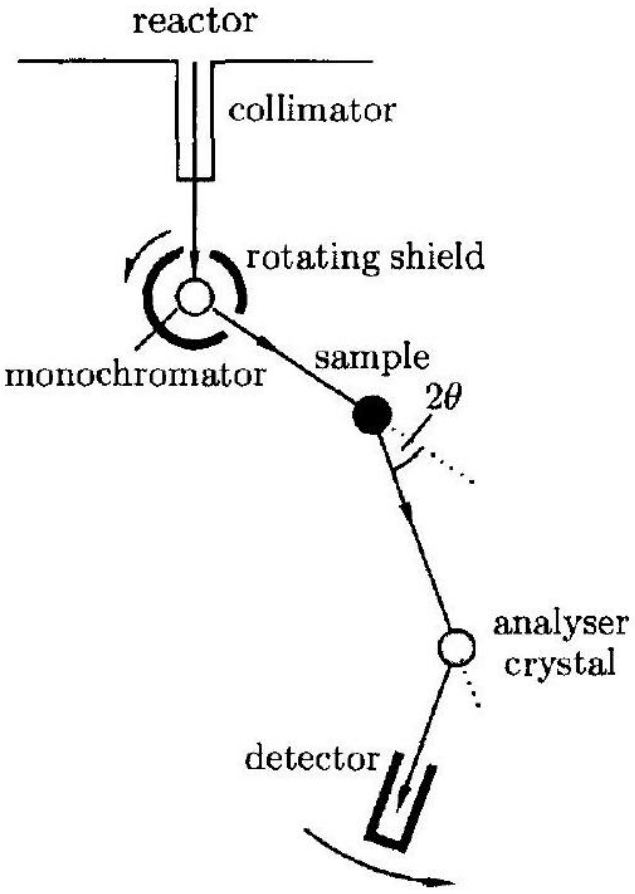


Fig. 6.16 Schematic diagram of a neutron triple-axis spectrometer. By setting the angle of the rotating shield and the monochromator crystal, the incident neutron energy can be selected. The sample can also be rotated. The analysing crystal can also be rotated, allowing the scattered neutron energy to be measured.

This result only applies to an isotropic Heisenberg ferromagnet. This possesses rotational

[07] For different dimensions, the function $g(\omega)$ in eqn 6.29 has a different dependence on ω.

[010] The addition of **G** is necessary because the dispersion relation of magnons, like that of phonons, is periodic in the reciprocal lattice.

symmetry so that all of the spin directions can be globally rotated without any additional energy cost. This means that long wavelength excitations, in which the spin state may deviate from its ground state value over a considerable distance, cost very little energy. [8] Thus a fluctuation of the spins can be excited with very little energy cost; in one and two dimensions they destroy the long range order. If, however, there is significant anisotropy, there will be an energy cost associated with rotating the spins from their ground state value.

It turns out that the anisotropy energy penalty incurred by allowing these fluctuations increases with the square of R, the radius of the excitation, and hence the anisotropy energy will suppress all but the smallest of these nonlinear fluctuations. [9] It is the presence of such symmetry-breaking fields which can stabilize long range order in two-dimensional systems. There is also a dipolar interaction between spins in real systems which, although much weaker than the exchange interaction, is anisotropic and can act in a similar way to suppress the growth of fluctuations.

It is observed in experiment that some ultra-thin magnetic films can show spontaneous ferromagnetism. Sometimes the anisotropy is such that it is energetically favourable for the spins to be perpendicular to the plane of the film. (The system is said to have a perpendicular easy axis for magnetization.) In this case the anisotropy leads to a gap in the spin wave excitation spectrum (between the ground state and the excited states) and long range magnetic order can exist at finite temperature. Dipolar interactions, or anisotropy of the energy of the spins in the plane of the film, can lead to a similar effect and stabilize the magnetization.

6.6.4 Measurement of spin waves

Elastic neutron scattering, which is used for magnetic structure determination, was discussed in Section 5.7.2. Spin wave dispersions can be measured using a technique known as inelastic neutron scattering. The magnitude of the incident neutron wave vector $\mathbf{k}$ is now no longer equal to the magnitude of the scattered neutron wave vector $\mathbf{k}'$. The energy of the neutron also changes from

$$E = \frac{\hbar^2 k^2}{2m_{\mathrm{n}}} \tag{6.33}$$

$$E' = \frac{\hbar^2 k'^2}{2m_{\mathrm{n}}} \tag{6.34}$$

This is because the neutron has produced in the sample an excitation of energy $\hbar\omega$ and wave vector $\mathbf{q}$. Conservation of energy and momentum implies that

$$E = E' + \hbar\omega \tag{6.35}$$

$$\mathbf{k} = \mathbf{k}' + \mathbf{q} + \mathbf{G} \tag{6.36}$$

where $\mathbf{G}$ is a reciprocal lattice wave vector, [10] so that a measurement of $\mathbf{k}, \mathbf{k}'$, E and E' allows a determination of ω and $\mathbf{q}$. Neutrons have energies similar to the energies of atomic and electronic processes, i.e. in the meV to eV range.

Thus we can probe energy scales from the μeV of quantum tunnelling, through molecular translations, rotations, vibrations and lattice modes, to eV transitions within the electronic structure of materials; magnon energies are typically in the range $10^{-3} - 10^{-2}$eV and therefore can be effectively measured using inelastic neutron scattering.

Experiments are typically performed using a neutron triple-axis spectrometer (see Fig. 6.16), so called because the angles of the monochromator, sample and analyser crystals can all be separately varied. This allows scattered neutrons corresponding to a large range of possible

values of ω and $\mathbf{q}$ to be measured. In particular, it allows scans to be performed in which q varies with fixed ω (or vice versa). Example data obtained by this technique are shown in Figures 6.17 for $Co_{0.92}Fe_{0.08}$ and 6.18 for the ferromagnetic oxide $La_{0.7}$ $Pb_{0.3}MnO_3$ (which exhibits colossal magnetoresistance, see Section 8.9.5).

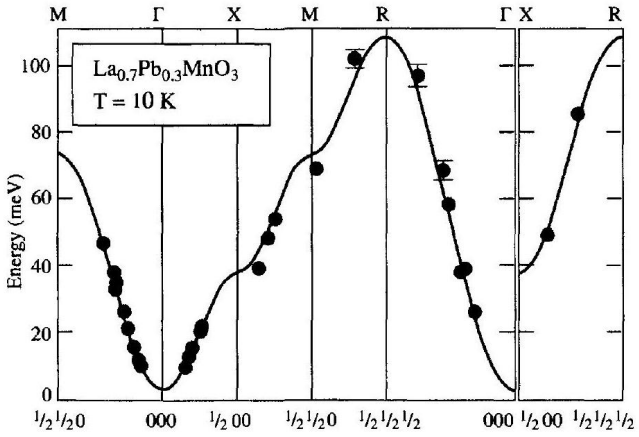


Fig. 6.18 Spin wave dispersion in the ferromagnetic oxide $\mathrm{La *{0.7} *mathrm{˜Pb}*{0.3} \mathrm{MnO}_{3}$ measured at 10 K by inelastic neutron scattering. The horizontal axis corresponds to the magnon wave vector. After Perring et al. 1996.}

9.6.7 6.7 Domains

If different regions of a macroscopic system break symmetry in different ways, then in the interface between these regions the rigidity can break down. In general we expect domain walls, defects, vortices, dislocations and other singularities. In ferromagnets the most important singularity is the domain wall.

Weiss first proposed that a ferromagnet contains a number of small regions called domains, within each of which the local magnetization reaches the saturation value. The direction of the magnetization of different domains need not be parallel. Domains are separated by domain walls. The existence of domains explains the surprising observation that in some ferromagnetic specimens [11] it can be possible, at room temperature, to attain saturation magnetization of the whole sample (corresponding to $\mu_0 M \sim 1$ T) by the application of a very weak magnetic field (as low as 10^{-6} T). Such low applied fields would

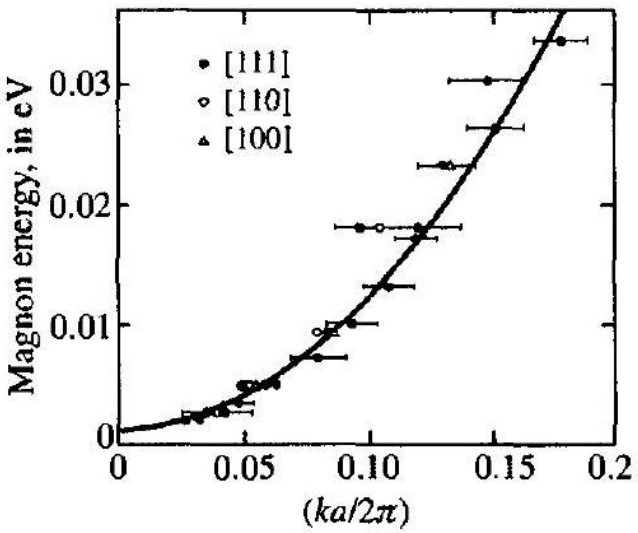


Fig. 6.17 The spectrum of magnon energy for spin waves in an alloy of $\mathrm{Co *{0.92}* *mathrm{Fe}*{0.08}$ obtained at room temperature (Sinclair and Brockhouse 1960). In this figure, k is the magnon wave vector corresponding to q in the text.}

[11] Which are magnetically 'soft', see Section 6.7.9.

[12] If $x \sim 10^{-3}$ for a particular paramagnet. see Table 2.1, then an applied field of 10^{-6} T would produce a magnetization of $\mu_0 M \sim 10^{-9}$·.

13 As we shall see, in some ferromagnetic materials the movement of domain walls can in fact be energenically costly. Such substances would require large applied magnetic fields to switch them from one magnetic state to the other.

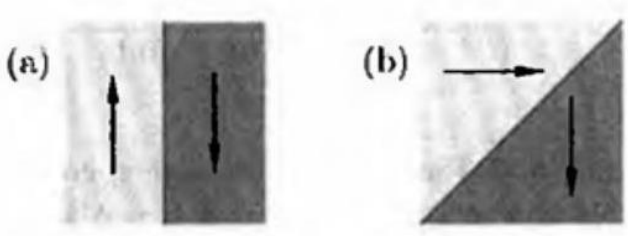


Fig. 6.19 (a) 180° domain wall. (b) 90° domain wall.}

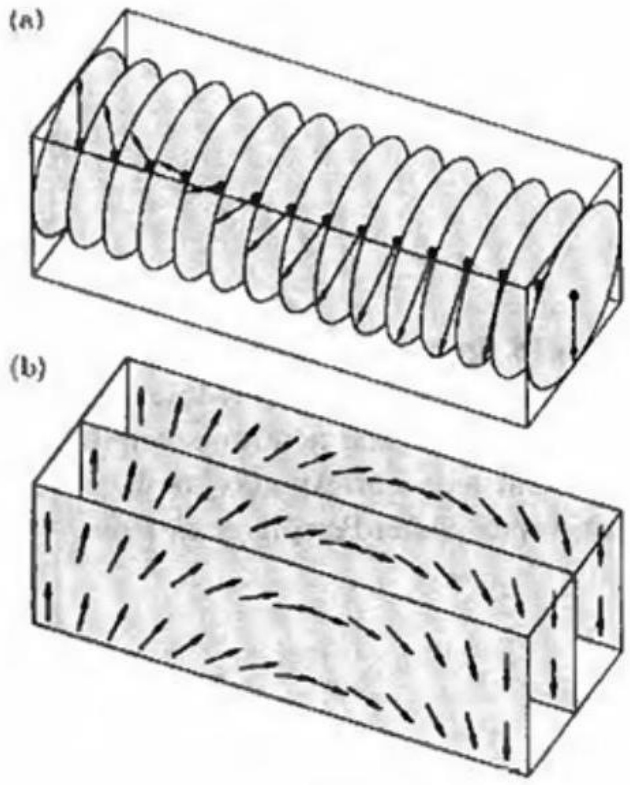


Fig. 6.20 (a) A Bloch wall. (b) A Néel wall.

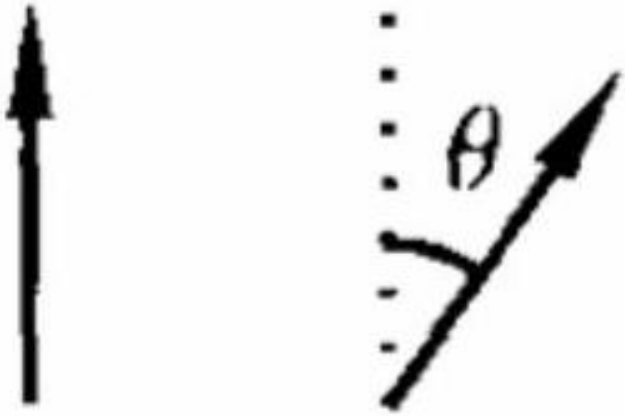


Fig. 6.21 Two spins at an angle A with respect to cach other. producing an energy cost $\frac{1 {2} J S^{2} A^{2}$.}

have negligible effect on a paramagnet. [12] The large effect in the ferromagnetic specimen is because the applied field does not have to order the magnetic moments macroscopically (they are all already ordered!) but merely to cause the domains to align. This can be achieved by the energetically painless process of domain wall motion. [13]

In the same ferromagnetic specimen, it is also possible for the magnetization to be zero in zero applied field. This is also a manifestation of domains; the magnetization is saturated within each domain, but the directions of the magnetization of each domain are such that the net magnetization of the specimen is zero.

[014] L sing $\cos\theta \approx \mathrm{I} - \mathrm{H}^2/2$ for $\theta \ll \mathrm{I}$.

6.7.1 Domain walls

Between adjacent domins there is a boundary called a domain wall. The domain walls can be classified according to the angle between the magnetization in the two domains (see Fig. 6.19). A 180° domain wall separates domains of opposite magnetization. A 90° domain wall separates domains of perpendicular magnetization.

The most common type of 180° wall is a Bloch wall (see Fig. 6.20(a)) in which the magnetization rotates in a plane parallel to the plane of the wall. Another possible configuration is the Néel wall (see Fig. 6.20(b)) in which the magnetization rotates in a plane perpendicular to the plane of the wall. Let us attempt to calculate the domain wall width in a Bloch wall.

In a ferromagnet it costs energy to rotate neighbouring spins. Two spins, S_1 and $\mathbf{S}_2$, which are at angle of θ with respect to each other (Fig. 6.21) have an energy $-2\sqrt{\mathbf{S}_1 \cdot \mathbf{S}_2} = -2\sqrt{S^2 \cos\theta}$. If $\theta = 0$, their energy is $-2\sqrt{S^2}$. Hence the energy cost of having $\theta \neq 0$ is approximately $^{14}\mathrm{JS}^2\theta^2$ if $\theta \ll 1$. In a Bloch wall, spins rotate over N sites by an angle π (see Fig. 6.20(a)). Hence the energy cost of a line of spins is equal to N contributions of $JS^2\theta^2$ where $\theta = \pi/N$. i.e. to $\mathrm{J}S^2\pi^2/N$. In a Bloch wall we have planes of spins (see Fig. 6.20) and so we are interested in σ_{BW}, the energy per unit area of the Bloch wall. In a square metre of wall. there are $1/u^2$ lines of spins like the one we have calculated. Hence

$$\sigma_{\mathrm{BW}} = \mathrm{J}S^2 \frac{\pi^2}{Na^2} \tag{6.37}$$

which tends to zero as $N \to \infty$. This result would seem to indicate that if a domain wall formed it would just unwind itself, growing in size throughout the entire system. This is because it costs energy to have spins twisted with respect to each other, and therefore they will all untwist unless some other interaction stops them. This other interaction is the magnetocrystalline anisotropy.

6.7.2 Magnetocrystalline anisotropy

Crystals possess a magnetic easy axis and a hard oxis. Along certain crystallographic directions it is easy to magnetize the crystal. along others it is harder (this is shown for single crystals of Fe, Co and Ni in Fig. 6.22). In Co . for example, this anisotropy leads to an additional energy of the form

$$E = K_1 \sin^2\theta + K_2 \sin^4\theta \tag{6.38}$$

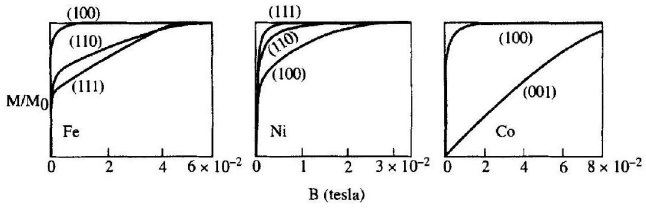


Fig. 6.22 Magnetization in Fe, Co and Ni for applied fields in different directions showing anisotropy. After Honda and Kaya 1926, Kaya 1928.}

where K_1 and K_2 are anisotropy constants and θ is the angle between the magnetization and the stacking direction of the hexagonally close packed planes. Because these constants are positive the energy is minimized when the magnetization lies along the stacking direction. In

eqn 6.38, E, K_1 and K_2 are energy densities (i.e. they are measured in Jm^{-3}). The anisotropy constants are found to be strongly temperature dependent.

Equation 6.38 is appropriate for uniaxial anisotropy, in which the energy depends on the angle to a single axis (in Co this is the stacking axis of the hexagonally close packed planes). In a cubic system, the appropriate expression is

$$E = K_1 \left(m_x^2 m_y^2 + m_y^2 m_z^2 + m_z^2 m_x^2\right) + K_2 m_x^2 m_y^2 m_z^2 + \cdots , \tag{6.39}$$

where $\mathbf{m} = (m_x, m_y, m_z) = \mathbf{M}/|\mathbf{M}|$. In spherical coordinates this is

$$E = K_1 \left(\frac{1}{4}\sin^2\theta \sin^2 2\phi + \cos^2\theta\right) \sin^2\theta + \frac{K_2}{16}\sin^2 2\phi \sin^2 2\theta \sin^2\theta + \cdots . \tag{6.40}$$

The anisotropy energy arises from the spin-orbit interaction and the partial quenching of the angular momentum. Anisotropy energies are usually in the range $10^2 - 10^7 \mathrm{Jm}^{-3}$. This corresponds to an energy per atom in the range $10^{-8} - 10^{-3}\mathrm{eV}$. The anisotropy energy is larger in lattices (of magnetic ions) of low symmetry and smaller in lattices of high symmetry. For example, cubic Fe and Ni have K_1 equal to $4.8 \times 10^4 \mathrm{Jm}^{-3}$ and $-5.7 \times 10^3 \mathrm{Jm}^{-3}$ respectively, but hexagonal Co has $K_1 = 5 \times 10^5 \mathrm{Jm}^{-3}$. Low symmetry permanent magnet materials Nd_2Fe_{14} B and $SmCo_5$ have K_1 equal to $5 \times 10^6 \mathrm{Jm}^{-3}$ and $1.7 \times 10^7 \mathrm{Jm}^{-3}$ respectively.
An additional energy term is due to the demagnetizing energy associated with the sample shape and is referred to as shape anisotropy. In thin films, a shape anisotropy $\frac{1}{2}\mu_0 M^2 \cos^2\theta$ (where θ is the angle between the film normal and $\mathbf{M}$) leads to an energetic saving for keeping the magnetization in the plane of the film.

6.7.3 Domain wall width

In the magnetic domains of a ferromagnet the magnetization will prefer to lie along the easy direction but between domains, in the domain wall, it will have to rotate and a component will lie along the hard axis which will cost energy. If we assume a simple form for the anisotropy energy density, namely $E = K\sin^2\theta$ where K is an anisotropy constant, then we can easily find an
expression for the anisotropy energy contribution to a Bloch wall. We take $K > 0$ so that spins prefer to line up along $\theta = 0$ or $\theta = \pi$. We add up the contribution from each of the N spins and replace the sum by an integral for simplicity (the continuum limit). Thus the energy density contribution is

$$\sum_{i=1}^{N} K\sin^2\theta_i \approx \frac{N}{\pi}\int_0^{\pi} K\sin^2\theta \, \mathrm{d}\theta = \frac{NK}{2}. \tag{6.41}$$

The energy contribution per unit area of wall can then be written as $NKa/2$. Thus the total energy per unit area of the domain wall, including the contribution from the exchange energy (eqn 6.37) and the contribution from the anisotropy energy (eqn 6.41) is

$$\sigma_{\mathrm{BW}} = JS^2\frac{\pi^2}{Na^2} + \frac{NKa}{2}. \tag{6.42}$$

This gives us the required behaviour since we have two terms, one proportional to $1/N$ (tending to unwind the wall and make it bigger) and the other proportional to N (tending to tighten it and make it smaller). We find the equilibrium configuration using $\mathrm{d}E_{\mathrm{BW}}/\mathrm{d}N = 0$ which leads to a value of N given by

$$N = \pi S\sqrt{2J/Ka^3} \tag{6.43}$$

so that the width of the Bloch wall is

$$\delta = Na = \pi S\sqrt{\frac{2\rfloor}{Ka}}, \tag{6.44}$$

where a is the lattice spacing. Larger J makes the wall thicker, larger K makes it smaller. The energy per unit area of the domain wall is

$$\sigma_{\mathrm{BW}} = \pi S\sqrt{\frac{2\sqrt{K}}{a}}. \tag{6.45}$$

In terms of the parameter A defined in eqn 4.19 (for a cubic crystal) we have the domain wall thickness

$$\delta = \pi\sqrt{\frac{A}{K}}, \tag{6.46}$$

and the energy per unit area as

$$\sigma_{\mathrm{BW}} = \pi\sqrt{AK} \tag{6.47}$$

6.7.4 Domain formation

Since it costs this energy to make a domain wall, one might wonder why more than one domain ever forms in the first place. The reason is that the formation of domains saves energy associated with dipolar fields. Because $\nabla \cdot \mathbf{H} = -\nabla \cdot \mathbf{M}$ (see Appendix B) then whenever $\mathbf{M}$ stops and starts, at the edges of a sample for example, the magnetic field diverges and this produces demagnetizing fields which fill space and which cost $B^2/2\mu_0$ Joules of energy per cubic metre. The energy associated with the demagnetizing field is called, variously, the
demagnetization energy, magnetostatic energy or dipolar energy. It takes the value

$$-\frac{\mu_0}{2}\int_V \mathbf{M}\cdot\mathbf{H}_{\mathrm{d}}\ \mathrm{d}\boldsymbol{\tau} \tag{\{6.48\}}$$

where $\mathrm{H_d}$ is the demagnetizing field and the integral is taken over the volume of the sample (see Appendix D). For an ellipsoidally shaped sample magnetized along one of its principal axes, this energy reduces to

$$\frac{\mu_0}{2}NM^2v \tag{6.49}$$

where N is the demagnetizing factor and V is the sample volume.
This dipolar energy can be saved by breaking the sample into domains, but each domain created costs energy because of the cost of the domain walls. Therefore. in the same way as the size of a domain wall is a balance between the exchange and anisotropy energies, so the formation of domains is a balance between the cost of a demagnetizing ficld and the cost of a domain wall.

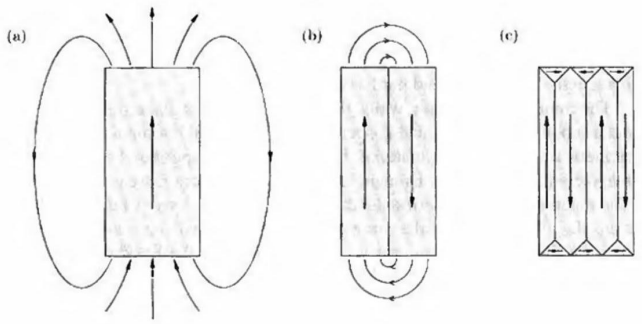


Fig. 6.23 A sample which is (a) uniformly magnetized. (b) divided into two domains. and (c) with a simple closure domain structure.

Figure 6.23 shows three different choices for the domain structure for a ferromagnetic sample. The single domain structure in Fig. 6.23(a) has no domain walls but a large dipolar energy. The dipolar energy can be reduced by breaking the sample into two domains as shown in Fig. 6.23(b), albeit at the cost of introducing a domain wall. The so-catled closure domain structure in Fig. 6.23(c) eliminates the dipolar energy but introduces a number of domain walls.

Dipolar energy can also determine the type of domain wall that can form. The Bloch wall is favoured in the bulk because it leads to a smaller dipolar energy. The Néel wall tends to be favoured in thin films (where there is a dipolar energy cost to rotating the spins out of the plane of the film).

6.7.5 Magnetization processes

Figure 6.24 shows a typical hysteresis loop expected for measuring the magnetization as a function of the applied magnetic field for a ferromagnetic sample. If the sample is magnetized to the saturation magnctization M_{s} by an applied field, then when the applied field is reduced to zero the magnetization

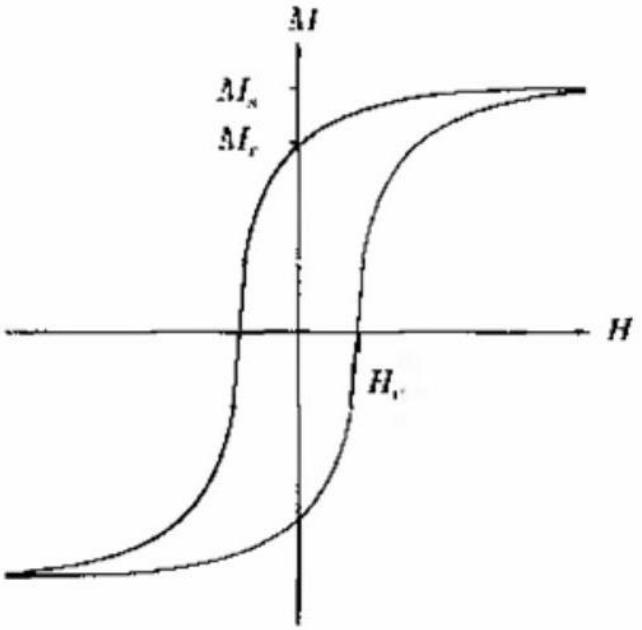


Fig. 6.24 A hysteresis loop showing the saturation magnetization $M_{s} $, the remanent magnetization M_r and the coercive field H_c.}

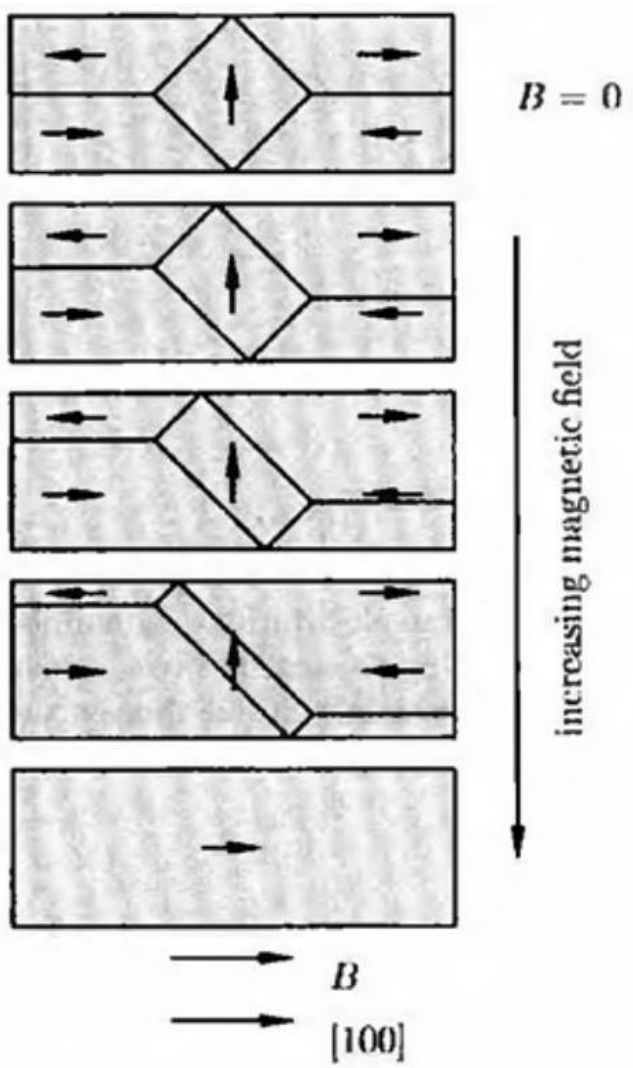


Fig. 6.25 Effect of an applied field on the domain pattern on the surface of a single crystal iron whisker showing domain wall displacement, as the applied B field increases from 0 up is a maximum value.

Heinrich Barkhausen (1881-1956)

John Kerr (1824-1907)

reduces to the remanent magnetization M_{r}. A magnetic field equal to the coercive field H_c is needed to switch the magnetization into the opposite direction. The parameters M_{r} and H_{c} can be used to characterize a ferronagnet.

The length of a ferromagnet changes slightly when it is magnetized. This effect is known as magnetostriction and demonstrates that the magnetic and elastic behaviour of a material are connected (the connection is called magnetoelastic coupling). The crystal deforms because it can lower its anisotropy energy which can depend on the strain of the crystal. For example, in a cubic system it may be energetically favourable for the crystal to deform slightly from exactly cubic symmetry if doing so will save more anisotropy energy than it will cost in elastic energy.

When a demagnetized ferromagnet is magnetized, various processes occur. First is domain-wall motion: domains which are aligned favourably with respect to the applied magnetic field grow at the expense of domains which are unfavourably aligned. (This process is shown in Fig. 6.25.) At higher fields, domain rotation can occur in which the anisotropy energy can be outweighed and the magnetization can suddenly rotate away from the original direction of magnetization to the crystallographic easy axis which is nearest to the field direction. The final

domain process at highest magnetic fields is coherent rotation of the domains to a direction aligned with the magnetic field, irrespective of the easy and hard axes.

The motion of domain walls through a magnetic material depends in detail upon the metallurgical properties of the material. Domain walls can be pinned by strains in the material, by surfaces and impurities because of the magnetoelastic coupling. Domain wall pinning therefore increases coercivity. The magnetization of a ferromagnet also changes by a series of discontinuous steps due to domain boundary motion, so that very small steps are sometimes seen on the magnetization curves. This is known as the Barkhousen effect. Low level acoustic emission also sometimes accompanies the magnetization process because of these sudden discontinous changes in magnetization and their coupling to the elastic modes of the system. This effect is known as magnetococoustic emission.

6.7.6 Domain wall observation

Domains can be observed using a variety of techniques. In the Bitter powder technique a colloidal suspension of fine magnetic particles is placed on the surface of the sample. The magnetic particles collect near domain walls where there are strong local fields which attract them. These can be observed using optical microscopy. It is also possible to use polarized light reflecting from a polished magnetized sample to observe domains, a technique known as Kerr microscopy. The incident beam of light is polarized and its plane of polarization is rotated by reflection from the magnetized sample (this is the Kerr effect, see Section 8.8). The degree of rotation can be measured using a second polarizer (the analyser). If a beam of light is focussed to a spot and scanned across the surface of a magnetic sample, then for a particular setting of the analyser polarizer, some domains will appear bright and others dark. If the sample is transparent, the experiment can be performed in transmission (the Faraday effect).

Electron microscopy can also be effective as the electron beam is deflected by the Lorentz force due to the internal field. Boundaries between domains (across which this deflection can suddenly change in sign as the beam is scanned over) show up particularly well. This technique is called Lorentz microscopy.
In the technique of magnetic force microscopy [15] a tiny cantilever with a magnetized tip is scanned across a sample surface. The stray fields produced by a domain wall cause the cantilever to bend, and this bending (detected by, for example, a change in the resonant frequency of the cantilever) is used to image the surface.

6.7.7 Small magnetic particles

In many ferromagnetic samples the lowest energy state at zero applied field is the demagnetized state, so that the overall magnetic moment is zero. If the sample size is reduced, surface energies (such as the domain wall energies) become progressively more costly in comparison with volume energies (such as the demagnetizing energy). [16] Thus a critical dimension may be reached below which it is energetically favourable to remove the domain walls so that the sample consists of a single magnetized domain. It would thus behave like a small permanent magnet.

In the following example, we estimate the critical size for a single domain particle. The calculation will assume that this critical size is still larger than the domain wall width.

Example 6.1

Consider a ferromagnetic particle of radius r. The energy of the single domain state (Fig. 6.26(a)) is, using eqn D. 12 and $N = \frac{1}{3}$ for a sphere, given by

$$E_{(a)} = \frac{1}{6}\mu_0 M^2 V = \frac{2}{9}\mu_0 \pi M^2 r^3. \tag{6.50}$$

The state in Fig. 6.26(b), appropriate for cubic anisotropy, approximately removes all demagnetizing energy, but introduces 90° domain walls of area $2 \times \pi r^2$, leading to an energy cost

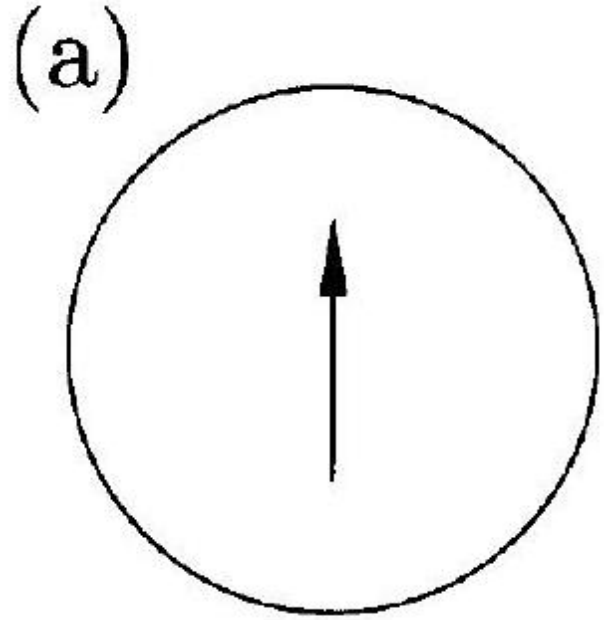
$$E_{(b)} = 2\pi r^2 \sigma_w^{90^\circ}, \tag{6.51}$$

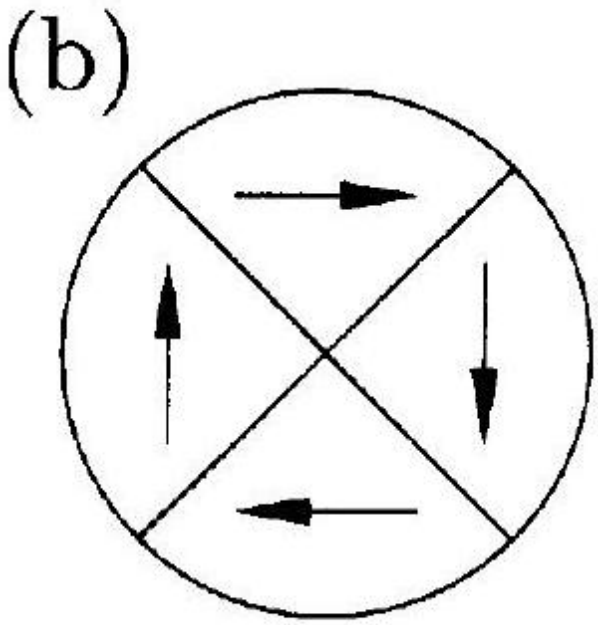
where $\sigma_W^{90^\circ}$ is the energy cost of a 90° domain wall per unit area. The state in Fig. 6.26(a) thus becomes more favourable than the state in Fig. 6.26(b) when $E_{(a)} < E_{(b)}$, i.e. when

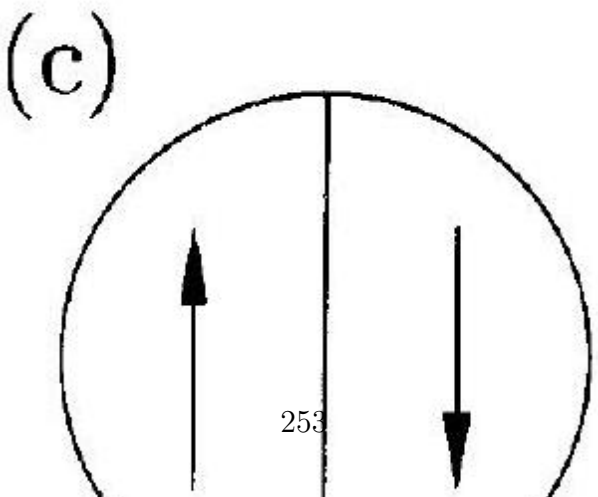
$$r < \frac{9\sigma_W^{90^\circ}}{\mu_0 M^2}. \tag{6.52}$$

[15] A magnetic force microscope is based on the well-known atomic force microscope.
[16] This is because surface energies scale as the (sample size) [2] whereas volume energies scale as the (sample size) [3]

The volume of a sphere is $V = \frac{4}{3}\pi r^3$.


(a)


(b)


(c)

The state in Fig. 6.26(c), appropriate for uniaxial anisotropy, is harder to consider because neither domain is ellipsoidal in shape and there is still some demagnetizing energy. Roughly, though, the demagnetizing energy of this configuration is half that of Fig. 6.26(a) since the domains are half the size. The domain wall cost is then $\pi r^2 \sigma_{\mathrm{w}}^{180^\circ}$ where $\sigma_{\mathrm{w}}^{180^\circ}$ is the energy of a 180° domain wall per unit area. Thus

$$E_{(\mathrm{c})} = \frac{1}{9}\mu_0 \pi M^2 r^3 + \pi r^2 \sigma_{\mathrm{w}}^{180^\circ} \tag{6.53}$$

Hence the state in Fig. 6.26(a) becomes more favourable than the state in Fig. 6.26(c) when $E_{(\mathrm{a})} < E_{(\mathrm{c})}$, i.e. when

$$r < \frac{9\sigma_{\mathrm{w}}^{180^\circ}}{\mu_0 M^2}. \tag{6.54}$$

If $\sigma_{\mathrm{w}}^{180^\circ} \sim 10^{-2}\mathrm{Jm}^{-2}$ and $\mu_0 M \sim 1$ T, the critical radius is $\sim 10^{-7}$ m. Note, if this value came out to be smaller than the domain wall width, then the calculation would be nonsensical.

6.7.8 The Stoner-Wohlfarth model

It is possible to calculate the magnetization curve for a single-domain particle, via the Stoner-Wohlfarth model. Because the particle is a single magnetic domain, there are no domain walls and one need only consider coherent domain rotation.

We consider a single-domain magnetic particle in a magnetic field H which is applied at an angle θ to the easy axis of its uniaxial anisotropy. If the magnetization of the particle then lies at an angle of ϕ to the magnetic field direction, the energy density of the system is

$$E = K \sin^2(\theta - \phi) - \mu_0 H M_{\mathrm{s}} \cos\phi. \tag{6.55}$$

The energy can be minimized to find the direction of the magnetization at any given value of the applied magnetic field.

This is shown in Fig. 6.27 where the applied field is written in dimensionless units using the parameter h given by

$$h = \frac{\mu_0 M_{\mathrm{s}} H}{2K} \tag{6.56}$$

for two choices of the direction of the applied magnetic field. Analytic solutions are possible for this model for $\theta = 0$ and $\theta = \pi/2$, but in the figure it is shown how they can be solved numerically. In each case the trajectory of the magnetization direction which lowers the energy is shown by a curve across an energy surface (energy as a function of ϕ and h) and the corresponding hysteresis loops are also shown. The hysteresis loops are shown for a larger choice of values of θ in Fig. 6.28, together with the calculated result for a polycrystalline average of directions, as would be appropriate for an array of single-domain particles (ignoring any interactions between the particles). This model demonstrates how the anisotropies present in a system can lead to hysteresis, even in a system in which there are no irreversible effects associated with domain-wall pinning.

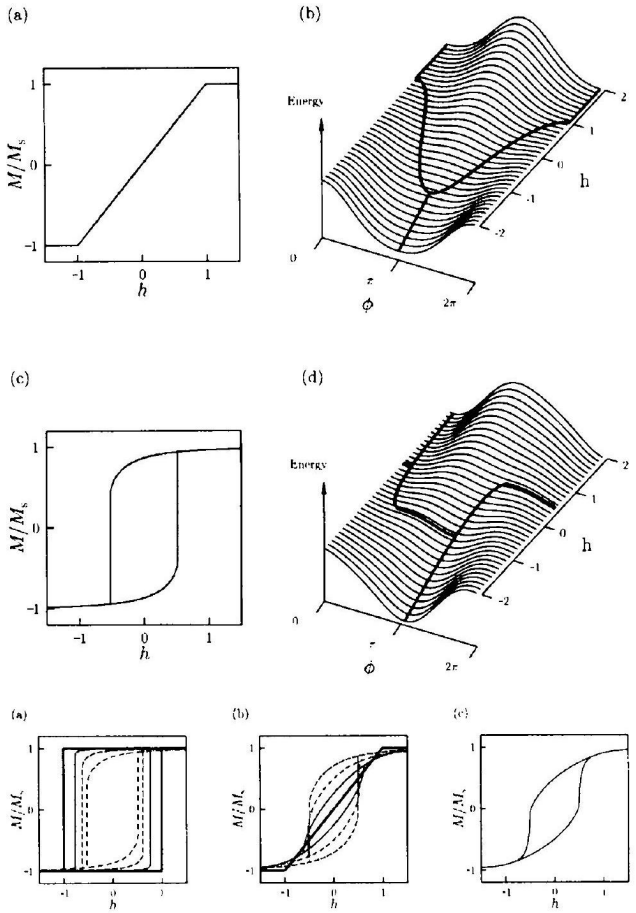


Fig. 6.27 The hysteresis loops in the StonerWohlfarth model for (a) $\theta=90^{\circ}$ and (c) $\theta = 30°$. These can be obtained by finding the minimum energy points on the energy surfaces as a function of h and ϕ, which are shown for (b) $\theta = 90°$ and (d) $\theta = 30°$.}

6.7.9 Soft and hard materials

The energy dissipated (as heat) by a ferromagnet as it is taken around a circuit of its hysteresis loop is proportional to the area of that loop. If the area is small, the material is said to be magnetically soft. If the area is large, the material is said to be magnetically hard. Domain walls traverse a sample as the field is cycled and we can distinguish these two broad classes of ferromagnetic materials based on how easy it is for domains to move through a sample.
(1) Soft magnetic materials are easy to magnetize. Soft magnets are used in transformer coils, generators and motors. In these applications the magnetization must be reversed many times a second and it is important that the energy dissipated in each cycle is minimized. Soft materials have broad domain walls (small anisotropy energy K) which are thus easy

Fig. 6.28 The hysteresis loops in the StonerWohlfarth model for (a) $\theta = 0°$ (bold). 5°. 15°.30°, (b) $\theta = 45°.60°.75°$ and 90° (bold). (c) A calculated hysteresis loop for a polycrystalline average.

to move. This leads to small coercive fields. Low magnetostriction is often desirable so that internal strains will not induce a local anisotropy energy. An example is permalloy (a commercial Ni/Fe alloy with an additional ingredient) which has a coercive field $B_c \sim 2 \times 10^{-7}$ T.
(2) Hard magnetic materials are difficult to magnetize and thus difficult to demagnetize. Hard magnets are used as permanent magnets (e.g. in the back of loudspeakers, in motors, and of course on the front of your refrigerator!) and in magnetic tape recording (in powder form). In these applications, the magnetization needs to be preserved for as long as possible. The energy

dissipated in a hysteresis loop cycle is as large as possible so that the magnetization will not occur spontaneously. Hard magnets have large hysteresis and narrow domain walls (large K) so that it is easy to have domain wall pinning. Large ion moments and large crystal fields are helpful for hard magnetic properties and suitable materials often involve rare earths, e.g. Nd_2Fe_{14} B which has $T_C = 585$ K and coercive field $B_c = 1.2$ T.

An important application of hard magnetic materials is magnetic recording. The areal density of bits that can be recorded using magnetic materials has shot up dramatically over the last few years, reflecting a steady improvement in the understanding and fabrication of magnetic materials. The approximate figures are ~ 0.1 Mbits in $^{-2}$ in 1960, ~ 10 Mbits in $^{-2}$ in 1980 and ~ 25 Gbits in $^{-2}$ in 2000 (1 bit in $^{-2}$ is a bit of information, a ' 0 ' or ' 1 ', stored in a squareinch - this most modern of industries has yet to modernize its units!) As the areal density increases in the future, superparamagnetic fluctuations (i.e. thermal fluctuations of the magnetic particles, see Section 8.3) of the magnetic moments will become important.

Small magnetic particles of Fe_2O_3 were traditionally used on recording tapes and disks. Improvements in properties for high-fidelity audio recording and video tapes are based on high coercivity particles such as Fe_2O_3 doped with Co, CrO_2, metal particles (usually Fe), and barium ferrite ($BaO \cdot 6Fe_2O_3$, see Section 5.3). Hard drives typically consist of a large rigid disk with a thin film deposited by vacuum deposition or sputtering, usually of a Co/Cr alloy. The hard disks rotate rapidly and a reading-head moves across the surface of the disk (it can hover a fraction of a micron above the surface) and is able to access any part of it.

9.6.8 Further reading

- The concept of broken symmetry is described in more detail in P. W. Anderson, Basic notions of condensed matter physics, Addison-Wesley 1984.
- The statistical mechanics of phase transitions is discussed in J. Yeomans, Statistical mechanics of phase transitions, OUP 1992.
- Also very helpful on critical phenomena is M. F. Collins, Magnetic critical scattering. OUP 1989.
- A very thorough description of phase transitions and broken symmetry phenomena, with particular emphasis on soft condensed matter, is P. M. Chaikin and T. C. Lubensky, Principles of condensed matter physics, CUP 1995.
- J. J. Binney, N. J. Dowrick, A. J. Fisher and M. E. J. Newman, The theory of critical phenomena, OUP 1992 contains an excellent account of renormalization group methods.
- Information on domains and magnetism may be found in G. Bertoni, Hysteresis in magnetism, Academic Press 1998 and A. Hubert and R. Schfer, Magnetic domains, Springer 1998.

9.6.9 References

- K. Honda and S. Kaya, Sci. Rep. Tohoku Univ., 15, 721 (1926)
- S. Kaya, Sci. Rep. Tohoku Univ., 17, 639, 1157 (1928).
- Useful general references are D. Craik, Magnetism: principles and applications, Wiley 1995 and J. Crangle, Solid state magnetism, Edward Arnold 1991.

- T. G. Perring, G. Aeppli, S. M. Hayden, S. A. Carter, J. P. Remeika and S-W. Cheong, Phys. Rev. Lett., 77, 711 (1996).
- R. N. Sinclair and B. N. Brockhouse, Phys. Rev., 120, 1638 (1960).

Exercises

(6.1) A one-dimensional ferromagnetic chain of N spins is described by the Ising Hamiltonian

$$\hat{\mathcal{H}} = -2J \sum_{i=1}^{N-1} \hat{S}_i^z \hat{S}_{i+1}^z - 2J \hat{S}_N^z \hat{S}_1^z$$

(The last term is used to give periodic boundary conditions.)
Show that the partition function Z of this system can be obtained by introducing new operators

$$\zeta_i = 4 S_i^z S_{i+1}^z$$

which have eigenvalues +1 or -1 , and hence show that the partition function is $Z = (2\cosh(\ \mathrm{J}/2k_{\mathrm{B}}T))^N$.
Using these results obtain an expression for the heat capacity per spin of the chain as $N \to \infty$. Deduce the low and high temperature behaviour of the heat capacity and sketch the heat capacity as a function of temperature. Discuss the result.
(6.2) A uniaxial ferromagnet is described by the Hamiltonian

$$\hat{\mathcal{H}} = -\sum_{ij} \mathbf{J}_{ij} \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j - \sum_{ij} K_{ij} \hat{S}_i^z \hat{S}_j^z \tag{6.57}$$

(a) Show that the state with all spins fully aligned along the z axis is an eigenstate of the Hamiltonian.
(b) Obtain an expression for the spin wave spectrum as a function of wave vector q.
(c) Simplify the expressions for the case where J_{ij} and K_{ij} are restricted to nearest neighbours, J_0 and K_0, and the ferromagnet is (i) a one-dimensional chain, (ii) a twodimensional square lattice and (iii) a three-dimensional body-centred cubic material.
(6.3) Using the results of Exercise 6.2 for small wave vectors, deduce the temperature dependence at low temperatures
of the number of spin waves for the Heisenberg model ($K_0 = 0$) and the Ising model ($\mathrm{J}_0 = 0$) for each of the structures (i), (ii) and (iii) of Exercise 6.2. Show that your results for the Ising model for case (i) agree with the results obtained in Exercise 6.1 and that there is no long range magnetic order above absolute zero for the Heisenberg model in one or two dimensions.
(6.4) Apply spin-wave theory to an antiferromagnet and show that the dispersion relation is

$$\hbar\omega = 4|JS\sin(qa)|. \tag{6.58}$$

Hence show that for an antiferromagnet $\omega \propto |q|$ at small q, whereas for a ferromagnet $\omega \propto q^2$.
(6.5) The Landau theory of a ferromagnet in a magnetic field H implies that the free energy is given by

$$F(M) = F_0 + a_0 (T - T_{\mathrm{C}}) M^2 + bM^4 - \mu_0 MH \tag{6.59}$$

where a_0 and b are positive constants. Show that this implies that

$$M^2 = u + v\frac{H}{M} \tag{6.60}$$

where u and v are constants that you should determine. By sketching M^2 against H/M for T just above T_C, just below T_C, and exactly at T_C, show how this method can be used to determine T_C. This idea is the basis of the Arrott plot which is a plot of M^2 against H/M.
(6.6) Find analytic solutions to the Stoner-Wohlfarth model for $\theta = 0$ and $\theta = \pi/2$.
(6.7) To illustrate the ideas behind the Mermin-WagnerBerenzinskii theorem, consider an excitation in two dimensions (see Fig. 6.29) in which the magnetization as a function of position $\mathbf{r} = (x, y)$ has constant magnitude $|\mathbf{M}|$ but varying direction specified by the vector $\mathbf{M}(\mathbf{r})$ given by

$$\mathbf{M}(\mathbf{r}) = \begin{cases} |\mathbf{M}|(\dot{\mathbf{y}}\cos(\pi r/R) - \dot{\mathbf{x}}\sin(\pi r/R)) & \text{if } r < R \\ -|\mathbf{M}|\dot{\mathbf{y}} & \text{if } r > R \end{cases} \quad (6.61)$$

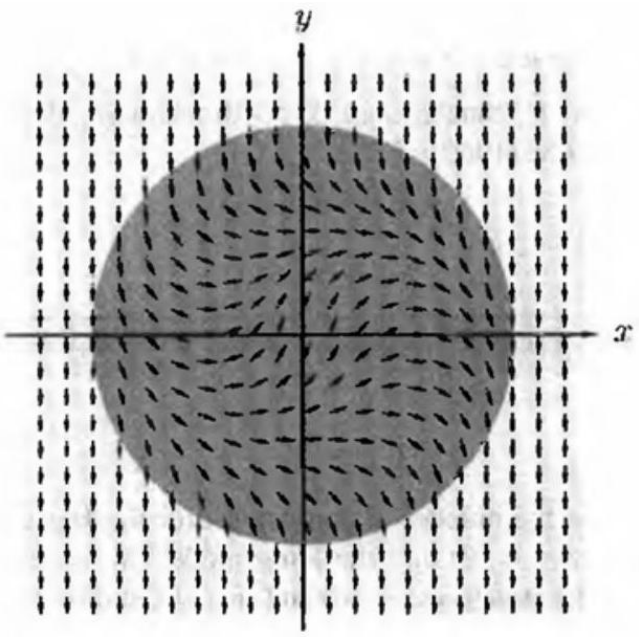


Fig. 6.29 An excitation in a 2D magnet.

Explain why the exchange energy cost due to thes excitation is proportional to

$$\int \left[(\Gamma M_1)^2 + (\Gamma M_v)^2\right] \mathrm{d}r \; \mathrm{d}y. \quad (6.62)$$

and show that it is a constant $i\pi^3|\mathbf{M}|^2$ 1 and therefore independent of R. This implies that very large scale fluctuations cun occur with no additional energy cost compared with a small-scale fluctuation, thereby destroying any long range order. This problem is based on an example presented in an article by A. S. Arrott and B. Heinrich. Journal of Magnetism and Magnetic Materals 93, 571 (1991).
16.81 Data are written onto a magnetic recording disk with an arcal density an 25 Gibits in ^{2}. By using somic educatied guesses, convert this areal density into other units: (a) bits per erosssectional area of a hydrogen atom. (b) copies of the complete plays by William Shakespeare (i.e. their information content) per postage stamp. Compare this density of intormation storage to that of the human genome.
16.9) Consider an isotropic cubic ferromagnet flattice parameter a) in a magnetic field, with Hamiltonian

$$\dot{H} = -J \sum_{ij=} \dot{\mathbf{S}}_i \cdot \dot{\mathbf{S}}_j + r\mu_B \sum_l \mathbf{B} \cdot \dot{\mathbf{S}}_j. \quad (6.63)$$

Show that it has a spin wate dispersion given by

$$\hbar\omega(q) = g\mu_{\mathrm{B}}B + 2JSq^2a^2$$

for $ya \ll 1$.
i6.JO; Show that the Hamiitomian

$$\dot{\mathrm{H}} = -\sum_{\mathrm{n}.\delta} \mathrm{J}(\delta)\dot{\mathrm{S}}_{\mathrm{n}} \quad \dot{\mathbf{S}}_{\mathrm{n}+\tilde{o}} \tag{6.65}$$

has a spin-wave dispersion grow by

$$\mathrm{hco}(\mathbf{q}) = \mathrm{hco}\left(O_1 + S\sum_{\delta} \mathrm{J}(\delta)(\mathbf{q}\cdot\delta)^{\lambda}\right. \tag{6.66}$$

for $|q \cdot \delta_{\mathrm{tnax}}| \ll I$ where $\delta_{\max}$ is the largent vector $\boldsymbol{\delta}$ for which Ji δ) gives any appreciable contribution.
(611) The sate {9} in eqn 6.23 represents a delocalized single flipped epin in a background of unfipped spins. In tinee dimensions it becomes

$$\left|\mathbf{q}_i' = \frac{1}{\sqrt{N}}\sum_{\mathbf{j}} \mathrm{e}^{\{\mathbf{q}\cdot\mathbf{R}_{\mathbf{j}}}\right| \mathbf{j}\Big\rangle \tag{6.67}$$

The probability that the flopped spin is at ste i is

$$P = |\langle \mathbf{q} \mid \mathbf{i}\rangle|^2. \tag{6.68}$$

Show that $\mu = 1\%N$ and hence that the lowered spin is distributed with equal probability. The operator for the tramsverse spin correlation function is

$$\dot{S}_{\mathbf{j}}^{\mathrm{x}}\dot{S}_{\mathbf{j}}^{\mathrm{x}} + \dot{S}_{\mathbf{j}}^{\mathrm{x}}\dot{S}_{j'}^{\mathrm{x}} \tag{6.69}$$

Show that

$$\left\{\boldsymbol{q}\left|\dot{S}_{\mathbf{j}}'\dot{S}_{\mathbf{j}}' + \dot{S}_{\mathbf{j}}^{y}\dot{S}_{\mathbf{j}}^{y} : \boldsymbol{q}\right\rangle = \frac{2S}{N}\cos\left(\boldsymbol{q}\cdot(\mathbf{R}_{\mathbf{j}} - \mathbf{R}_{\mathbf{j}'})\right)\right. \tag{6.70}$$

(for $i \neq i'$) and interpret this with reference to Fig. 6.13.
(6.12) Our calculation for the domain wall width of a Bloch wall made the assumption that the angle between adjacent spins in the wall was τ/N everywhere throughout the N spins of the wall. This was elearly an approximation! Here is an improved calculation.

Let the spin as a function of $\therefore$ a coordinate running perpendicular to the plane of the wall. be described by an angle $H(-)$. We assume that $H(-) \to 0$ as $z \to -\infty$ and $H(E)\cdots\pi$ as $z \to +\infty$. corresponding to a 180° wall. The energy is given by

$$E = \int_{-\infty}^{\infty}\left[A\left(\frac{\partial\theta}{\partial z}\right)^2 + f(\theta)\right]\mathrm{d}z. \tag{6.71}$$

where A is the contonuum exchange constant and $f(\theta)$ is a function which describes the anivorropy. Show that the energy is minimized when

$$\frac{\partial f}{\partial\dot{\theta}} - 2\Lambda\frac{a^2 f}{iz^2} = 0. \tag{6.72}$$

Multiply this equation by $\partial\theta/\partial$: and integrate with respect to : to show that

$$f(\theta) = A\left(\frac{\partial f}{\partial z}\right)^2 \tag{6.73}$$

Now suppose that $f(f)$ is given by

$$f(\theta) = K\sin^2 H. \tag{6.74}$$

where K wan anisotropy constant. Show that this implies that

$$\cos A(z) = -\tanh\left(\frac{\pi z}{\delta}\right). \tag{6.75}$$

where $\delta = \pi\sqrt{A/K}$. Sketch $A(\Sigma)$ and argue that δ is a suitable measure of the domain wall width. Show further that

$$E = 4\sqrt{\prime}\bar{A}\bar{K} \tag{6.76}$$

in this case.

(6.13) Consider the elosure domain structure in a ferromagnetic film as shown in Fig. 6.30. The easy axis is vertical and the hard axis is horizontal. The energy of the long 180° walls per unit area is m_W and the anisotropy energy density is K. Ignoring the contribution from the $9()^c$ walls

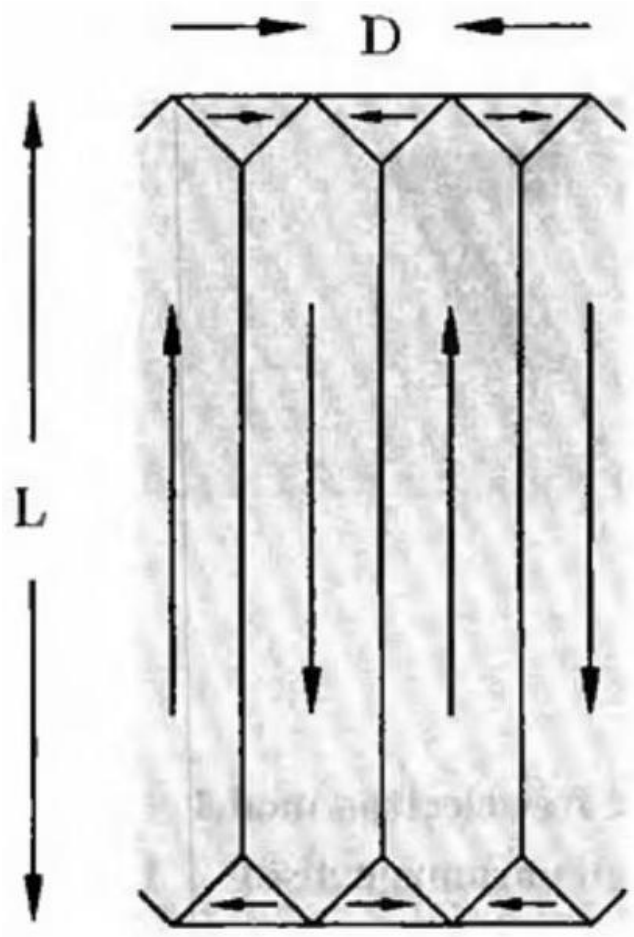


Fig. 6.30 A closure domain.

(ic. assume $L \gg D$), show that

$$D = \sqrt{\frac{2\sigma \mathrm{w} l}{K}}. \tag{6.77}$$

Assuming that $\sigma_{\mathrm{W}} = 2 \times 10^{-\frac{1}{3}} \mathrm{Jm}^{-2}$ and that $K = 4 \times 10^{4} \mathrm{Jm}^{-3}$ and $L = 4$ mm, estimate D.

9.7 Magnetism in metals

In this chapter the magnetic properties of metals are considered. In previous chapters we have concentrated on interacting but localized magnetic moments. The conduction electrons in metals are delocalized and can wander freely through the sample; they are known as itinerant electrons. In some cases the magnetic moments in metals are associated with the conduction electrons, in other cases the magnetic moments remain localized. In both cases paramagnetic and diamagnetic behaviour can occur. Ferromagnetism is possible under certain conditions. Most of the discussion in this chapter will be centred around the free electron model, which is introduced in the following section. The free electron model is a crude approximation to most real situations, but it is simple to consider and will allow the discussion to proceed a long way. Subsequent sections contain derivations of the magnetic properties of the electron gas which include Pauli paramagnetism, Landau diamagnetism, the origin of RKKY interactions, instabilities of the electron gas such as spindensity wave formation, and the Kondo effect which occurs when localized moments interact with the electron gas.

9.7.1 7.1 The free electron model

We begin our discussion of the magnetism of itinerant electrons by reviewing the free electron model. In this model, the periodic potential due to the lattice is ignored, and the electrons fill states up to the Fermi wave vector $k_{\mathbf{F}}$. Points in k-space are separated by $2\pi/L$ (see Fig. 7.1(a)) where $V = L^3$ is the volume of the sample, so that the number of states between k and $k + \mathrm{d}k$ is equal to $4\pi k^2\ \mathrm{d}k$, the volume of a spherical shell of radius k and width $\mathrm{d}k$ (see Fig. 7.1(b)) divided by $(2\pi/L)^3$, the volume occupied by one point in k-space. Each state is doubly occupied, by an electron with spin-up and an electron with spin-down, so there is an additional factor of two. Hence the density of states $g(k)\mathrm{d}k$ can be written as

$$g(k)\mathrm{d}k = \frac{2}{(2\pi/L)^3} \times 4\pi k^2\ \mathrm{d}k \tag{7.1}$$

where the factor of 2 takes care of the two spin-states of the electrons. Hence

$$g(k)\mathrm{d}k = \frac{Vk^2\ \mathrm{d}k}{\pi^2}. \tag{7.2}$$

If the material has N electrons, then at absolute zero ($T = 0$) these electrons will fill up the states up to a maximum wave vector of k_{F}. Hence

$$N = \int_0^{k_{\mathrm{F}}} g(k)\mathrm{d}k = \frac{Vk_F^3}{3\pi^2}, \tag{7.3}$$

so that

$$k_{\mathrm{F}}^3 = 3\pi^2 n \tag{7.4}$$

where $n = N/V$ is the number of electrons per unit volume. The Fermi energy E_{F} is defined by

$$E_{\mathrm{F}} = \frac{\hbar^2 k_{\mathrm{F}}^2}{2m_{\mathrm{e}}} \tag{7.5}$$

The density of states as a function of energy $E \propto k^2$ is proportional to $E^{1/2}$, i.e.

$$g(E) = \frac{\mathrm{d}n}{\mathrm{d}E} \propto E^{1/2} \tag{7.6}$$

so that

$$n = \int_0^{E_{\mathrm{F}}} g(E)\mathrm{d}E \propto E_{\mathrm{F}}^{3/2} \tag{7.7}$$

and hence

$$\frac{\mathrm{d}n}{n} = \frac{3}{2}\frac{\mathrm{d}E_{\mathrm{F}}}{E_{\mathrm{F}}}. \tag{7.8}$$

The density of states at the Fermi energy is therefore given by [1]

$$g\left(E_{\mathrm{F}}\right) = \left(\frac{\mathrm{d}n}{\mathrm{d}E}\right)_{E=E_{\mathrm{F}}} = \frac{3}{2}\frac{n}{E_{\mathrm{F}}}. \tag{7.9}$$

We note in passing that another useful expression for $g\left(E_{\mathrm{F}}\right)$ can be obtained by combining eqn 7.4 , eqn 7.5 and eqn 7.9 to yield

$$g\left(E_{\mathbf{F}}\right) = \frac{m_{\mathrm{e}} k_{\mathbf{F}}}{(\pi\hbar)^2} \tag{7.10}$$

which shows that $g\left(E_{\mathrm{F}}\right) \propto m_{\mathrm{e}}$. Many important properties depend on the density of states at the Fermi energy and therefore it is useful to know that it is proportional to the electron's mass. In many systems of interest, the electron's mass is enhanced above its free space value due to the effect of the band structure or interactions.

In the free electron model we ignore the periodic potential due to the lattice. However if it is included as a perturbation (the nearly free electron model) it turns out that it has very little effect except when the wave vector of the electron is close to a reciprocal lattice vector. At such points in k-space energy gaps appear in the dispersion relation (see Fig. 7.2).

So far, everything has been treated at $T = 0$. When $T > 0$, the density of states $g(E)$ is unchanged but the occupancy of each state is governed by the Fermi function $f(E)$ which is given by

$$f(E) = \frac{1}{\mathrm{e}^{(E-\mu)/k_{\mathrm{B}}T} + 1} \tag{7.11}$$

where μ is the chemical potential which is temperature dependent. This function is plotted in Fig. 7.3. At $T = 0, f(E)$ is a step function, taking the value 1 for $E < \mu$ and 0 for $E > \mu$. The step is smoothed out as the temperature T increases. When the Fermi function is close to a step function, as is the usual case for most metals at pretty much all temperatures below their melting temperature, the electrons are said to be in the degenerate limit. The Fermi

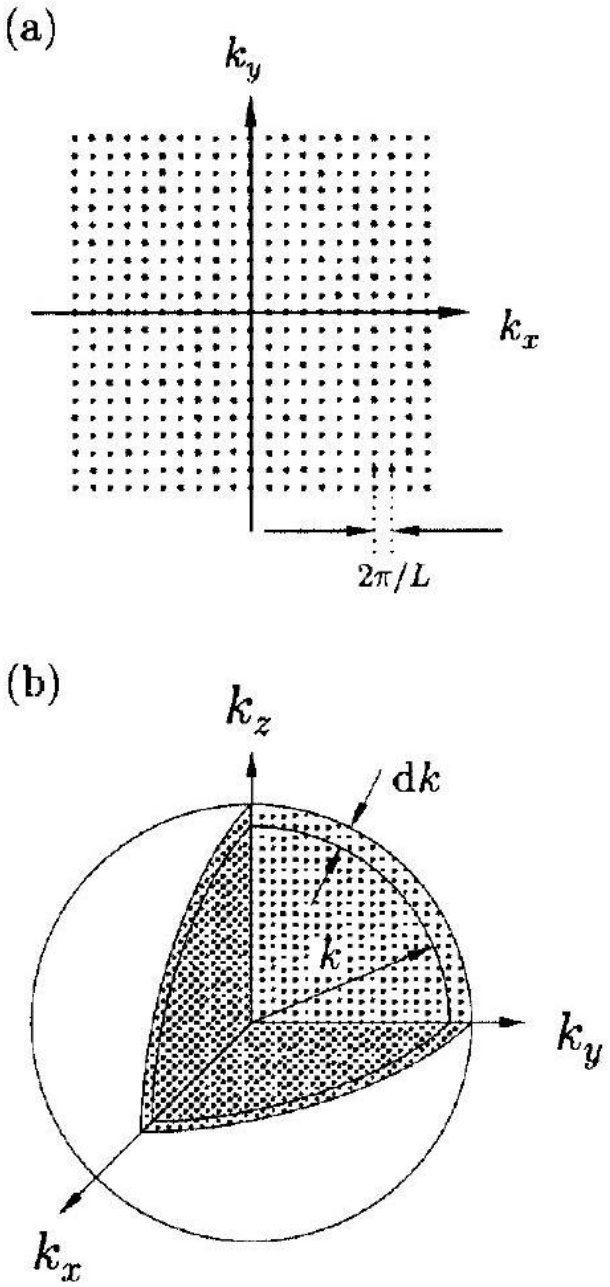


Fig. 7.1 (a) Electron states are separated by $2\pi/L$. Each state can be doubly occupied and occupies a volume $(2 \pi / L)^{3
$. (b) The density of states can be calculated by considering the volume in k-space between states with wave vector k and states with wave vector $k + \mathrm{d}k$, namely $4\pi k^2\ \mathrm{d}k$.}

[1] Equation 7.9 could also be derived using eqns 7.2, 7.4 and 7.5 directly.

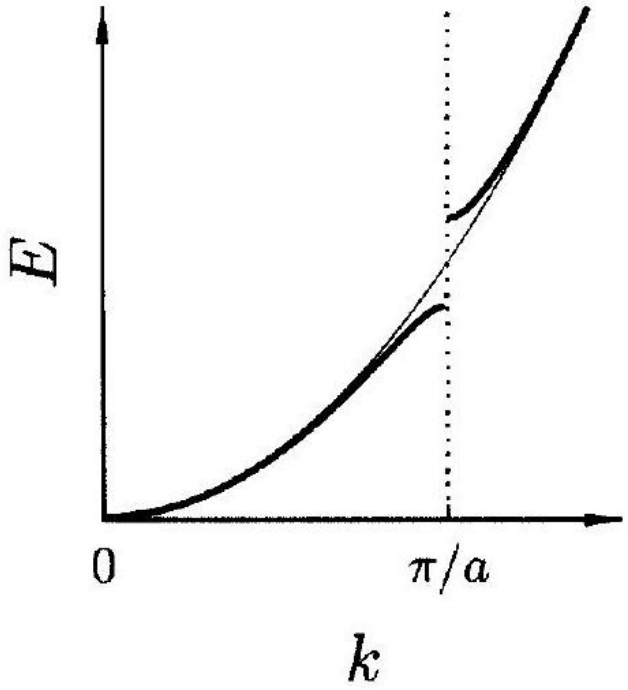
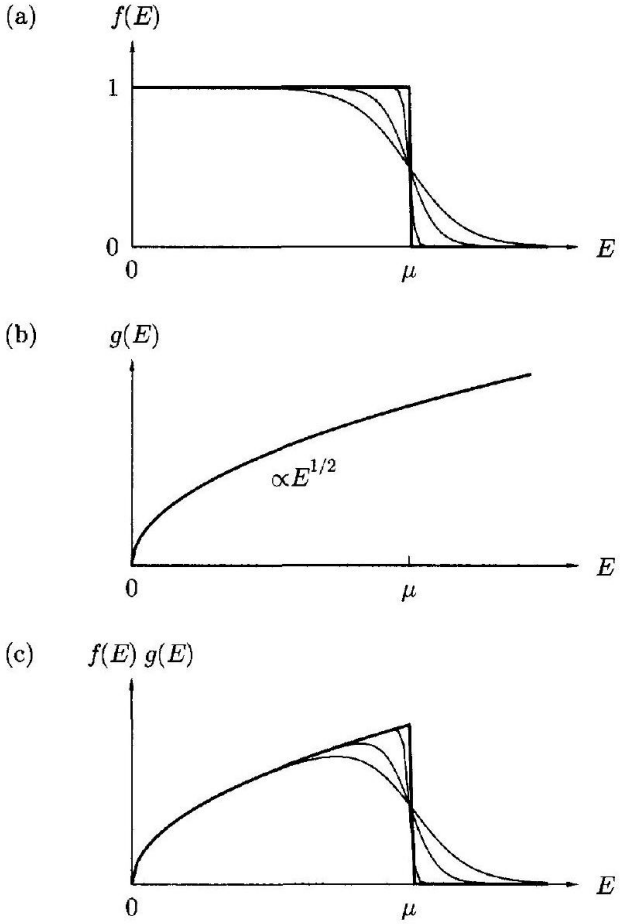


Fig. 7.2 The energy gap at the Brillouin zone boundary.

Fig. 7.3 (a) The Fermi function $f(E)$ defined by eqn 7.11. The thick line is for $T = 0$. The step function is smoothed out as the temperature increases. The temperatures shown are $T=0, T=0.01 \mu / k_{\mathrm{B}}$, $T=0.05 \mu / k_{\mathrm{B}}$ and $T = 0.1\mu/k_{\mathrm{B}}$. (b) The density of states $g(E)$ for a free electron gas is proportional to $E^{1/2}$. (c) $f(E)g(E)$ for the same temperatures as in (a).}



function is a consequence of the Pauli exclusion principle, that each electron must have a unique set of quantum numbers and no two electrons can sit in the same state. For $(E - \mu) \gg k_{\mathrm{B}}T$ the Fermi function approaches the MaxwellBoltzmann form $\mathrm{e}^{-(E-\mu)/k_{\mathrm{B}}T}$ which is known as the non-degenerate limit.

The Fermi energy is the energy of the highest occupied level at $T = 0$ and is determined by the equation

$$\int_0^{E_{\mathrm{F}}} f(E)g(E)\mathrm{d}E = n \tag{7.12}$$

The function $f(E)g(E)$ is shown in Fig. 7.3(c). At $T = 0$ we easily find that the Fermi energy precisely equals the chemical potential: $E_{\mathbf{F}} = \mu$. For $T > 0$, a tedious calculation gives

$$\mu = E_{\mathrm{F}}\left[1 - \frac{\pi^2}{12}\left(\frac{k_{\mathrm{B}}T}{E_{\mathrm{F}}}\right)^2 + O\left(\left(\frac{k_{\mathrm{B}}T}{E_{\mathrm{F}}}\right)^4\right)\right] \tag{7.13}$$

but this means that equating E_{F} and μ is good to 0.01% for typical metals even at room temperature, although it is worthwhile keeping in the back of one's mind that the two quantities are not the same. The Fermi surface is the

set of points in k-space whose energy is equal to the chemical potential. If the chemical potential lies in a gap, then the material is a semiconductor or an insulator and there will be no Fermi surface. Thus a metal is a material with a Fermi surface.

9.7.2 7.2 Pauli paramagnetism

Each k-state in a metal can be doubly occupied because of the two possible spin states of the electron. Each electron in a metal is therefore either spinup or spin-down. When a magnetic field is applied, the energy of the electron is raised or lowered depending on its spin. This gives rise to a paramagnetic susceptibility of the electron gas and is known as Pauli paramagnetism.

7.2.1 Elementary derivation

Initially, we neglect the orbital contribution and take $g = 2$. We also neglect smearing of the Fermi surface due to finite temperature. As shown in Fig. 7.4, in an applied magnetic field, the electron band is spin-split into two spin subbands separated by $g\mu_{\mathrm{B}}B = 2\mu_{\mathrm{B}}B$. We will assume that $g\mu_{\mathrm{B}}B$ is a very small energy so that the splitting of the energy bands is very small. The number of extra electrons per unit volume with spin-up is $n_{\uparrow} = \frac{1}{2}g\left(E_{\mathrm{F}}\right)\mu_{\mathrm{B}}B$. This is also the number per unit volume of the deficit of electrons with spin-down, $n_{\downarrow} = \frac{1}{2}g\left(E_{\mathrm{F}}\right)\mu_{\mathrm{B}}B$. Thus the magnetization is given by

$$M = \mu_{\mathrm{B}}\left(n \wedge -n_{\downarrow}\right) = g\left(E_{\mathrm{F}}\right)\mu_{\mathrm{B}}^{2}B \tag{7.14}$$

and the magnetic susceptibility χ_{P} (the subscript ' P ' denoting the Pauli susceptibility) by

$$\chi_{\mathrm{P}} = \frac{M}{H} \approx \frac{\mu_0 M}{B} = \mu_0\mu_{\mathrm{B}}^{2}g\left(E_{\mathrm{F}}\right) \tag{7.15}$$

$$= \frac{3n\mu_0\mu_{\mathrm{B}}^{2}}{2E_{\mathrm{F}}} \tag{7.16}$$

where the final equality is obtained using eqn 7.9 . Because $\chi_{\mathrm{P}} \ll 1$ we are justified in writing $\chi_P \approx \mu_0 M/B$ (see Section 1.1.4).

Our expression for Pauli paramagnetism is temperature independent, although admittedly this is because we started out by ignoring the smearing of the Fermi surface due to finite temperature. However, if temperature is included, it makes only a rather small correction (see Exercise 7.1). Pauli paramagnetism is a weak effect, much smaller than the paramagnetism observed in insulators at most temperatures due to Curie's law. This is because in paramagnetic insulators at least one electron on every magnetic atom in the material contributes. but in a metal, it is only those electrons close to the Fermi surface which play a rôle. The small size of the paramagnetic susceptibility of most metals was something of a puzzle until Pauli pointed out that it was a consequence of the fact that electrons obeyed Fermi Dirac, rather than classical, statistics.

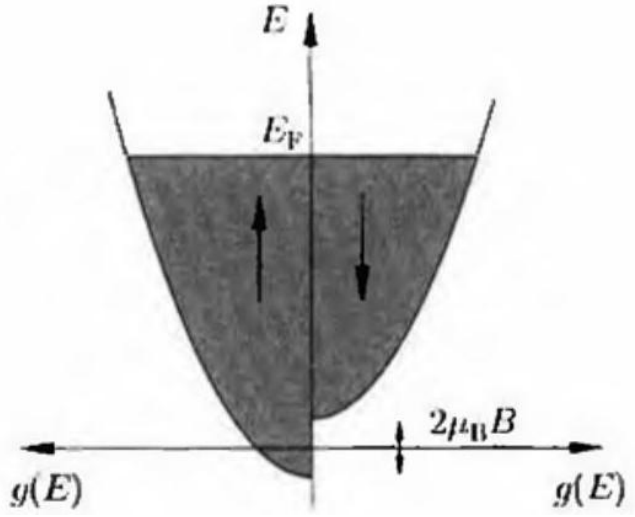


Fig. 7.4 Density of states showing splitting of energy bands in a field B. The splitting is shown greatly exaggerated.

7.2.2 Crossover to localized behaviour

The effect of the Fermi Dirac statistics and the crossover between Pauli paramagnetism and localized moment behaviour can be illustrated by rederiving eqn 7.15. One should strictly write the number of electrons per unit volume of each spin state as

$$\begin{aligned} n_\uparrow &= \frac{1}{2}\int_0^\infty g\left(E+\mu_\mathrm{B}B\right) f(E)\mathrm{d}E \\ n_\downarrow &= \frac{1}{2}\int_0^\infty g\left(E-\mu_\mathrm{B}B\right) f(E)\mathrm{d}E \end{aligned} \tag{7.17}$$

Thus for small B, the magnetization is $M = \mu_\mathrm{B}\left(n_\uparrow - n_\downarrow\right)$ and hence

$$\begin{aligned} M &\approx \mu_\mathrm{B}^2 B \int_0^\infty \frac{\mathrm{d}g}{\mathrm{d}E} f(E)\mathrm{d}E \\ &= \mu_\mathrm{B}^2 B \left(\left[g(E)f(E)\right]_0^\infty - \int_0^\infty \frac{\mathrm{d}f}{\mathrm{d}E} g(E)\mathrm{d}E \right) \end{aligned} \tag{7.18}$$

where the second line is obtained by performing the integration of the first line by parts. Using the fact that $g(0) = 0$ and $f(\infty) = 0$ (see Fig. 7.3), the first term is zero. Hence

$$M \approx \mu_\mathrm{B}^2 B \int_0^\infty \left(-\frac{\mathrm{d}f}{\mathrm{d}E}\right) g(E)\mathrm{d}E \tag{7.19}$$

In the degenerate limit at $T = 0, -\mathrm{d}f/\mathrm{d}E$, the differential of a step function, is a delta function at E_F, i.e.

$$-\frac{\mathrm{d}f}{\mathrm{d}E} = \delta\left(E - E_\mathbf{F}\right) \tag{7.20}$$

so that we recover $M = \mu_\mathrm{B}^2 B g\left(E_\mathrm{F}\right)$ and hence $\chi = \mu_0 \mu_\mathrm{B}^2 g\left(E_\mathrm{F}\right)$ in agreement with eqn 7.15. In the non-degenerate limit, $f(E) \approx \mathrm{e}^{-(E-\mu)/k_\mathrm{B}T}$ so that

$$-\frac{\mathrm{d}f}{\mathrm{d}E} = \frac{f}{k_\mathrm{B}T} \tag{7.21}$$

and the magnetization is

$$M = \frac{\mu_B^2 B}{k_B T} \int_0^\infty f(E) g(E) \mathrm{d}E = \frac{n \mu_B^2 B}{k_B T} \tag{7.22}$$

so that

$$\chi = n \mu_0 \mu_B / k_B T \tag{7.23}$$

in agreement with eqn 2.28 . The susceptibility is then equivalent to n localized moments (of magnitude μ_B) per unit volume. As stated above, for most metals E_F is a few eV and so the degenerate limit holds well at all temperatures below the melting temperature. However, for a material in which the carrier concentration is low (for example in a doped semiconductor) and $E_F \propto n^{2/3}$ is much smaller, the non-degenerate limit may be reached and a Curie-like susceptibility obtained.

7.2.3 Experimental techniques

The expressions that we have derived for the Pauli paramagnetism of metals agree moderately well with experiment, but can be improved by correcting for the effect of electron-electron interactions. The spin susceptibility of a metal can be extracted from NMR measurements which are much more sensitive to the field due to the spin magnetic moment of the conduction electrons than to the fields which arise from the electron's orbital motion (which give rise to the diamagnetic effects considered in Section 7.6).

The effect of the contact interaction between the conduction electron spin and the nuclear spin leads to a small shift $\Delta\omega$, known as a Knight shift, in the nuclear resonance frequency ω. It can be understood by imagining that individual conduction electrons hop on and off a given nucleus; the net hyperfine coupling which the nucleus experiences is the result of averaging over all the electron spin orientations. This net hyperfine coupling will be zero without an applied field because the average of the electronic spin orientations will vanish; the net hyperfine coupling will be non-zero in a nonzero static field because this will polarize the electron spins. The Knight shift, $K = \Delta\omega/\omega$, is therefore proportional to the conduction electron density at the nucleus (which expresses the dependence on the coupling strength) and also to the Pauli spin susceptibility (which expresses the extent to which an applied field polarizes the electrons).

The static average of the hyperfine interactions causes the Knight shift. Fluctuations about this average provide a mechanism for T_1 relaxation (known as Korringa relaxation). The dominant T_1 processes are flip-flop transitions of the electron and nuclear (or muon) spins, in which the difference in electron and nuclear Zeeman energies is taken up by a change in kinetic energy of the conduction electron. The exchange in energy between the nucleus and the conduction electrons is very small, so only electrons within $k_B T$ of the Fermi surface are able to participate since only these have empty states nearby into which they can make a transition. Thus for simple metals the spin-lattice relaxation rate T_1^{-1} is proportional to temperature. The Knight shift, usually expressed in the dimensionless form $\Delta\omega/\omega$, and the Korringa relaxation rate T_1^{-1} are usually connected by the equation

$$\frac{1}{T_1} \propto T \left(\frac{\Delta\omega}{\omega} \right)^2 \tag{7.24}$$

which is known as the Korringa relation.

9.7.3 7.3 Spontaneously spin-split bands

The magnetic moment per atom in iron is about $2.2\mu_B$ (see Table 5.1). This non-integral value is not possible to understand on the basis of localized moments on atoms. It is therefore

strong evidence for band ferromagnetism (also known as itinerant ferromagnetism) in which the magnetization is due to spontaneously spin-split bands. In this section we will explore some models which can be used to understand how bands in some materials can become spontaneously spin-split.

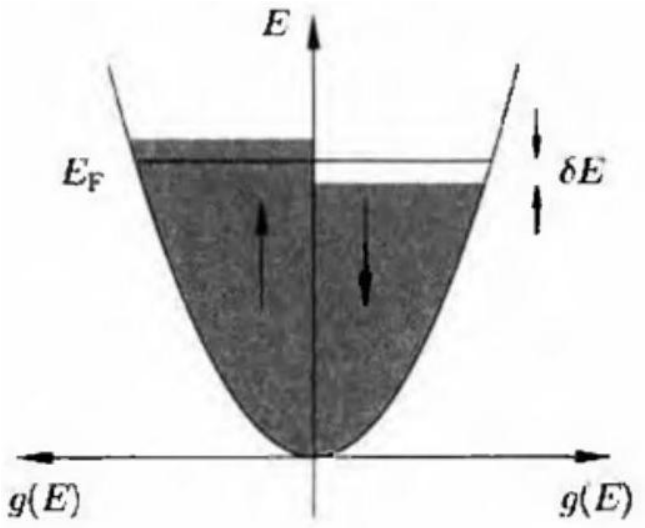


Fig. 7.5 Density of states showing spontaneous splitting of cnergy bands without an applied magnetic field.

[2] The molecular ficld is due to exchange. Exchange is due to the Coulomb interaction. In more advanced treatments. the Coulomb energy is included directly as $U_n \uparrow^n \downarrow$ which yields the same result.

Edmund C. Stoner (1899-1968)

In molecular field theory we say that all spins 'feel' an identical average exchange field λM produced by all their neighbours. In a metal, the molecular field can magnetize the electron gas because of the Pauli paramagnetism XP. The resulting magnetization of the electron gas M would in turn be responsible for the molecular field. This is a chicken-and-egg scenario (also known as bootstrupping); can this positive feedback mechanism lead to spontaneous ferromagnetism? Presumably yes, if λ (expressing how much molecular field you get for a given M) and χ_{P} (expressing how much magnetization you get for a given molecular field) are both large enough.

It is desirable to make the above heuristic argument a little more rigorous! The question that we need to ask is: can the system as a whole save energy by becoming ferromagnetic?

Let us first imagine that in the absence of an applied magnetic field we take a small number of electrons at the Fermi surface from the spin-down band and place them in the spin-up band. Specifically we take spin-down electrons with energies from $E_{\mathrm{F}} - \delta E$ up to E_{F} and flip their spins, placing them in the spinup band where they sit with energies from E_{F} up to $E_{\mathrm{F}} + \delta E$. This situation is illustrated in Fig. 7.5. The number of electrons moved is $g\left(E_{\mathrm{F}}\right) \delta E/2$ and they increase in energy by δE. The total energy change is $g\left(E_{\mathrm{F}}\right) \delta E/2 \times \delta E$. The total kinetic energy change $\Delta E_{\mathrm{K.E}}$ is therefore

$$\Delta E_{\mathrm{K.F.}} = \frac{1}{2} g\left(E_{\mathrm{F}}\right) (\delta E)^2 \tag{7.25}$$

This is an energy cost so this process looks unfavourable. However, the interaction of the magnetization with the molecular field gives an energy reduction which can outweigh this cost.

The number density of up-spins is $n_{-} = \frac{1}{2}\left(n + g\left(E_{\mathrm{F}}\right)\delta E\right)$ and the number density of down-spins is $n_{\downarrow} = \frac{1}{2}\left(n - g\left(E_F\right)\delta E\right)$. Hence the magnetization is $M = \mu_{\mathrm{B}}\left(n_{\uparrow} - n_{\downarrow}\right)$, assuming each electron has a magnetic moment of $1\mu_{\mathrm{B}}$. The molecular field energy is

$$\Delta E_{\mathrm{P.E.}} = -\int_0^M \mu_0\left(\lambda M'\right)\mathrm{d}M' = -\frac{1}{2}\mu_0\dot{\lambda}M^2 = -\frac{1}{2}\mu_0\mu_{\mathrm{B}}^2\dot{\lambda}\left(n_{\star} - n_{\frac{1}{\mathrm{i}}}^2\right) \tag{7.26}$$

Writing $U = \mu_0\mu_{\mathbf{B}}^2\lambda$, where U is a measure of the Coulomb energy, [2] we have

$$\Delta E_{\mathrm{P.E.}} = -\frac{1}{2}U\left(g\left(E_{\mathrm{F}}\right)\delta E\right)^2 \tag{7.27}$$

Hence the total change of energy ΔE is

$$\Delta E = \Delta E_{\mathrm{K.E.}} + \Delta E_{\mathrm{P.E.}} = \frac{1}{2}g\left(E_{\mathrm{F}}\right)\left(\delta E\right)^2\left(1 - Ug\left(E_{\mathrm{F}}\right)\right) \tag{7.28}$$

Thus spontaneous ferromagnetism is possible if $\Delta E < 0$ which implies that

$$Ug\left(E_{\mathrm{F}}\right) \geq 1 \tag{7.29}$$

which is known as the Stoner criterion. This condition for the ferromagnetic instability requires that the Coulomb effects are strong and also that the density of states at the Fermi energy is large. If there is spontaneous ferromagnetism, the spin-up and spin-down bands will be split by an energy Δ. where Δ is the exchange splitting, in the absencc of an applied magnetic field.

If the Stoner criterion is not satisfied, then spontaneous ferromagnetism will not occur. But the susceptibility may be altered. We can calculate this easily by including both the effects of an applied magnetic field and the interactions. The magnetization produced by an energy shift δE is simply $M = \mu_{\mathrm{B}}\left(N_{\uparrow} - N_{\downarrow}\right) = 2\mu_{\mathrm{B}}g\left(E_{\mathrm{F}}\right)\delta E$. Thus

$$\Delta E = \frac{1}{2}g\left(E_{\mathrm{F}}\right)\left(\delta E\right)^2\left(1 - Ug\left(E_{\mathrm{F}}\right)\right) - MB \tag{7.30}$$

$$= \frac{M^2}{2\mu_{\mathrm{B}}^2 g\left(E_{\mathrm{F}}\right)}\left(1 - Ug\left(E_{\mathrm{F}}\right)\right) - MB \tag{7.31}$$

This is minimized when

$$\frac{M}{\mu_{\mathrm{B}}^2 g\left(E_{\mathrm{F}}\right)}\left(1 - Ug\left(E_{\mathrm{F}}\right)\right) - B = 0 \tag{7.32}$$

so that the magnetic susceptibility χ is given by

$$\chi = \frac{M}{H} \approx \frac{\mu_0 M}{B} = \frac{\mu_0\mu_{\mathrm{B}}^2 g\left(E_{\mathrm{F}}\right)}{1 - Ug\left(E_{\mathrm{F}}\right)} = \frac{\chi_{\mathrm{P}}}{1 - Ug\left(E_{\mathrm{F}}\right)} \tag{7.33}$$

This is larger than the value χ_{P} expected without the presence of Coulomb interactions by a factor $\left(1 - Ug\left(E_{\mathrm{F}}\right)\right)^{-1}$, a phenomenon known as Stoner enhancement. It is responsible for the enhanced Pauli susceptibility measured in the metals Pd and Pt which can both be thought of as systems on the verge of ferromagnetism; they have a large enough value of the parameter $Ug\left(E_{\mathbf{F}}\right)$ to cause a significant enhancement of the magnetic susceptibility but not large enough (i.e. not sufficiently close to 1) to cause spontaneous ferromagnetism.

9.7.4 7.4 Spin-density functional theory

So far we have used free electron models or nearly free electron models in our discussion. It is possible to improve on this with more advanced methods, and one of these will be discussed in this section. In real systems one cannot ignore Coulomb interactions between electrons and the effect of exchange interactions on the motion of the electrons. The positions and motions of all the particles are correlated because the particles interact with each other and exert forces upon each other as they move. Thus the interactions lead to correlations appearing between particles. Such correlations can be very difficult to deal with theoretically, but a useful and successful approach is that of density functional theory.

In this theory it is recognized that the ground state energy of a many electron system can be written as a functional [3] of the electron density $n(\mathbf{r})$. The functional contains three contributions, a kinetic energy, a Coulomb energy due to the electrostatic interactions between the charged particles in the system, and a term called the exchange-correlation energy that captures all the manybody interactions. Rather than dealing with the wave function $\psi(\mathbf{r})$, in density functional theory one only has to consider the electron density $n(\mathbf{r}) = |\psi(\mathbf{r})|^2$, and this results in a considerable simplification. Minimizing the energy functional leads to an equation which can be used to find the ground state energy.

[3] A function is a rule which maps one number into another number. For example the function

$$f(x) = x^2$$

maps the number 2 into the number 4. A functional is a rule which maps an entire function into a number. For example, the functional

$$F[f] = \int_{-1}^{1} f(x)\mathrm{d}x$$

maps the function $f(x) = x^2$ into the number 2/3.

[4] The theory works very well for metals, but not so well for localized systems. This is because the Coulomb energy represents the response of a particular electron to the electron density, but the density is due to all electrons, including the particular electron. So electron self-interactions are unavoidably, but mistakenly, included. This is not so bad for a metal where any given electron is just one part of the vast conduction electron ocean, but disastrous in a localized system in which a particular state may be occupied by only one electron.

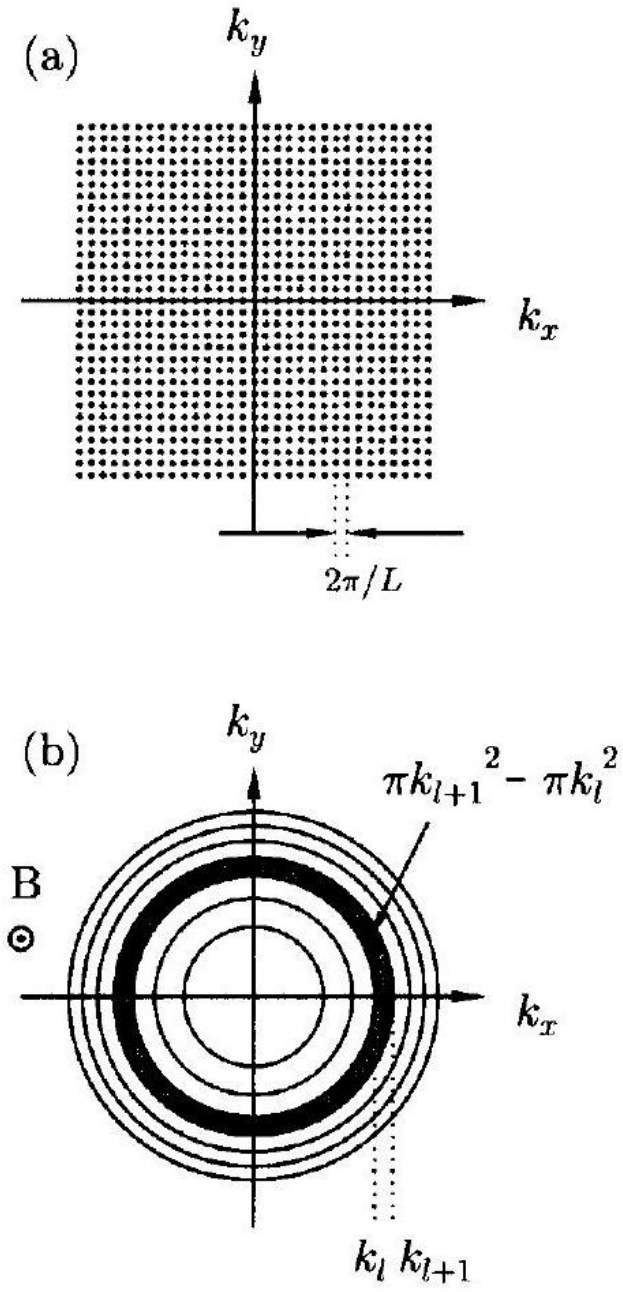


Fig. 7.6 (a) With $B = 0$ the electron states are uniformly spaced in k-space, with each point separated by $2\pi/L$. (b) The application of a magnetic field results in only certain allowed energies. These Landau level states have a degeneracy which can be calculated by considering the area in k-space between the level with wave vector k_{l+1} and the level with wave vector k_l.

It has been proved that minimizing the functional really does yield the exact ground state energy, even though one is using the electron density, not the wave function.

The problem is that the energy functional is not known in full! The kinetic energy and Coulomb energy can be written down but the exchangecorrelation energy is unknown. However, one can use the value of the exchange-correlation energy from known results of many-electron interactions in a homogenous electron gas (an electron system of constant density). This is called the local density approximation (LDA) and it amounts to saying that at each point in space, $\mathbf{r}$, where the electron density is $n(\mathbf{r})$, an electron responds to the many-body interactions as if the electron density throughout all space (i.e. at positions $\mathbf{r}' \neq \mathbf{r}$) takes the value $n(\mathbf{r})$.

The contributions to the energy from different positions in space are all added together (one integrates over all volume elements); these contributions are different for each volume element, depending on the local electron density. The LDA is exact for a perfect metal (which does have uniform electron density) but is not so good for systems with wildly spatially varying electron density. [4]

The extension of density functional theory to include the effects of spin polarization is called spin-density functional theory. In magnetic systems this is used together with the local spin-density approximation (often referred to as the LSDA) in which the exchange correlation potential depends not only on the local electron density but also on the local spin density (the difference between the electron density of spin-up and spin-down electrons). This technique can be used to perform realistic calculations of electronic band structure and obtain quantitative information concerning the spin density of real systems. In the next section, however, we retreat to the comparitive safety of the simple free electron model!

9.7.5 7.5 Landau levels

Pauli paramagnetism is an effect associated with the spin of the electrons. What about the orbital contribution of the electrons? This will be evaluated using the following argument. We assume that the free electron Hamiltonian is

$$\frac{\hat{\mathbf{p}}^2}{2m_\mathrm{e}}\psi = -\frac{\hbar^2}{2m_\mathrm{e}}\nabla^2\psi = E\psi, \tag{7.34}$$

and the wave functions are plane waves

$$\psi = \frac{1}{\sqrt{V}}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}, \tag{7.35}$$

so that the energy E is given by

$$E = \frac{\hbar^2 k_x^2}{2m_\mathrm{e}} + \frac{\hbar^2 k_y^2}{2m_\mathrm{e}} + \frac{\hbar^2 k_z^2}{2m_\mathrm{e}} \tag{7.36}$$

As before, with no magnetic field applied, the electron states are uniformly spaced in k-space, with each point separated by $2\pi/L$ (see Fig. 7.6(a)).

In the presence of a magnetic field $\mathbf{B} = (0, 0, B)$, the momentum operator $\hat{\mathbf{p}} = -\mathrm{i}\hbar\nabla$ must be replaced by $-\mathrm{i}\hbar\nabla + e\mathbf{A}$ where a useful choice for the magnetic vector potential is $\mathbf{A} = (0, Bx, 0)$. Hence eqn 7.34 becomes

$$\left(\frac{\hat{p}_x^2}{2m_\mathrm{e}} + \frac{\left(\hat{p}_y + eB\hat{x}\right)^2}{2m_\mathrm{e}} + \frac{\hat{p}_z^2}{2m_\mathrm{e}}\right)\psi = E\psi. \tag{7.37}$$

Since this expression contains the operators $\hat{x}, \hat{p}_x, \hat{p}_y$ and $\hat{p}_z$, the wave function remains a plane wave in the y and z directions, so that $\hat{p}_y\psi = \hbar k_y\psi$ and $\hat{p}_z\psi = \hbar k_z\psi$. Thus eqn 7.37 can be rearranged (using separation of variables) to give

$$\left(\left[\frac{\hat{p}_x^2}{2m_\mathrm{e}} + \frac{1}{2}m_\mathrm{e}\omega_\mathrm{c}^2\left(x - x_0\right)^2\right] + \frac{\hbar^2 k_z^2}{2m_\mathrm{e}}\right)\psi = E\psi, \tag{7.38}$$

where $x_0 = -\hbar k_y/eB$ and

$$\omega_\mathrm{c} = \frac{eB}{m_\mathrm{e}} \tag{7.39}$$

is the cyclotron frequency. The term in square brackets in eqn 7.38 is the Hamiltonian for a one-dimensional harmonic oscillator. Hence the energy eigenvalues correspond to

$$E = \left(l + \frac{1}{2}\right)\hbar\omega_c + \frac{\hbar^2 k_z^2}{2m_e} \tag{7.40}$$

where l is an integer. The energy eigenfunctions are a product of plane waves in the y and z directions and a one-dimensional harmonic oscillator wave function in the x direction.

In the treatment presented here, the quantity x_0 is the 'centre' of the harmonic oscillator motion in the x direction. This takes different values depending on the value of k_y.[5] A further useful way of attacking this problem is to choose a gauge which reflects the cylindrical symmetry inherent in this situation (this is done in Exercise 7.2) and this results in identical energy eigenvalues to those in eqn 7.40 but the eigenfunctions are a product of a plane wave in the z direction and a function of $\sqrt{k_x^2 + k_y^2}$. In whichever gauge the problem is solved, the allowed states can be thought of as occurring on Landau tubes which lie parallel to k_z (see Fig. 7.7). The application of a magnetic field causes the electrons to form Landau levels (as they are viewed when thinking of the states as a function of energy) and each Landau level can be labelled by the values of the quantum number l (see Fig. 7.6(b)). The degeneracy of a Landau level can be calculated by noting that the area between one Landau level and the next in k-space (see Fig. 7.6(b)) is given by

$$\pi k_{l+1}^2 - \pi k_l^2 = \frac{2m_e\pi}{\hbar^2}\left(\left(l + 1 + \frac{1}{2}\right)\hbar\omega_c - \left(l + \frac{1}{2}\right)\hbar\omega_c\right) = \frac{2m_e\pi\omega_c}{\hbar} \tag{7.41}$$

so that since two electrons occupy each state, and one state occupies an area of $(2\pi/L)^2$ in the $k_x k_y$ plane, the degeneracy p of one Landau level is

$$p = \frac{4m_e\pi\omega_c}{\hbar(2\pi/L)^2} = \frac{m_e L^2\omega_c}{\pi\hbar} \tag{7.42}$$

Notice that the same energy eigenvalue would have been obtained if the magnetic vector potential was chosen to have been $A = (-By, 0, 0)$ (which gives the same value for $B = \nabla \times A = (0, 0, B)$) but the energy eigenfunctions would have been in this case a product of plane waves in the x and z directions and a one-dimensional harmonic oscillator wave function in the y direction. The "choice of gauge" (i.e. the way you write down $\mathbf{A}$) doesn't affect the energy eigenvalues but does affect the set of wave functions that you derive.

[5] If you make the choice A $= (-By, 0, 0)$ you obtain a parameter y_0 which is proportional to k_x which is the centre of the harmonic oscillator motion in the y direction. It turns out that the operators corresponding to x_0 and y_0 do not commute, so that x_0 and y_0 cannot take definite values simultaneously.

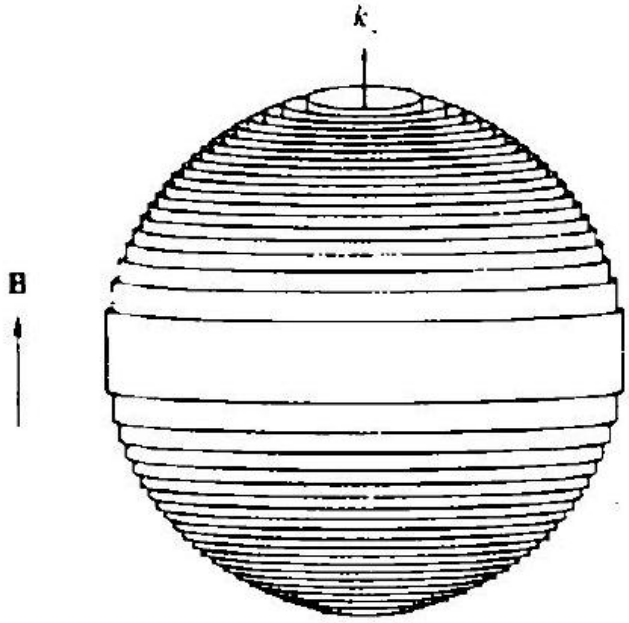


Fig. 7.7 Landau tubes

9.7.6 7.6 Landau diamagnetism

Using the results from the previous section, we are now in a position to calculate the diamagnetic response of the electron gas. The idea is that as a magnetic field is applied, and the clectron distribution breaks up into a series of Landau levels, the total energy of the system may change. This change of energy with field is equivalent to a magnetization of the system. The phenomenon is known as Landou diamagnetism.

The calculation is based on the observation that for different values of k_z the highest occupied Landau level is different (sce Fig. 7.7 which shows that the Fermi surface cuts different Landau tubes for different values of k_n). For a particular value of k_2. we define the energy $E_\perp$ associated with the electronic motion perpendicular to the field by the equation

$$E_\perp = E_\mathrm{F} - \frac{\hbar^2 k_2^2}{2m_\mathrm{e}} \tag{7.43}$$

In the absence of a magnetic field electrons can fill up the states to a maximum value of $E_\perp$ for a particular choice of k_z, but the application of a magnetic field allows discrete Landau levels. We give the highest occupied Landau level the label E_l. The electrons in the $B = 0$ problem can be broken up into blocks whose average energy is the corresponding Landau level energy (except for the top block which is partially filled).

Let us define a parameter x by

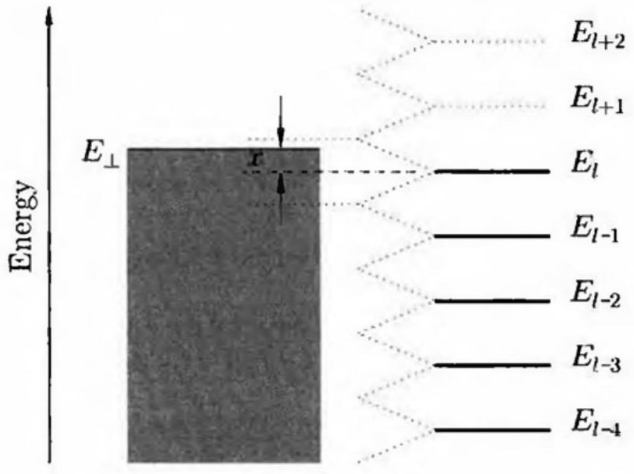
$$x = E_- - E_l \tag{7.44}$$

(see Fig. 7.8). The mean energy in the highest occupied block is

$$\frac{1}{2}\left[E_\perp + E_l - \frac{h\omega_c}{2}\right] = E_\perp - \frac{x}{2} - \frac{h\omega_c}{4} \tag{7.45}$$

(a) $B = 0$
(b) $B \neq 0$

Fig. 7.8 (a) In the absence of a magnetic field electrons can fill up the states to a maximum value of $E_{\perp}$ which is detined by eqn 7.43. The filled states at $T = 0$ arc shown shaded. The 'bighest occupied block' described in the text is labelled E_j. The parameter $x = E_\perp - E_l$ measures the energy between the middle of this block (dot-dash line) and the highest occupied state. (b) The magnetic field allows only discrete Landau levels with energy $\ldots E_{l..2}, E_{l-1}, E_l, E_{l+1} \ldots$. The occupied Landau levels at $T = 0$ are shown as heavy lines and the highest occupied Landau level is labelled as E_l.}



and its occupancy is

$$\frac{p}{\hbar\omega_c}\left[E_\perp - \left(E_l - \frac{\hbar\omega_c}{2}\right)\right] = \frac{p}{2}\left[1 + \frac{2x}{\hbar\omega_c}\right]. \tag{7.46}$$

Multiplying these together gives the energy of the top block. This has been evaluated for a particular choice of k_z but it is necessary to average over all the electrons on the Fermi surface, and hence over all different values of k_z. This can be achieved by averaging over all possible values of x assuming that x is uniformly distributed between $-\hbar\omega_c/2$ and $\hbar\omega_c/2$. Hence

$$\langle x \rangle = \int_{-\hbar\omega_c/2}^{\hbar\omega_c/2} x \, \mathrm{d}x = 0 \tag{7.47}$$

and

$$\langle x^2 \rangle = \int_{-\hbar\omega_c/2}^{\hbar\omega_c/2} x^2 \, \mathrm{d}x = \frac{\hbar\omega_c^2}{12} \tag{7.48}$$

This leads to an average energy of the top block (the one labelled E_l) associated with motion perpendicular to the field equal to

$$\frac{p}{2}\left[E_\perp - \frac{\hbar\omega_c}{3}\right] \tag{7.49}$$

so that the total average energy is (using eqn 7.43)

$$\frac{p}{2}\left[E_F - \frac{\hbar\omega_c}{3}\right] \tag{7.50}$$

With a non-zero magnetic field, the top block is full of p electrons with energy E_l if $0 < x < \hbar\omega_c/2$ and the block is totally empty if $-\hbar\omega_c/2 < x < 0$. (Note that this calculation is for a

particular value of k_z so that states are redistributed amongst those for different values of k_z.) Thus the average energy of the top block is

$$E = \begin{cases} p\left(E_\perp - x\right) & \text{if } x > 0 \\ 0 & \text{if } x < 0 \end{cases} \tag{7.51}$$

so that the mean energy associated with motion perpendicular to the field is

$$\frac{p}{2}\left[E_\perp - \frac{\hbar\omega_\mathrm{c}}{4}\right] \tag{7.52}$$

and hence the total average energy (again using eqn 7.43) is

$$\frac{p}{2}\left[E_\mathrm{F} - \frac{\hbar\omega_\mathrm{c}}{4}\right] \tag{7.53}$$

Including a factor $k_\mathrm{F}L/\pi$ for the degeneracy of the states in the k_z direction, the difference in total energy for $B \neq 0$ and $B = 0$ is

$$\begin{aligned} \Delta U &= U(B \neq 0) - U(B = 0) && (7.54)\\ &= \frac{pk_\mathrm{F}L}{2\pi}\left[E_\mathrm{F} - \frac{\hbar\omega_\mathrm{c}}{4}\right] - \frac{pk_\mathrm{F}L}{2\pi}\left[E_\mathrm{F} - \frac{\hbar\omega_\mathrm{c}}{3}\right] && (7.55)\\ &= \frac{pk_\mathrm{F}L\hbar\omega_\mathrm{c}}{24\pi} && (7.56) \end{aligned}$$

Using the expression for the degeneracy p (eqn 7.41) we have that

$$\Delta U = \frac{Vk_\mathrm{F}e^2B^2}{24\pi^2 m} \tag{7.57}$$

Hence the magnetic susceptibility is given by

$$\begin{aligned} \chi_\mathrm{L} &= -\frac{\mu_0}{V}\frac{\partial^2 \Delta U}{\partial B^2} && (7.58)\\ &= -\frac{\mu_0 k_\mathrm{F} e^2}{12\pi^2 m_\mathrm{e}} && (7.59)\\ &= -\frac{1}{3}\mu_0\left(\frac{e\hbar^2}{2m_\mathrm{e}}\right)^2\left[\frac{m_\mathrm{e}k_\mathrm{F}}{\pi^2\hbar^2}\right] && (7.60)\\ &= -\frac{1}{3}\mu_0\mu_\mathrm{B}^2 g\left(E_\mathrm{F}\right) && (7.61) \end{aligned}$$

This is the Landau diamagnetism and it is related to the Pauli paramagnetism χ_P by

$$\chi_\mathrm{L} = -\frac{\chi_\mathrm{P}}{3} \tag{7.62}$$

We might therefore conclude that all metals will be paramagnetic since the (positive) Pauli paramagnetism is three times larger than the (negative) Landau diamagnetism. This would be rather rash, because we have assumed a free electron model and have therefore ignored band structure effects. If the electron has an effective mass m^*, then the density of states at the Fermi energy $g\left(E_\mathrm{F}\right)$ will be enhanced by a factor m^*/m_e. This therefore enhances both the Pauli paramagnetism and the Landau diamagnetism by m^*/m_e. However, in eqn 7.61 we have used $\mu_\mathrm{B} = e\hbar/2m_\mathrm{e}$ to obtain the final result. A Bohr magneton is a fixed constant, independent

of effective mass. This implies that the Landau diamagnetism must be further multiplied by $(m_e/m^*)^2$. The net result is that

$$\chi_L = -\left(\frac{m_e}{m^*}\right)^2 \frac{\chi_P}{3} \tag{7.63}$$

and the total susceptibility of the metal will be

$$\chi = \chi_P \left[1 - \frac{1}{3}\left(\frac{m_e}{m^*}\right)^2\right] \tag{7.64}$$

For most metals $m^* \sim m_e$ and they are therefore paramagnetic. If $m^* < m_e/\sqrt{3}$, the net susceptibility becomes negative. This explains the strong diamagnetism of bismuth, where $m^* \sim 0.01 m_e$.

A further effect which we have so far ignored is the diamagnetism of the bound electrons in the ion cores. This effect gets larger as the atomic number increases (because there are more electrons to contribute) but it is still usually much smaller than the effects associated with the conduction electrons which we have been discussing.

The electron motion on the Fermi surface can be described by the equation

$$\hbar\dot{\mathbf{k}} = -e\mathbf{v} \times \mathbf{B} \tag{7.65}$$

where $\mathbf{k}$ is the electron wave vector on the Fermi surface, and $\mathbf{v} = \hbar^{-1}\partial E(\mathbf{K})/\partial \mathbf{k}$ is the electron velocity, and is always perpendicular to the

Fermi surface. Hence the electron motion is perpendicular to $\mathbf{B}$ but stays on the Fermi surface. [6] Diamagnetic orbital contributions to the magnetic susceptibility arise from closed orbits on the Fermi surface. As shown in Fig. 7.9, open orbits are also possible. In the simple free electron sphere, only closed orbits are possible. Both open and closed orbits are possible in real materials.

Interesting oscillatory changes in properties with magnetic field occur when the Landau levels break through the Fermi surface as the magnetic field is increased. Properties change because the density of states at the Fermi energy can oscillate as a function of magnetic field. This effect is maximized when a Landau tube crosses an extremal cross-section of the Fermi surface (for the Fermi sphere, this would be an equatorial cross-section). The oscillation of the density of states leads to an oscillatory magnetization and this is the de Haas-van Alphen effect and can be used to measure the cross-sectional area of the Fermi surface. There is an analogous effect on the electrical resistance known as the Shubnikov-de Haas effect. Both effects give rise to oscillations in properties as a function of $1/B$. The period of these oscillations gives the area A_{ext} of the maximal and minimal cross-sectional area of the Fermi surface normal to the magnetic field according to

$$\Delta\left(\frac{1}{B}\right) = \frac{2\pi e}{\hbar A_{ext}}. \tag{7.66}$$

9.7.7 7.7 Magnetism of the electron gas

The Pauli paramagnetism effect describes the paramagnetic response of an electron gas to an applied magnetic field. But what if the applied magnetic field is spatially varying? The electron gas will respond to this perturbation according to the wave vector-dependent susceptibility, which we now calculate. (The diamagnetic response will be ignored for the moment.)

7.7.1 Paramagnetic response of the electron gas

A spatially varying magnetic field

$$\mathbf{H}(\mathbf{r}) = \mathbf{H}_\mathbf{q} \cos \mathbf{q} \cdot \mathbf{r} \tag{7.67}$$

provides a perturbation $\hat{\mathcal{H}}'$ equal to

$$\hat{\mathcal{H}}' = \pm \frac{g\mu_0\mu_\mathrm{B}}{2} |\mathbf{H}_\mathbf{q}| \cos \mathbf{q} \cdot \mathbf{r} \tag{7.68}$$

where the $\pm$ refers to the electron spin. A plane wave state

$$\psi_{\mathbf{k}\pm}(\mathbf{r}) = \frac{1}{\sqrt{V}} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} \Big| \pm \Big) \tag{7.69}$$

is weakly perturbed into states $\psi_{(\mathbf{k}+\mathbf{q})}(\mathbf{r})$ and $\psi_{(\mathbf{k}-\mathbf{q})}(\mathbf{r})$. The amplitude of these states can be calculated using first-order perturbation theory (see Appendix C). The perturbed wave function is therefore

$$\psi_{\mathbf{k}\pm}(\mathbf{r}) = \frac{1}{\sqrt{V}} \left(\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}} \pm \frac{g\mu_0\mu_\mathrm{B} H_\mathbf{q}}{4} \left[\frac{\mathrm{e}^{\mathrm{i}(\mathbf{k}+\mathbf{q})\cdot\mathbf{r}}}{E_{\mathbf{k}+\mathbf{q}} - E_\mathbf{k}} + \frac{\mathrm{e}^{\mathrm{i}(\mathbf{k}-\mathbf{q})\cdot\mathbf{r}}}{E_{\mathbf{k}-\mathbf{q}} - E_\mathbf{k}} \right] \right) |\pm\rangle \tag{7.70}$$

[6] This is because in a time dt, the change in k is $\mathrm{d}\mathbf{k}\alpha(\partial E(\mathbf{K})/\partial\mathbf{k}) \times \mathbf{B}$ and hence is both normal to $\mathbf{B}$ and in the plane of the Fermi surface.

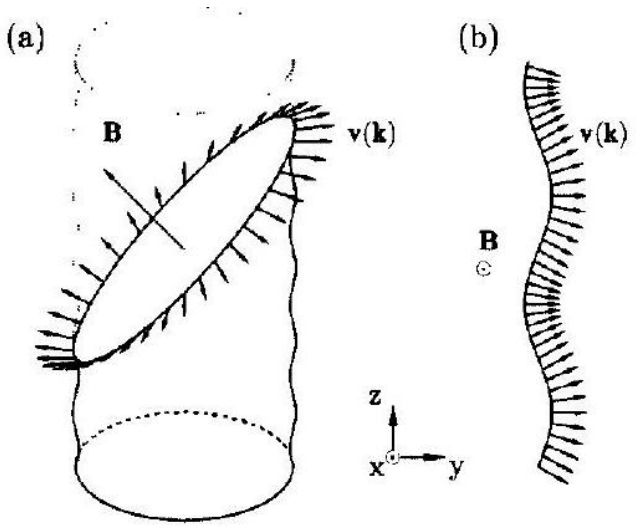


Fig. 7.9 Examples of (a) a closed orbit and (b) an open orbit. The vectors on the orbits show the real-space velocity of the electron, given by $\mathbf{v} = \hbar^{-1} \partial E(\mathbf{K}) / \partial \mathbf{k}$, which is a vector normal to the Fermi surface. With the magnetic field $\mathbf{B}$ oriented as in (a), a closed orbit results. With $\mathbf{B}$ along the x direction, an open orbit on the same Fermi surface can result, as shown in (b).}

P. M. van Alphen (1906-1967)
W. J. de Haas (1878-1960)
L. V. Shubnikov (1901-1945)

where $E_\mathbf{k} = \hbar^2 k^2 / 2m_\mathrm{e}$. Keeping only leading orders of $H_\mathbf{q}$ this implies that

$$|\psi_{\mathbf{k}\pm}(\mathbf{r})|^2 = \frac{1}{V}\left(1 \pm \frac{g\mu_0\mu_\mathrm{B}H_\mathbf{q}m_\mathrm{e}}{\hbar^2} \times \left[\frac{1}{(\mathbf{k}+\mathbf{q})^2 - k^2} + \frac{1}{(\mathbf{k}-\mathbf{q})^2 - k^2}\right]\cos\mathbf{q}\cdot\mathbf{r}\right). \tag{7.71}$$

The magnetization $M(\mathbf{r})$ due to this perturbed wave function is then

$$\begin{aligned} M(\mathbf{r}) &= \frac{g\mu_0\mu_\mathrm{B}}{2}\left(|\psi_{k+}(\mathbf{r})|^2 - |\psi_{k-}(\mathbf{r})|^2\right) \\ &= \frac{g^2\mu_0\mu_\mathrm{B}^2 m_\mathrm{e}H_\mathrm{q}\cos\mathbf{q}\cdot\mathbf{r}}{\hbar^2 V}\left[\frac{1}{(\mathbf{k}+\mathbf{q})^2 - k^2} + \frac{1}{(\mathbf{k}-\mathbf{q})^2 - k^2}\right] \end{aligned} \tag{7.72}$$

To calculate the magnetic response of the electron gas as a whole, it is necessary to sum the expression in eqn 7.72 over all the electrons in the Fermi sphere. Hence the spatially varying magnetization of the electron gas $M(\mathbf{r})$ becomes

$$M(\mathbf{r}) = M_\mathrm{q}\cos\mathbf{q}\cdot\mathbf{r} \tag{7.73}$$

where $M_\mathbf{q}$ is given by

$$\begin{aligned} M_\mathbf{q} &= \frac{g^2\mu_0\mu_\mathrm{B}^2 m_\mathrm{e}H_\mathrm{q}}{\hbar^2 V}\int_{|\mathbf{k}|<k_\mathrm{F}} g(\mathbf{k})\mathrm{d}^3\mathbf{k}\left[\frac{1}{(\mathbf{k}+\mathbf{q})^2 - k^2} + \frac{1}{(\mathbf{k}-\mathbf{q})^2 - k^2}\right] \\ &= \frac{k_\mathrm{F}m_\mathrm{e}g^2\mu_0\mu_\mathrm{B}^2 H_\mathrm{q}}{\pi^2\hbar^2}\left[1 + \frac{4k_\mathrm{F}^2 - q^2}{4k_\mathrm{F}q}\log\left|\frac{q+2k_\mathrm{F}}{q-2k_\mathrm{F}}\right|\right] \end{aligned} \tag{7.74}$$

where the density of states $g(|k|) = Vk^2/\pi^2$ and use has been made of the integral

$$\begin{aligned} \int_{|\mathbf{k}|<k_\mathrm{F}} \frac{\mathrm{d}^3\mathbf{k}}{(\mathbf{k}+\mathbf{q})^2 - k^2} &= \int_0^{k_\mathrm{F}} 2\pi k^2\,\mathrm{d}k\int_0^\pi \frac{\sin\theta\mathrm{d}\theta}{q(q-2k\cos\theta)} \\ &= \int_0^{k_\mathrm{F}} \frac{\pi k\,\mathrm{d}k}{q}\log\frac{q+2k}{|q-2k|} \qquad (7.75) \\ &= \frac{\pi k_\mathrm{F}}{2}\left[1 + \frac{4k_\mathrm{F}^2 - q^2}{4k_\mathrm{F}q}\log\left|\frac{q+2k_\mathrm{F}}{q-2k_\mathrm{F}}\right|\right]. \qquad (7.76) \end{aligned}$$

Using the equation

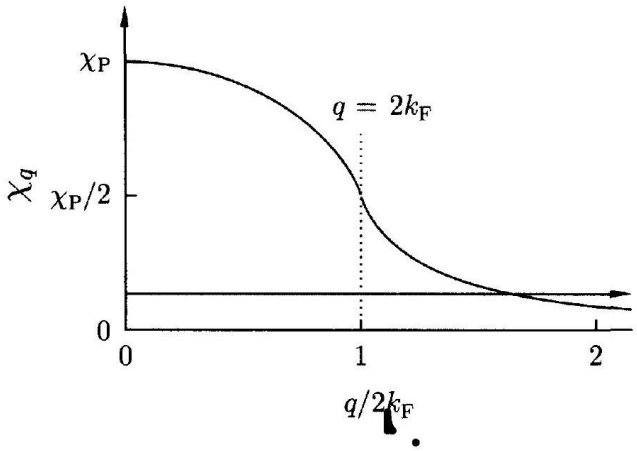
$$\chi_\mathbf{q} = M_\mathbf{q}/H_\mathbf{q} \tag{7.77}$$

which we shall justify below, and the expression for $g\left(E_\mathrm{F}\right)$ in eqn 7.10 , the wave vector-dependent susceptibility $\chi_\mathbf{q}$ is then given by

$$\chi_\mathbf{q} = \chi_\mathrm{P} f\left(\frac{q}{2k_\mathrm{F}}\right) \tag{7.78}$$

where the function $f(x)$ is given by

$$f(x) = \frac{1}{2}\left(1 + \frac{1-x^2}{2x}\log\left|\frac{x+1}{x-1}\right|\right) \tag{7.79}$$



$\chi_{\mathbf{q}}$ is plotted in Fig. 7.10. At $q = 0, f(0) = 1$ so that $\chi_{\mathbf{q}=0} = \chi_P$ and so this result happily reduces to the expression for Pauli paramagnetism derived in Section 7.2. The kink in the curve which appears at $q = 2k_F$ is due to the existence of the Fermi surface (the diameter of the Fermi sphere is $2k_F$). This kink becomes progressively more severe as the dimensionality is reduced down from three dimensions, as will be discussed later in the chapter.

In general the q-dependent magnetic field $\mathbf{H_q}$ can be defined by a Fourier transform of $\mathbf{H}(\mathbf{r})$ by

$$\mathbf{H_q} = \int \mathrm{d}^3\mathbf{r}\mathbf{H}(\mathbf{r})\mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\mathbf{r}}, \quad \mathbf{H}(\mathbf{r}) = \frac{1}{(2\pi)^3}\int \mathrm{d}^3\mathbf{q}\mathbf{H_q}\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}} \tag{7.80}$$

In the same way the q-dependent magnetization $\mathbf{M_q}$ is defined by

$$\mathbf{M_q} = \int \mathrm{d}^3\mathbf{r}\mathbf{M}(\mathbf{r})\mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\mathbf{r}}, \quad \mathbf{M}(\mathbf{r}) = \frac{1}{(2\pi)^3}\int \mathrm{d}^3\mathbf{q}\mathbf{M}\mathbf{q}^{\mathrm{e}}\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}} \tag{7.81}$$

and the q-dependent susceptibility $\chi_{\mathbf{q}}$ is similarly

$$\chi_{\mathbf{q}} = \int \mathrm{d}^3\mathbf{r}\chi(\mathbf{r})\mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\mathbf{r}}, \quad \chi(\mathbf{r}) = \frac{1}{(2\pi)^3}\int \mathrm{d}^3\mathbf{q}\chi_{\mathbf{q}}\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}} \tag{7.82}$$

where for each of eqns 7.80, 7.81 and 7.82 the inverse transformation is also shown. The q-dependent susceptibility provides a means of calculating the magnetization of the electron gas. The magnetization $\mathbf{M}(\mathbf{r})$ at any point $\mathbf{r}$ in the electron gas is related to the magnetic field $\mathbf{H}(\mathbf{r})$ by

$$\mathbf{M}(\mathbf{r}) = \int \mathrm{d}^3\mathbf{r}'\chi\left(\mathbf{r}-\mathbf{r}'\right)\mathbf{H}\left(\mathbf{r}'\right) \tag{7.83}$$

which is a convolution. Thus $\mathbf{M}(\mathbf{r})$ responds not just to the field at $\mathbf{r}$, but to a weighted average over nearby values. The convolution theorem implies that the relationship between the corresponding q-dependent quantities is

$$\mathbf{M_q} = \chi_{\mathbf{q}}\mathbf{H_q} \tag{7.84}$$

which is the same as eqn 7.77 which we used above.

Fig. 7.10 The q-dependent paramagnetic susceptibility. At $q = 0$ it takes the value of the Pauli paramagnetic susceptibility.

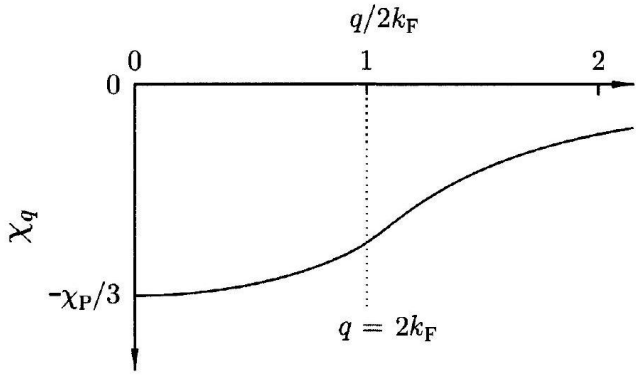
7.7.2 Diamagnetic response of the electron gas

The q-dependent diamagnetic susceptibility can also be calculated, although this is considerably more complicated. Only the result is quoted here, and it is

$$\chi_{\mathbf{q}} = \chi_{\mathrm{L}} \frac{3k_{\mathrm{F}}^2}{2q^2} \left[1 + \frac{4k_{\mathrm{F}}^2}{q^2} - \frac{k_{\mathrm{F}}}{q} \left(1 - \frac{q^2}{4k_{\mathrm{F}}^2} \right)^2 \log \left| \frac{q + 2k_{\mathrm{F}}}{q - 2k_{\mathrm{F}}} \right| \right] \tag{7.85}$$

The diamagnetic susceptibility $\chi_{\mathbf{q}}$ is plotted in Fig. 7.11. At $q = 0$, the expression reduces to $\chi_{\mathbf{q}=0} = \chi_{\mathbf{L}} = -\chi_{\mathbf{P}}/3$. Expression 7.85 also has a singularity at $q = 2k_{\mathrm{F}}$ because of the existence of the Fermi surface.

Fig. 7.11 The q-dependent diamagnetic susceptibility. At $q = 0$ it takes the value of the Landau diamagnetic susceptibility.



7.7.3 The RKKY interaction

The q-dependent paramagnetic susceptibility has consequences for the real space behaviour of a system with a delta-function perturbation. A deltafunction perturbation $\mathbf{H}(\mathbf{r}) = \delta(\mathbf{r})\mathbf{H}$ can be decomposed into a sum of all possible spatial frequencies

$$\mathbf{H}(\mathbf{r}) = \delta(\mathbf{r})\mathbf{H} = \frac{1}{(2\pi)^3} \int \mathbf{H}_{\mathbf{q}} \mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}} \, \mathrm{d}^3 q \tag{7.86}$$

where $\mathbf{H}_{\mathbf{q}} = \mathbf{H}$. Thus a delta function can be considered as a sum over all frequencies, comprising both very long wavelength and short wavelength oscillations. If $\chi_{\mathbf{q}} = \chi$ is independent of q, then $\mathbf{M}_{\mathbf{q}} = \chi \mathbf{H}_{\mathbf{q}}$ and

$$\mathbf{M} = \frac{1}{(2\pi)^3} \int \mathbf{M}_{\mathbf{q}} \mathbf{e}^{\mathbf{i}\mathbf{q}\cdot\mathbf{r}} \mathbf{d}^3 q = \chi \delta(\mathbf{r}) \mathbf{H} \tag{7.87}$$

in other words the magnetization is also a delta function and the electron gas responds completely to the perturbation, shadowing its behaviour perfectly. (In this case $\chi(\mathbf{r} - \mathbf{r}') = \chi\delta(\mathbf{r} - \mathbf{r}')$.)

However the susceptibility of the electron gas is a q-dependent. Because of the presence of the Fermi surface, the electron gas therefore becomes much less responsive for wave vectors above $q = 2k_{\mathrm{F}}$. Thus although it can follow the long wavelength components of the delta-function perturbation, the short wavelengths (that is wavelengths below about π/k_{F}) get attenuated because the
electron gas is less able to respond to fields with such large spatial frequencies. This produces a spatial 'ringing' effect in the magnetization which is analogous to the oscillations observed in the diffraction pattern from a slit. The resulting magnetization has spatial oscillations.

This effect can be calculated as follows. The real-space susceptibility is given by

$$\begin{aligned}\chi(\mathbf{r}) &= \frac{1}{(2\pi)^3}\int \mathrm{d}^3\mathbf{q}\chi_{\mathbf{q}}\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}} \\ &= \frac{1}{(2\pi)^3}\int \mathrm{d}^3\mathbf{q}\frac{\chi_{\mathrm{P}}}{2}\left(1+\frac{4k_{\mathrm{F}}^2-q^2}{4k_{\mathrm{F}}q}\log\left|\frac{q+2k_{\mathrm{F}}}{q-2k_{\mathrm{F}}}\right|\right)\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}} \\ &= \frac{2k_{\mathrm{F}}^3\chi_{\mathrm{P}}}{\pi}F\left(2k_{\mathrm{F}}r\right) \end{aligned} \tag{7.88}$$

where $r = |\mathbf{r}|$ and the function $F(x)$ is given by

$$F(x) = \frac{-x\cos x + \sin x}{x^4} \tag{7.89}$$

and is plotted in Fig. 7.12. This is known as the RKKY interaction (which was presented earlier in Section 4.2.4). The integral in eqn 7.88 is far from trivial and is examined in more detail in Exercise 7.5. The susceptibility (and hence the magnetization due to a delta function perturbation) is proportional to $\cos\left(2k_{\mathrm{F}}r\right)/r^3$ at large distances $\left(r \gg k_{\mathrm{F}}^{-1}\right)$ and is oscillatory. At small r the susceptibility diverges, which is a consequence of the delta function perturbation which we have been assuming. The RKKY interaction was first proposed in order to understand the effect of a localized nuclear moment on the electron gas; the nuclear moment is pretty much as close as one can get to a delta function perturbation, but is not infinitely localized, but spread over the (albeit tiny) nuclear volume. The RKKY interaction provides an important mechanism for magnetic coupling between localized electronic moments in metals; one moment produces an oscillatory magnetization of the electron gas which can interact with a second moment. The interaction can be ferromagnetic or antiferromagnetic depending on the distance between the two moments.

A similar phenomenon is observed in the non-magnetic case, i.e. the response of the electron gas to a delta-function charge. This is of interest in the study of localized charges in metallic alloys. Oscillations in the charge density are produced by a similar mechanism and are known as Friedel oscillations.

9.7.8 7.8 Excitations in the electron gas

The electron gas also contains a rich spectrum of excitations. These can be treated elegantly by calculating $\chi(\mathbf{q},\omega)$, the q - and ω-dependent susceptibility, but we shall not proceed with this approach in this book.

In an itinerant ferromagnet in which the bands are spin-split in the absence of a magnetic field, there are two types of excitations. First there are the familiar spin-wave excitations which follow a conventional dispersion relation. Our treatment of spin-waves in Section 6.6 was based on the Heisenberg model

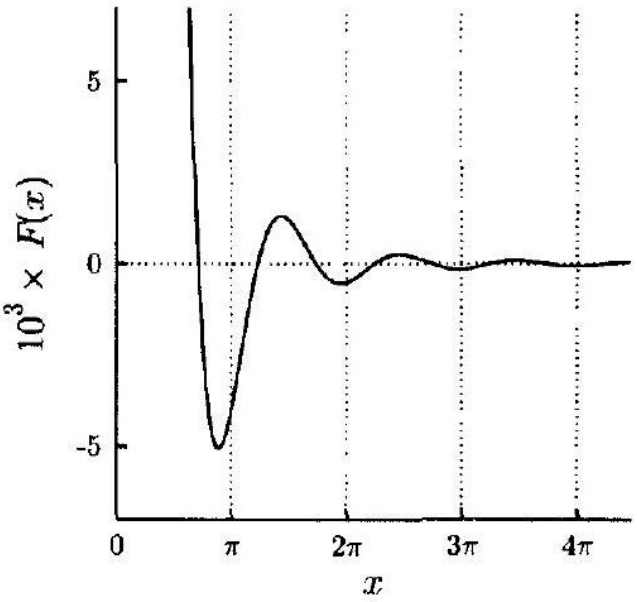


Fig. 7.12 The function $F(x)$ of eqn 7.89 which describes the real space magnetization produced by a delta function field.

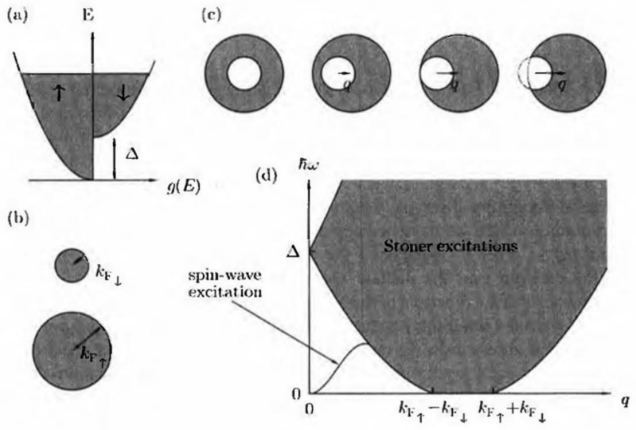
and is therefore applicable to localized moment systems, but it turns out that spin waves can also be derived for metallic systems. Second, there is also a continuum of electron-hole excitations in which an electron transfers from a filled state in one of the spin-split bands to an empty state in the other spinsplit band. These latter excitations are known as Stoner excitations. In a Stoner excitation an electron with wave vector $\mathbf{k}+\mathbf{q}$ and spin down is excited to a state with wave vector $\mathbf{k}$ and spin up. The energy of the excitation is given by

$$\hbar\omega = E_{\mathbf{k}+\mathbf{q}} - E_{\mathbf{k}} + \Delta. \tag{7.90}$$

where $E_{\mathbf{k}} = \hbar^2 k^2/2m_c$ and Δ is the exchange splitting, the energy cost to flip a spin. These results are illustrated in Fig. 7,13.

Fig. 7.13 Excitations in an electron gas. (a) The spin-split bands are separated by Δ, the exchange splitting. (b) The Fermi surfaces for spin-up and spin-duwn electrons. (c) A Stoner excitation can be created if there is a filled $\uparrow$ state with wave vector k and an $\downarrow$ empty state with wave vector $\mathbf{k

$+\mathbf{4}$. The shaded area shows the possible choices of $\mathbf{k}$ for different values of q. (d) The dispersion relation (a graph of ω against q, shows a spin-wave branch and a continuum of Stoner excitations.}



In paramagnetic metals the two spin-split bands are not spin-split in the absence of a magnetic field. The spin-wave excitations have a short lifetime and thus are very heavily damped (they are in the overdamped limit, if considered as damped harmonic oscillators). They are known as poramagnons and, like conventional spin-wave excitations. they can be studied by inelastic neutron scattering.

In this book we have been working with two entirely distinct pictures: insulating materials in which the magnetism is associated with localized moments on atoms and metals in which the magnetism is associated with entirely delocalized moments. Many real materials are somewhere in between, with magnetism associated with spin density fluctuations intermediate between the localized moment and band ferromagnet regimes. A local moment fluctuation is localized in real space while a spin fluctuation in a weakly ferromagnetic metal may be regarded as being localized in q-space.

9.7.9 7.9 Spin-density waves

The Stoner criterion demonstrates that the electron gas is unstable against spontaneous ferromagnetism if $Ug(E_\mathrm{F})$ becomes large enough (see Section 7.3) and that the magnetic susceptibility can be enhanced by a factor $\left(1-Ug(E_\mathrm{F})^{-1}\right)$. It turns out that the q-dependent susceptibility, which we will write as $\chi^0_\mathbf{q}=\chi_\mathrm{P}f(q/2k_\mathrm{F})$ without the presence of Coulomb interactions, is also enhanced by the Coulomb interactions and becomes

$$\chi_\mathbf{q}=\frac{\chi_\mathrm{P}f(q/2k_\mathrm{F})}{1-Ug(E_\mathrm{F})f(q/2k_\mathrm{F})}=\frac{\chi^0_\mathbf{q}}{1-\alpha\chi^0_\mathbf{q}} \tag{7.91}$$

where $\alpha=U/\mu_0\mu_\mathbf{B}^2$. It may occur that $\chi^0_\mathbf{q}$ has a maximum at some value of $\mathbf{q}$ not equal to zero. In this case the interactions, parametrized by α, can make the susceptibility diverge at this value of q if $\alpha\chi^0_q$ reaches unity. Hence an oscillatory static magnetization could spontaneously develop in the sample. If $\mathbf{q}=0$ this would correspond to ferromagnetic order; if $|\mathbf{q}|=\pi/a$ then antiferromagnetic order could develop. In general, one would expect spiral structures or spin-density wave structures with wave vector $\mathbf{q}$.

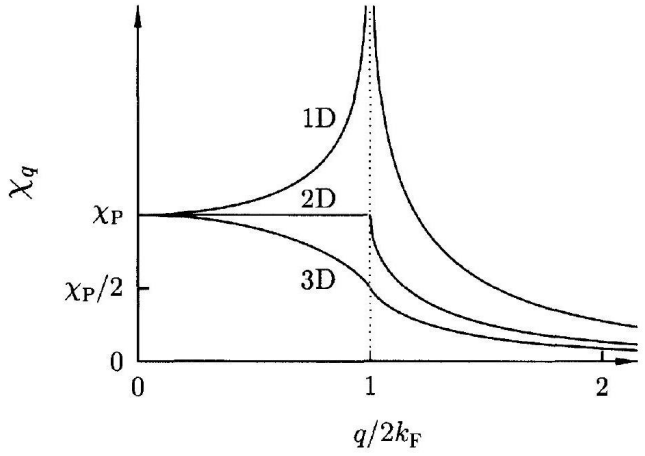


Fig. 7.14 The q-dependent paramagnetic susceptibility in one, two and three dimensions. The response function diverges at $q=2 k_{\mathrm{F}}$ in one-dimension. At $q = 0$ it takes the value of the Pauli paramagnetic susceptibility in all dimensions.}

We have evaluated the q-dependent susceptibility for an electron gas in three dimensions, but the calculation can be repeated (see Exercise 7.6) for two dimensions and one dimension. The results are shown in Fig. 7.14 and demonstrate that as you lower the dimension, the knee in the curve for three dimensions at $q = 2k_{\mathrm{F}}$ becomes a kink in two dimensions and a singularity in one-dimension. The peak in one-dimension is known as a Kohn anomaly and shows that the one-dimensional electron gas is unstable to the formation of spin-density waves with wave vector $2k_{\mathrm{F}}$. This instability is associated with the fact that a periodic modulation of the magnetization with wavelength $2\pi/q$ and wave vector q opens up a gap in the energy dispersion at wave vectors

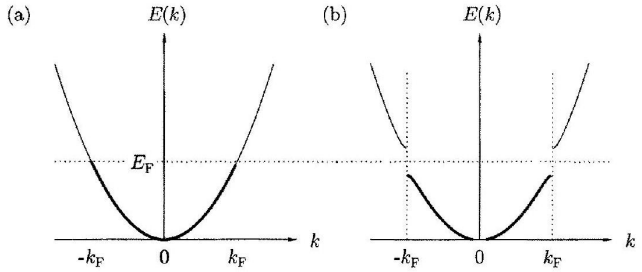
Rudolf E. Peierls (1907-1995)

Fig. 7.16 (a) Commensurate spin-density wave, with wave vector $q = \pi/a$. (b) Incommensurate spin-density wave.

$\pm q/2$ (any periodic potential can do this). If $q/2 = k_{\mathrm{F}}$ the gap is at the Fermi surface and this can lead to a lowering of the total energy. This opportunity to lower the electronic energy can therefore itself drive the formation of the spin-density wave.

There is an analogous instability for the formation of charge-density waves with wave vector $2k_{\mathrm{F}}$. Both effects are connected with the well-known Peierls instability in which the dimerization of a one-dimensional chain of atoms can occur spontaneously because the electronic energy saved by opening up a gap in the energy spectrum near the Fermi surface can outweigh the elastic energy cost of the dimerization. This is illustrated in Fig. 7.15.

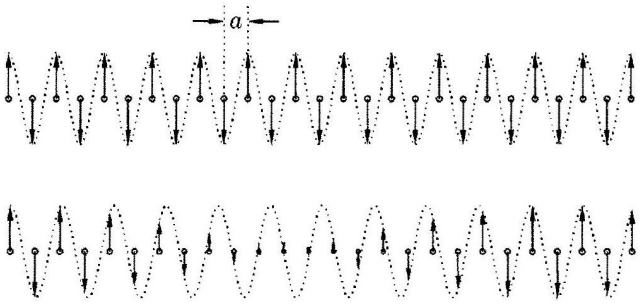
Fig. 7.15 (a) Electrons in a band fill up to the Fermi energy $E_{\mathrm{F
}}$. (b) The energy can be lowered by a periodic potential of wave vector $q = 2k_{\mathrm{F}}$ which lowers the total electronic energy by producing a gap at the Fermi surface. The gapping of the Fermi surface (i.e. going from (a) to (b)) is accompanied by a metal-insulator transition. The corresponding real space distortion is known as a Peierls distortion. A one-dimensional electron gas is unstable with respect to this Peierls instability.}



In three dimensions a spin-density wave cannot produce an energy gap at all points on the Fermi surface. Sometimes the Fermi surface is such that translation of one part of the Fermi surface by a vector **q** can place it on top of another part of the Fermi surface. This phenomenon is known as nesting. A spin-density wave, with wave vector **q**, can produce energy gaps along the region of the Fermi surface for which this nesting is possible. In real metals it very often happens that two pieces of Fermi surface are approximately translated from one another in k-space by a fixed wave vector **q**. This can give rise to a peak in the susceptibility and a resulting instability. The formation of the density wave is said to nest the Fermi surface.

If the nesting wave vector q turns out to be π/a where a is the spacing between atoms, the spin-density wave is commensurate with the lattice and antiferromagnetic order results (Fig. 7.16(a)). However it is much more usual

(a)



(b)

for the nesting wave vector **q**, which is equal to $2k_{\mathrm{F}}$, not to be a simple multiple of π/a so that the spin-density wave is incommensurate (Fig. 7.16(b)).

Chromium has a body-centred cubic structure, with the side of the cube equal to a, and develops a spin-density wave structure below $T_{\mathrm{N}} = 310$ K. The wave vector $\mathbf{q} \approx 0.96(2\pi/a)$ (thus the material is 'almost' an antiferromagnet) varies slowly with temperature and is in the [100] direction. The polarization is longitudinal for $T < 115$ K and transverse for $T > 115$ K. The electrical resistivity of Cr increases just below T_{N} due to the energy gaps introduced into parts of the Fermi surface by the spin-density wave which reduce the number of carriers. Spin density waves are also found in a number of organic metals.

9.7.10 7.10 The Kondo effect

In very dilute alloys of magnetic ions in a non-magnetic host, the magnetic moments of the magnetic ions can be considered as independent if we ignore the RKKY coupling. The only

remaining interaction is between the magnetic moment and the conduction electron spins. At high temperatures the magnetic moments behave like free, paramagnetic moments but below a characteristic temperature, known as the Kondo temperature T_{K}, the interaction between the magnetic moment and the conduction electrons leads to the impurity spin becoming non-magnetic. What is happening is that the conduction electrons begin to form a cloud of opposite spin-polarization around the impurity spin resulting in a quasi-bound state. This process of magnetic screening of a magnetic impurity by the conduction electrons is known as the Kondo effect and has two profound experimental consequences. First, the magnetization falls below its free-moment value and the susceptibility therefore falls below the value expected from Curie's law. Second, because the impurity moment is strongly interacting with the conduction electrons, the electron scattering cross-section of the moment is strongly enhanced and there is a new term in the resistivity which is proportional to $\mathrm{J} \ln T$ where $\mathrm{J} < 0$ is the exchange coupling between the local moment and the conduction electrons.

The resistivity of most materials decreases with decreasing temperature. This is because the number of phonons decreases as the material cools, and the resistivity is strongly determined by the number of phonons. Thus it is usually dominated by a term proportional to T at high temperatures and T^5 at lower temperatures. There is also a constant term, determined by the (temperatureindependent) concentration of impurities which becomes important at the lowest temperatures.

The new $J \ln T \propto -\ln T$ (because $J < 0$) term from the Kondo effect kicks in at low temperature and leads to the appearance of a resistivity minimum (see Fig. 7.17), below which the resistivity increases with decreasing temperature.

9.7.11 7.11 The Hubbard model

Metallic behaviour occurs because electrons save kinetic energy by being delocalized over the whole of a crystal. However, in some systems the onsite Coulomb energy (the cost of putting two electrons on the same lattice site) is so strong that electrons cannot move freely through the crystal; the high

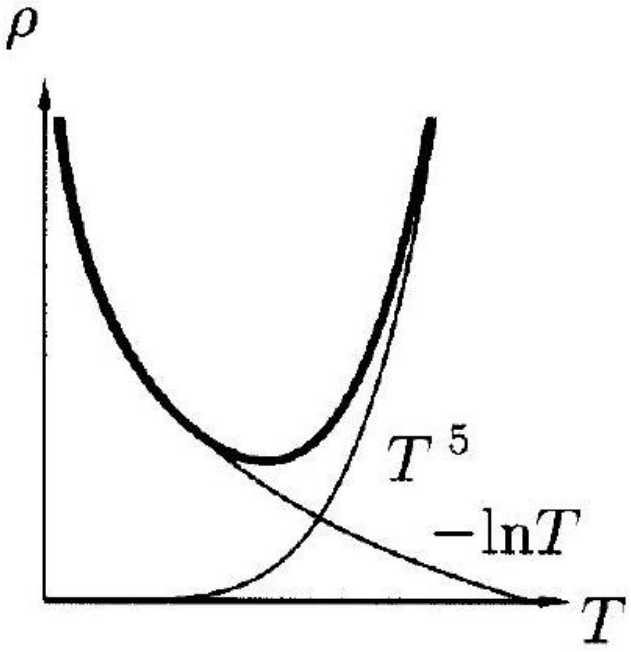


Fig. 7.17 The temperature dependence of the resistivity shows a minimum. The resistivity is the sum of a term proportional to $T^{5} and a term given by $J \ln T \propto -\ln T$.}

energetic cost of double occupancy of a site means that metallic behaviour breaks down. Electrons can no longer be treated as free particles but electron correlations, the features associated with the way electrons have to avoid each other if the Coulomb effects are large, become very significant. The competition between metallic behaviour (parametrized by the transfer integrals in the tight-binding model, or the size of the electronic bandwidth) and Coulomb energy (parametrized by U, sometimes known as the Hubbard-U) is expressed by the Hubbard model.

Each site in a material can be doubly occupied, with one electron having spin-up and one having spin-down. The Pauli exclusion principle precludes double occupancy of a site by electrons with the same spin. In the half-filled limit, where there are just enough electrons to have one electron per lattice site, the on-site Coulomb energy can be minimized if one electron sits on each site and doesn't move. This is the ground state if $U = \infty$. If U is large but not infinite, each site remains singly occupied but it is possible for an electron briefly to visit its neighbour's site if their spins are antiparallel. This saves a bit of kinetic energy and therefore there is an energetic saving for the electrons on neighbouring sites to be antiparallel, hence the system will be an antiferromagnet. However an electron cannot get on to its next-nearest neighbour's site because that will have the same spin and a site cannot be occupied by two electrons of the same spin by the Pauli exclusion principle. Therefore the system is also an insulator. Such an antiferromagnetic insulator, with a large U and one electron per site, is known as a Mott insulator. When U is smaller than the electronic bandwidth W, metallic behaviour is realized. The crossover occurs when $U \sim W$ where there is a metal-insulator transition.

9.7.12 7.12 Neutron stars

We conclude this chapter with a rather extreme example of magnetic condensed matter: the neutron star. Formed during the rapid collapse of a star in a supernova explosion, neutron stars consist of highly dense nuclear matter. Conservation of angular momentum during the stellar

collapse means that any initial rotational rate of the precursor star is greatly increased (as the star collapses, its moment of inertia plummets and so its angular rotation rate shoots up). The radius of a neutron star is typically ~ 10 km and the rotation rate can be up to 10^3 s^{-1}. The collapse also is associated with flux compression, so that the magnetic field of neutron stars is enormous, up to $\sim 10^8$ T. For comparison, the field at the surface of the Sun is $\sim 10^{-4}$ T and at the surface of the Earth is $\sim 5 \times 10^{-5}$ T. The magnetic field is most intense near the poles, but the magnetic field axis of a neutron star, like the Earth's, is at an angle to the rotation axis, so that it precesses as the neutron star rotates. The resulting changing magnetic field leads to a large electric field by Faraday's law and this accelerates electrons along the magnetic axis. The electrons are accelerated along curved trajectories and therefore emit electromagnetic radiation. The result is a tightly concentrated beam of radiation that revolves as the neutron star spins. When the Earth happens to be in the way, we receive regular pulses of radiation. Hence neutron stars were first called pulsars, i.e. pulsating stars.

We can make some (very rough) estimates of the behaviour of neutrons inside the neutron star by using a 'free neutron' model and treating them non-relativistically. Neutrons, like electrons, are fermions and must obey the Pauli exclusion principle. Therefore, as a first-guess model of neutron stars, we can consider the filling up of k-states with neutrons in a box the size of a neutron star following the method of Section 7.1. The neutron star interior contains mainly neutrons at a density about 12 orders of magnitude above that of conduction electrons in a metal. The neutron mass is of course about 3 orders of magnitude higher than the electron mass. Since the Fermi energy is proportional to $n^{2/3}/m$, it should be about 5 orders of magnitude higher in a neutron star than in a metal. Thus the Fermi temperature [7] of the neutrons in a neutron star should be about 10^9 K which is well above the temperature of the neutron star (up to $\sim 10^6$ K) showing that the neutrons in the neutron star are in the degenerate limit. The surface of a neutron star is thought to be solid, but the matter in the interior may possibly be in a superfluid state. This is analogous to the pairing of helium atoms that is found in superfluid helium, but in neutron stars it may be pairs of neutrons, and also pairs of protons, that condense. Neutrons and protons have a pairing energy of $1 - 2$MeV, which is equivalent to a temperature of $\sim 10^{10}$ K, much higher than the temperature of the neutron star, so that pairing can be expected to occur.

[7] By comparison, for the case of electrons in a metal we expect $E_{\rm F} \sim 1$eV, so that $T_{\rm F} \sim 10^4$ K

9.7.13 Further reading

- Landau diamagnetism is discussed in A. Dupré, American Journal of Physics, 49, 34 (1981), and the treatment in this chapter has been adapted from that presented in this article.
- The band theory of metals and the de Haas-van Alphen effect is clearly described in J. Singleton, Electronic properties of solids, OUP 2001.
- A useful treatment of the physics of Landau levels is in Chapter 6 of J. H. Davies, The physics of lowdimensional semiconductors, CUP 1998.
- An advanced treatment of the magnetism of band electrons is in K. Yosida, Theory of magnetism, Springer 1996.
- Spin density functional theory is described in J. Kübler, Theory of itinerant electron magnetism, OUP 2000.
- Useful information on fluctuations in metallic magnets may be found in T. Moriya, Spin fluctuations in itinerant electron magnetism, Springer 1985.

- The Kondo effect is treated in A. C. Hewson, The Kondo problem to heavy fermions, CUP 1993.
- The physics of neutron stars and related astrophysical objects are described in S. L. Shapiro and S. A. Teukolsky, Black holes, white dwarfs and neutron stars, Wiley 1983.

Exercises

(7.1) Show that Pauli paramagnetism leads to a magnetic susceptibility that can be written in a Curie-like form

$$\chi_{\mathrm{P}} = \frac{n\mu_0\mu_{\mathrm{B}}^2}{k_{\mathrm{B}}T_0} \tag{7.92}$$

where the constant T_0 is given by

$$T_0 = \frac{2T_{\mathrm{F}}}{3} \tag{7.93}$$

where $T_{\mathrm{F}} = E_{\mathrm{F}}/k_{\mathrm{B}}$ is the Fermi temperature. Show further that if $T \ll T_{\mathrm{F}}$ the temperature-dependent correction to χ_{P} is

$$\chi_{\mathrm{P}}(T) = \frac{\mu_0\mu_{\mathrm{B}}^2}{3k_{\mathrm{B}}T_0}\left(1 - \frac{\pi^2T^2}{12T_{\mathrm{F}}^2}\right) \tag{7.94}$$

and estimate the size of this correction for a typical metal at room temperature.

(7.2) (a) Show that the paramagnetic susceptibility of a nondegenerate electron gas containing N electrons is identical to that of N independent localized electrons whose orbital motion is quenched.

(b) Consider a semiconductor with 3×10^{22} electrons per cubic metre in its conduction band and an effective mass $m^* = 0.1m_{\mathrm{e}}$. Estimate the temperature below which the magnetic susceptibility is independent of temperature. Below this temperature, calculate the magnitude of the Pauli paramagnetism and the Landau diamagnetism.

(7.3) In the presence of a magnetic field $\mathbf{B} = (0, 0, B)$, the momentum operator $\hat{\mathbf{p}} = -\mathrm{i}\hbar\nabla$ must be replaced by $-\mathrm{i}\hbar\nabla + e$ A leading to the energy eigenvalues in eqn 7.40. In Section 7.5 the magnetic vector potential was chosen to be $\mathbf{A} = (0, Bx, 0)$. Show that the same results could have been obtained if $\mathbf{A} = (-By, 0, 0)$ was used instead. The problem can be solved in cylindrical polars where $A_\phi = \frac{1}{2}Br, A_z = A_r = 0$. Show that this choice yields the same value for $\mathbf{B}$ and for the energy eigenvalues.

(7.4) Show that the Fourier transform of the Fermi sphere is related to a function related to the RKKY function in eqn 7.89 , namely that

$$\begin{aligned}\int_{|\mathbf{k}|<k_{\mathrm{F}}} \mathrm{d}^3\mathbf{k}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} &= \int_0^{k_{\mathrm{F}}} 2\pi k^2\ \mathrm{d}k \int_0^{\pi} \mathrm{e}^{\mathrm{i}kr\cos\theta}\sin\theta\ \mathrm{d}\theta \\ &= \frac{4\pi}{r^3}\left[\sin k_{\mathrm{F}}r - k_{\mathrm{F}}r\cos k_{\mathrm{F}}r\right]\end{aligned} \tag{7.95}$$

(7.5) The integral in eqn 7.88 is a bit complicated to evaluate and this problem is designed to take it a stage at a time. First show that

$$\int \mathrm{d}^3\mathbf{q}\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}}\varphi(\mathbf{q}) = \frac{2\pi}{\mathrm{i}r}\int_{-\infty}^{\infty}\mathrm{d}q q\mathrm{e}^{\mathrm{i}qr}\varphi(q), \tag{7.96}$$

where $\varphi(q) = \varphi(-q)$ is a function of the variable q. Hence show that

$$\begin{aligned}\tfrac{1}{(2\pi)^3}\int \mathrm{d}^3\mathbf{q}\tfrac{\chi\mathbf{P}}{2}\left(1+\tfrac{4k_\mathrm{F}^2-q^2}{4k_\mathrm{F}q}\log\left|\tfrac{q+2k_\mathrm{F}}{q-2k_\mathrm{F}}\right|\right)\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}} \\ =\tfrac{\chi\mathrm{P}k_\mathrm{F}^2}{\mathrm{i}^2 r}\int_{-\infty}^{\infty}\mathrm{d}x x\mathrm{e}^{\mathrm{i}yx}f(x)\end{aligned} \tag{7.97}$$

where $x=q/2k_\mathrm{F}, y=2k_\mathrm{F}r$ and where $f(x)$ is the function defined in eqn 7.79 . To do this integral it is necessary to resort to a trick of contour integration.

Consider the two integrals

$$\begin{aligned}I_1&=\frac{1}{\mathrm{i}r}\int_{-\infty}^{\infty}\mathrm{d}x\frac{x\mathrm{e}^{\mathrm{i}yx}}{2}\left(1+\frac{1-x^2}{2x}\log\left|\frac{x+1}{x-1}\right|\right)\\ I_2&=\frac{1}{\mathrm{i}r}\int_{-\infty}^{\infty}\mathrm{d}x\frac{x\mathrm{e}^{\mathrm{i}yx}}{2}\left(1+\frac{1-x^2}{2x}\log\frac{x+1}{x-1}\right)\end{aligned} \tag{7.99}$$

The two integrands are equal when $|x|>1$. Show that I_2 is zero by completing the contour with an infinite semicircle in the upper half of the x plane and using Jordan's lemma. Hence, using this result and a contour for I_1 that is displaced infinitesimally above the real axis from $x=-1$ to $x=1$, show that

$$\begin{aligned}I_1&=\frac{1}{\mathrm{i}r}\int_{-1}^{1}\mathrm{d}x\frac{\mathrm{e}^{\mathrm{i}yx}\left(1-x^2\right)}{4}\left(\log\left|\frac{x+1}{x-1}\right|-\log\frac{x+1}{x-1}\right)\\ &=\frac{\pi}{r}\int_{-1}^{1}\mathrm{d}x\frac{\mathrm{e}^{\mathrm{i}yx}\left(1-x^2\right)}{4}\\ &=\frac{\pi}{r}\left[\frac{\sin y-y\cos y}{y^3}\right]\\ &=2\pi k_\mathrm{F}F\left(2k_\mathrm{F}r\right)\end{aligned} \tag{7.101}$$

and hence verify eqn 7.88 .

(a) Show that the density of the states at the Fermi energy levels in one, two and three dimensions are given by

$$\begin{array}{rc}\text{one-dimension:} & g\left(E_\mathrm{F}\right)=\frac{n}{2E_\mathrm{F}}=\frac{2m}{\hbar^2\pi k_\mathrm{F}}\\ \text{two-dimensions:} & g\left(E_\mathrm{F}\right)=\frac{n}{E_\mathrm{F}}=\frac{m}{\hbar^2\pi}\\ \text{three-dimensions:} & g\left(E_\mathrm{F}\right)=\frac{3n}{2E_\mathrm{F}}=\frac{mk_\mathrm{F}}{\hbar^2\pi^2}\end{array} \tag{7.102}$$

(b) Show that the q-dependent susceptibility of the electron gas in one dimension is given by

$$\chi_\mathbf{q}=\chi_\mathrm{P}\frac{k_\mathrm{F}}{q}\log\left|\frac{q+2k_\mathrm{F}}{q-2k_\mathrm{F}}\right| \tag{7.103}$$

and in two-dimensions is given by

$$\chi_\mathbf{q}=\begin{cases}\chi_\mathrm{P} & q<2k_\mathrm{F}\\ \chi_\mathrm{P}\left(1-\sqrt{1-\left(4k_\mathrm{F}^2/q^2\right)}\right) & q>2k_\mathrm{F}\end{cases} \tag{7.104}$$

These results are plotted in Fig. 7.14.

(7.7) Show that the continuum of Stoner excitations described in Section 7.8 and pictured in Fig. 7.13 are bounded by the two parabolae

$$\hbar\omega=\Delta+\frac{\hbar^2 q}{2m_\mathrm{e}}\left(q\pm 2k_{\mathrm{F}\uparrow}\right) \tag{7.105}$$

and that the $\omega = 0$ excitations are contained within the range of q given by

$$k_{\mathrm{F}\uparrow} - k_{\mathrm{F}\downarrow} \leq q \leq k_{\mathrm{F}\uparrow} + k_{\mathrm{F}\downarrow} \tag{7.106}$$

as shown in Fig. 7.13 (Hint: use the fact that $\hbar^2 k_{\mathrm{F}\uparrow}^2/2m_{\mathrm{e}} = \hbar^2 k_{\mathrm{F}_\downarrow}^2/2m_{\mathrm{e}} + \Delta.$)

(7.8) Consider the splitting of energy bands which is shown in Fig. 7.5, but now in the case in which δE is not an infinitesimal quantity. The Fermi energy has now become spin-dependent. If the system of electrons has n electrons per unit volume, show that n_+, the number of up electrons per unit volume, and show that n_-, the number of down electrons per unit volume, are given by

$$n_\pm = \int_0^{E_{\mathrm{F}\pm}} g_\pm(E)\mathrm{d}E \tag{7.107}$$

$$= \int_0^{E_{\mathrm{F}\pm}} \frac{3}{4}\frac{n}{E_{\mathrm{F}}^{3/2}} E^{1/2}\ \mathrm{d}E \tag{7.108}$$

$$= \frac{n}{2}\left(\frac{E_{\mathrm{F}\pm}}{E_{\mathrm{F}}}\right)^{3/2} \tag{7.109}$$

where E_{F} is the Fermi energy in the absence of exchange splitting.
Using $n_+ + n_- = n$ and defining $x = (n_+ - n_-)/n$, show that

$$\frac{E_{\mathrm{F}\pm}}{E_{\mathrm{F}}} = (1 \pm x)^{2/3} \tag{7.110}$$

The kinetic energy is given by

$$T = \sum_{\sigma=\pm} \int_0^{E_{\mathrm{F}\sigma}} E g(E)\mathrm{d}E \tag{7.111}$$

Show that

$$T = \frac{3}{10} n E_{\mathrm{F}} \left[(1+x)^{5/3} + (1-x)^{5/3}\right]. \tag{7.112}$$

The interaction energy can be written
$V = -\int_0^M \mu_0 (\lambda M')\,\mathrm{d}M' = -\frac{1}{2}\mu_0 \lambda M^2 = -\frac{1}{2} U n^2 x^2$ (7.113) where $\lambda, U = \mu_0 \mu_{\mathrm{B}}^2 \lambda$ and $M = (n_+ - n_-)^2$ are the molecular field parameter, Coulomb energy and magnetization respectively. The total energy is therefore

$$E = T + V \tag{7.114}$$

Show that the condition $\mathrm{d}E/\mathrm{d}x = 0$ leads to stable solutions when

$$\frac{(1+x)^{2/3} + (1-x)^{2/3}}{x} = \frac{4U g(E_{\mathrm{F}})}{3} \tag{7.115}$$

(Hint: use $g(E_{\mathrm{F}}) = 3n/2E_{\mathrm{F}}$.) Interpret this result graphically.
Explain why the non-magnetic state is unstable when $\mathrm{d}^2E/\mathrm{d}x^2 < 0$ at $x = 0$. Show that this is equivalent to

$$U g(E_{\mathrm{F}}) > 0, \tag{7.116}$$

which is the Stoner criterion.

8.1 Frustration 167

8.2 Spin glasses 168
8.3 Superparamagnetism 171
8.4 One-dimensional magnets 172
8.5 Two-dimensional magnets 177
8.6 Quantum phase transitions 179
8.7 Thin films and multilayers 181
8.8 Magneto-optics 183
8.9 Magnetoresistance 184
8.10 Organic and molecular magnets 192
8.11 Spin electronics 193

9.8 Competing interactions and low dimensionality

In this chapter we will examine some of the ways in which competing interactions and low dimensionality can lead to some extremely subtle, complex, and sometimes even useful magnetic behaviour in solids. Competing interactions and low dimensionality both occur naturally in some systems; for example some crystals grow in such a way that the magnetic moments couple strongly in chains or occupy sites on frustrated lattices. However, these features can be introduced artificially, for example, by fabricating a ferromagnetic multilayer or using electron-beam lithography to produce ferromagnetic wires. Many of these topics are currently under active study in research groups throughout the world. This chapter presents an account of some of these developments, many of which contain mysteries which are not yet fully unravelled. Because many of the topics are new and complex, the discussion will be somewhat oversimplified in some cases; the interested reader can consult the references in the further reading at the end of the chapter for more details. Nevertheless I hope to give some flavour of a few of the problems which are, at the time of writing, providing enormous challenges to condensed matter physicists.

We begin with frustration, which can lead to a number of different ground states including spin glasses, and then continue to the effects of low dimensionality.

9.8.1 8.1 Frustration

In some lattices it is not possible to satisfy all the interactions in the system to find the ground state. Often this leads to there being no single unique ground state but a variety of similar low energy states of the system in which the unhappiness (by which I mean non-minimization of the energy) is shared around as much as possible. In this case the system is said to show frustration. As an example, consider the lattices in Fig. 8.1 in which only nearest-neighbour antiferromagnetic interactions operate. On the square lattice (Fig. 8.1(a)) it is easily possible to satisfy the requirement that nearest-neighbour spins must be antiparallel. However on a triangular lattice things are not so simple. As shown, if two adjacent spins are placed antiparallel, one is faced with a dilemma for the third spin. Whichever choice is made, one of the two neighbours will not have their energy minimized. The system therefore cannot achieve a state that entirely satisfies its microscopic constraints, but does possess a multiplicity of equally unsatisfied states. As a result the frustrated system shows metastability, hysteresis effects (dependence on the sample's magnetic or thermal history), and time-dependent relaxation towards equilibrium.

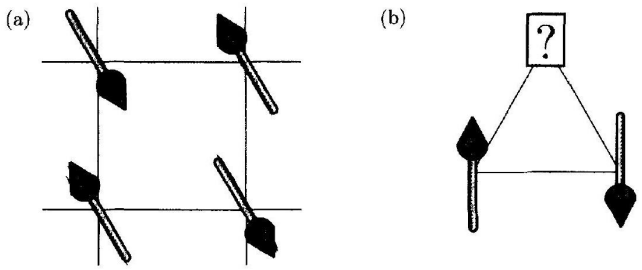


Fig. 8.1 Antiferromagnetic nearestneighbour interactions on the (a) square lattice and (b) triangular lattice. The triangular lattice shows frustration, because it is not possible to orient the spin on the third site to satisfy the requirement of antiferromagnetic nearest-neighbour interactions with the other two spins.

Hence in some systems the geometry of the lattice can frustrate the ordering of the spins. In two dimensions, Heisenberg spins on triangular and corner sharing kagomé lattices (see Fig. 8.2) show this effect, while in three dimensions the most well studied systems have a pyrochlore structure, in which the magnetic ions occupy a lattice of corner sharing tetrahedra. These systems are believed not to order but to display a classical ground state with macroscopic degeneracy, sometimes described as cooperative paramagnetism in which only short-range correlations between spins are found for all temperatures. Some of these systems are also believed to possess a dispersionless spin-wave branch, known as a zero mode, which strongly affects the low temperature thermodynamic behaviour and leads to the persistence of spin fluctuations down to zero temperature.

9.8.2 8.2 Spin glasses

Spin glasses were described earlier in Section 5.5. In that section a spin glass was defined as a random, mixed-interacting magnetic system characterized by a random, yet cooperative, freezing of spins at a well defined temperature T_{f} below which a highly irreversible, metastable frozen state appears without the usual magnetic long range ordering. We now examine the different parts of this definition in much more detail, beginning with the randomness which can be introduced in several ways.

First is site-randomness, which can be achieved in an alloy. A commonly studied spin glass is $Cu_{1-x}Mn_x$ with $x \ll 1$ in which the substitution of small amounts of Mn into the Cu matrix occurs completely randomly with no shortrange ordering. This directly leads to a random distance between magnetic Mn ions in the non-magnetic Cu matrix (see Fig. 5.1(c)). Random distances between magnetic ions can also be engineered by taking an intermetallic such as $GdAl_2$ and making it amorphous (melting it and then cooling very rapidly from the melt) so that the crystalline lattice is destroyed.

Another possibility is bond-randomness in which the nearest neighbour interactions vary between +J and -J . This can be achieved by fixing the magnetic ions in a regular crystalline array but modulating the indirect exchange interactions between the magnetic ions. This is performed in $Rb_2Cu_{1-x}Co_x\ F_4$ for which both Cu and Co have effective spins of $\frac{1}{2}$ (the crystal field on the Co splits the $S = \frac{3}{2}$ levels in such a way that only the lowest $S = \frac{1}{2}$ doublet is

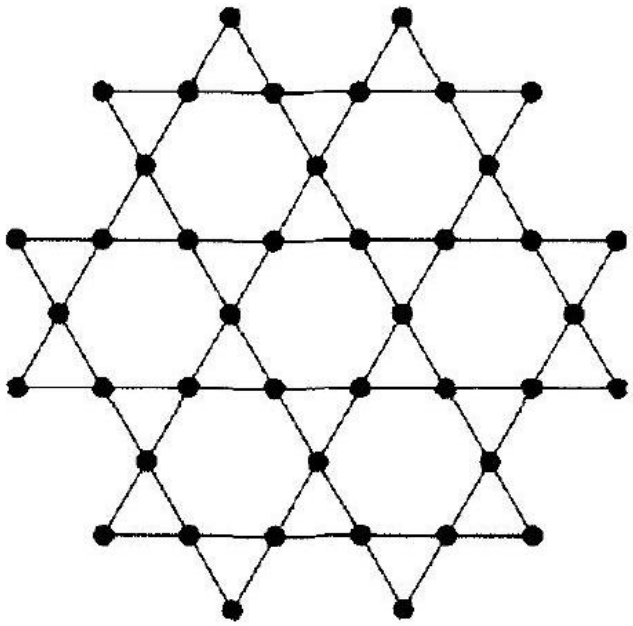


Fig. 8.2 A section of the kagomé lattice.

populated at low temperatures). The crystal field also causes the spins to point either parallel or antiparallel to the z axis (a uniaxial single ion anisotropy, see Section 3.2.2) which is said to give them Ising character. However the size and sign of the superexchange interaction between magnetic ions depends crucially on whether the coupling is between Co and Co , or Co and Cu , or Cu and Cu , and which orbital on the Cu is occupied. The net result is that the bonds between ions can have different values of the exchange J and these bonds are randomly distributed throughout the sample.

The randomness inherent in a spin glass is important, but equally important is the presence of competing interactions. The distribution of distances between moments in a random-site spin glass leads to competing interactions because the interactions are of RKKY-type and therefore their sign (ferromagnetic or antiferromagnetic) depends on the distance between the spins. Another contributing feature is the magnetic anisotropy, due to single-ion anisotropy or Dzyaloshinsky-Moriya interactions. In an amorphous material these anisotropies vary from site to site so that so-called amorphous magnets often possess random anisotropy, so that there can be a locally varying 'easyaxis' for the magnetization.

Competing interactions are automatically present in a random-bond spin glass, because different bonds 'pull' the system in different ways. These competing interactions lead to frustration, so that as in systems such as the kagomé lattices, there is a multidegenerate ground state. Spin glasses share this multidegenerate ground state but also show a new effect, a cooperative freezing transition, which will now be described.

At high temperature the behaviour of all magnetic systems is dominated by thermal fluctuations, so that in a spin glass all the spins are independent. As a spin glass is cooled from high temperature, the independent spins slow down and build up into locally correlated units, known as clusters. The spins which are not in clusters take part in interactions between clusters. As the temperature cools to T_{f} the fluctuations in the clusters progressively slow down. The interactions between spins become more long range so that each spin becomes more aware of spins in a progressively growing region around it. At T_{f} the system finds one of its many ground states and freezes. This process is not fully understood but it appears to be a cooperative phase transition. However this is not a phase transition to a magnetically ordered state; there are no magnetic Bragg peaks found in scattering experiments as would be the case if the system

showed long range magnetic order. Below T_f the ground state appears to be 'glassy', possessing metastability and slow relaxation behaviour. There is a divergence between the field-cooled and zero-field cooled magnetic susceptibility below T_f, reflecting this metastability (see Fig. 8.3).

One of the signatures of spin glass behaviour is a sharp peak close to T_f in the real part of the a.c. susceptibility, $\chi(\omega)$. In this technique the magnetic susceptibility is measured using a very small alternating magnetic field of frequency ω, sometimes with a constant (d.c.) magnetic field also applied. The position of the peak varies slightly with ω. The imaginary part of $\chi(\omega)$, related to the absorption, shows a sudden onset near T_f. The dynamics of the fluctuations associated with the freezing process can be studied using a.c. susceptibility, neutron spin echo (a method of measuring time correlation functions on relatively long time scales) and μSR. These techniques show

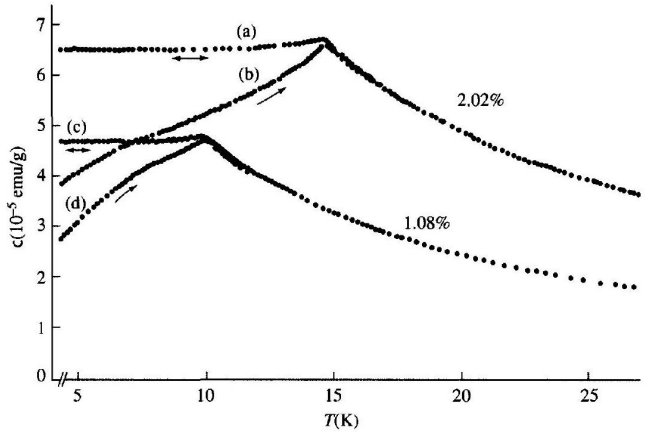


Fig. 8.3 The static susceptibility of $\mathrm{Cu
{1-x} |*mathrm{Mn}*{x}$ for $x = 0.0108$ and $x = 0.0202$. After zero-field cooling, χ was measured in a small field of 0.59 mT for increasing temperature ((b) and (d)). If the samples were field-cooled in 0.59mTχ was reversible ((a) and (c)) (Adapted from Nagata et al. 1997).}

that there is a large distribution of relaxation times for the clusters which form above T_f. As the sample is cooled, these clusters grow and fluctuate at a rate dependent on the size and nature of each cluster. A spin outside the clusters will fluctuate rapidly but as soon as it joins a growing cluster (something of a 'hostile takeover'!) it is forced to slow down and keep step with its new host. At T_f the random anisotropy becomes important and the system freezes into a random orientation in which the clusters are fixed. A wide range of relaxation times is also observed below T_f, showing that some free spins or small superparamagnetic (see below) clusters are still there. The co-operative spin freezing in spin glasses is still not yet fully understood.

The $\mathrm{Cu}_{1-x}\mathrm{Mn}_x$ spin glass is usually studied when x is a few percent or less. In a concentrated spin glass the behaviour is very different because the greater number of magnetic nearest neighbours, which can be enhanced by chemical clustering, leads to the presence of large, ordered magnetic clusters which then dominate the magnetic behaviour. The system then becomes known as a cluster glass (sometimes known as a mictomagnet) which has some characteristics of a spin glass but long range magnetic order is always in the background waiting to spring (the technical term is that the long range magnetic order is incipient). Sometimes the high temperature state can actually be an ordered ferromagnet, rather than a paramagnet, and the system actually freezes to a low temperature non-magnetic disordered state. Such a system is called a re-entrant spin glass and the behaviour is believed to be connected with either temperature-dependent random anisotropy or some sort of anisotropy associated with the freezing. In very high concentrations of magnetic ions, the system approaches the percolation

limit at which it is possible to find a path of nearest-neighbour links through the whole sample along which each ion is magnetic; this results in the sample having long range magnetic order. The transitions between these different states can be seen in the phase diagram of $Au_{1-x}Fe_x$ shown in Fig. 8.4.

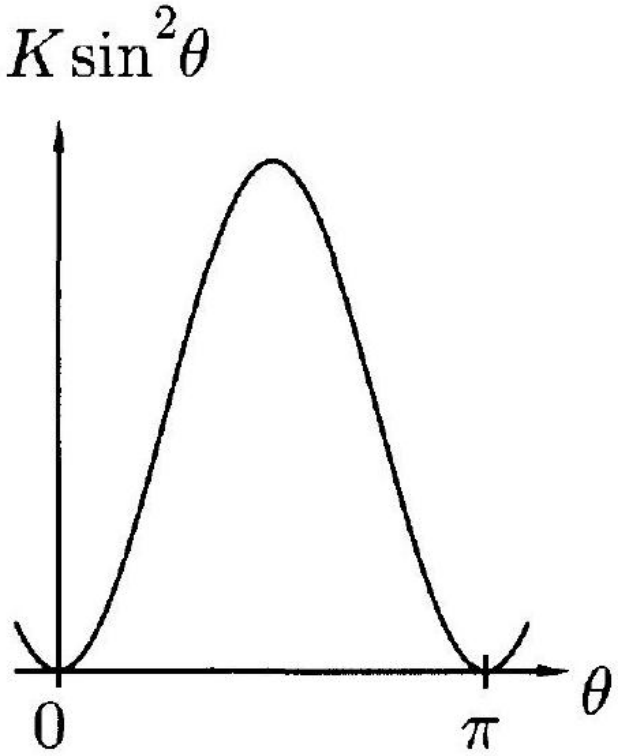
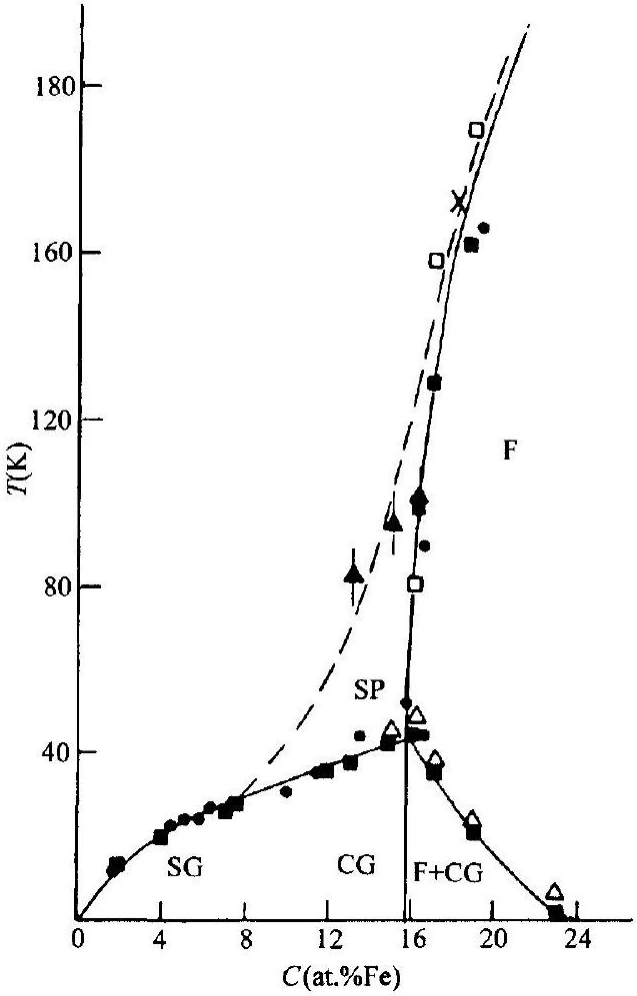


Fig. 8.5 The energy density of a magnetic particle contains a term $K \sin ^{2} \theta$. The energy is minimized when $\theta = 0$ or π.}

Fig. 8.4 The phase diagram of $\mathrm{Au

{1-x} \mathrm{Fe}{x}$. P = paramagnetic, SP = superparamagnetic, SG = spin glass, CG = cluster glass, F = ferromagnetic. After Coles et al. 1978.}



9.8.3 8.3 Superparamagnetism

If ferromagnetic particles are small enough, they will be single-domain (see Section 6.7.7), because the energy cost of domain wall formation does not outweigh any saving in demagnetizing energy. The magnetization of a small, single-domain ferromagnetic particle is often constrained to lie parallel or antiparallel to a particular direction. This can be due to magnetocrystalline anisotropy, or shape anisotropy (associated with the demagnetizing energy and the shape of the particle), or a variety of other reasons. However, we will assume that the energy density of the particle contains a term $K \sin^2 \theta$ where θ is the angle between the magnetization and this particular direction and K is a constant which quantifies the energy density associated with this anisotropy. Thus the energy is minimized when $\theta = 0$ or π (see Fig. 8.5). A particle of volume V needs an activation energy of $\Delta E = KV$ to flip its magnetization from $\theta = 0$ to π or from π to 0 . For very small particles, such that KV is small compared to $k_{\mathrm{B}}T$, the magnetization can be easily flipped in this way by thermal fluctuations.

Consider now a distribution of these small ferromagnetic particles in a nonmagnetic matrix and assume that the particles are separated far enough apart so that interparticle interactions can be neglected. For $k_{\mathrm{B}}T \gg KV$ the system will therefore behave like a paramagnet, albeit one in which the independent
moments are not atomic moments but large groups of moments, each group inside a ferromag-

netic particle. The system is therefore called a superparamagnet. At high temperatures the moments on the particles are able to fluctuate rapidly. The relaxation time τ of the moment on a particle is given by

$$\tau = \tau_0 \exp\left(\frac{KV}{k_{\mathrm{B}}T}\right), \tag{8.1}$$

where τ_0 is typically 10^{-9} s. The fluctuations therefore slow down (τ increases) as the sample is cooled (see Fig. 8.6) and the system appears static when τ becomes much longer than the measuring time t of the particular laboratory experimental technique which you are using. If we define 'much longer' to be $\tau > \alpha t$ where $\alpha = 100$, then below the blocking temperature T_{B}, given by $T_{\mathrm{B}} = KV/k_{\mathrm{B}} \ln(\alpha t/\tau_0)$, each magnetic particle appears to be locked into one of its two minima.

The experimental measuring time t is in the range $10^{-12} - 10^{-10}$ s for inelastic neutron scattering, $10^{-10} - 10^{-7}$ s for Mössbauer (the decay processes in the Mössbauer transition) are $\sim 10^{-7}$ s, $10^{-10} - 10^{-5}$ s for μSR (a measurable fraction of muons live for up to $\sim 10\tau_\mu$ where $\tau_\mu = 2.2\mu$ s is the average muon lifetime) and a.c. susceptibility typically probes $10^{-1} - 10^{-5}$ s. Because of the logarithmic dependence on $\alpha t/\tau_0$, any of α, t or τ_0 can be changed by a few orders of magnitude with only a relatively small change to T_{B}. If the particles in the superparamagnetic system have a range of sizes then they will 'block' at different temperatures.

Superparamagnets have some similarities with spin glasses but are quite different in a number of respects: the interactions between the magnetic particles are not important in a superparamagnet, whereas they are vital in a spin glass; also, the spin glass shows a cooperative phase transition while the superparamagnet shows a gradual blocking of the superparamagnetic particles. Superparamagnetism is technologically important since many important recording materials are particulate. This is crucial because magnetic tape is expected to store information for years, not μs, which puts a limit on the minimum size of the particles. In studying rock magnetism it is sometimes necessary to consider the exponential decay of magnetism over millions of years. Thus the stability of particles against superparamagnetic fluctuations depends on the timescale which you are using; a particle may be stable in a Mössbauer experiment but fluctuate in an a.c. susceptibility; it may be stable for a century, but decay on geological time scales.

9.8.4 8.4 One-dimensional magnets

No long range order is possible in one-dimension for $T > 0$, so it might be thought that one-dimensional magnets were rather tedious and uninteresting. In fact nothing could be further from the truth! The one-dimensionality implies the possibility of complex excitations which are still far from being completely understood.

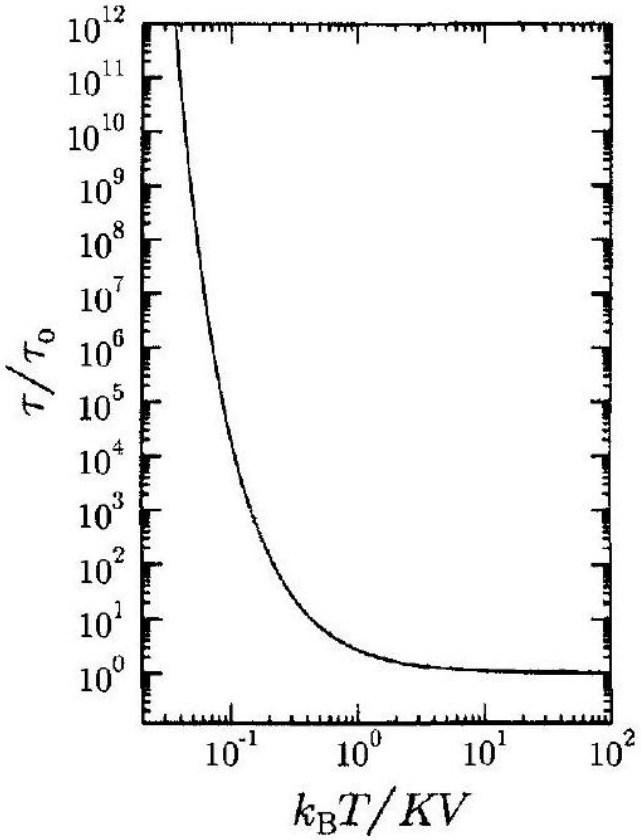


Fig. 8.6 The dependence of the relaxation time τ as a function of temperature T (scaled by $k_{\mathrm{B}} / K V$) according to eqn 8.1. The fluctuations therefore slow down (τ increases) as the temperature is reduced.}

8.4.1 Spin chains

A one-dimensional ($d = 1$) line of spins is known as a spin chain. The individual spins can be constrained to lie parallel or antiparallel to a particular direction (Ising spins), or may be free to point anywhere in a fixed plane (XY spins), or free to point in any direction (Heisenberg spins). Spin chains can be approximately realized in crystals, if the crystal structure is such so as to keep the chains reasonably far apart. The single ion anisotropy due to the crystal field may lead to the magnetic moments behaving as Ising spins ($D = 1$), XY spins ($D = 2$), Heisenberg spins ($D = 3$), or somewhere in between. A commonly studied family of systems is based on crystals of the type ABX_3 where A is a non-magnetic cation of a single charge, B is a doubly charged magnetic cation and X is a halide anion. This leads to a simple hexagonal lattice with transition metal ions forming chains along the c direction. For example $CsCoCl_3$ behaves almost as a one-dimensional Ising spin chain since the anisotropy constrains the spins along a particular direction; $KCuF_3$ behaves like a one-dimensional Heisenberg spin chain, as do a number of Cu salts with organic ligands.

Very often these systems show three-dimensional long range order at very low temperatures because there will always be some small interchain interaction which can couple the chains together. $CsCoCl_3$ shows three-dimensional long range magnetic order below 21 K because of this interchain coupling. Nevertheless, there is a wide region of temperature above the crossover to a three-dimensional region, where the magnetic behaviour is that of a onedimensional system.

The spin quantum number for each spin on the chain depends on the atom. For chains with Cu^{2+} $\left(3\ d^9\right)$ the spin quantum number is $S = \frac{1}{2}$, with Mn^{2+} $\left(3\ d^5\right)$ $S = \frac{5}{2}$ and for Co^{2+} $\left(3\ d^7\right)$

in $CsCoCl_3$ the ground state has an effective spin of $S = \frac{1}{2}$ (the $S = \frac{3}{2}$ free-ion ground state is split by the crystal field leaving a ground state doublet and excited states). In the next section we will just consider chains with $S = \frac{1}{2}$ on each site.

8.4.2 Spinons

What makes these chains interesting is not their ordering (because unless the interchain interactions are strong enough, they do not show order). but their excitations. As shown in Section 6.6.1, in three-dimensional Heisenberg magnets the excitations are magnons, which are bosons. (In a metallic magnetic material there may also be Stoner excitations, see Section 7.8.) Each magnon is an excitation with $S = 1$ and so a single magnon can interact with a neutron in an inelastic scattering experiment.

In Ising spin chains, the excitations are associated with the creation of domain walls. There are no gapless excitations (i.e. excitations with vanishingly small energy in the long wavelength limit; 'massless Goldstone modes' to use the particle physics parlance) because even to create one domain wall costs a finite amount of energy (and in fact for an excitation you have to create the domain walls in pairs). Once the excitation is created it can move freely along the chain. The excitation energy has no wave vector dependence, but if the chain is not perfectly Ising-like (as very often happens in real systems) then there will be some modulation of the dispersion relation.

In Heisenberg spin chains, the excitations are known as spinons. These have spin $-\frac{1}{2}$ (in contrast to magnons which have spin-1) and are fermions. They have a dispersion relation which is given by

$$\hbar\omega = \pi|J\sin(qa)|, \tag{8.2}$$

where J is the antiferromagnetic exchange coupling, and a and q are the lattice constant and wave vector, both measured along the chain direction. Equation 8.2 is the bold line in Fig. 8.7. This can be compared with the conventional spin wave dispersion relation in eqn 6.58 with $S = \frac{1}{2}$, but there is an additional factor of $\pi/2$. The excitations in both cases are gapless because when $q \to 0$ (long wavelength limit) $\omega \to 0$. A neutron scattering experiment involves a change of spin of one and so although this implies a creation or annihilation of a single magnon in a three-dimensional system, it implies a creation or annihilation of two spinons in a one-dimensional system. Neutron experiments therefore measure the momentum $q = q_1 + q_2$ and energy $\hbar\omega = \hbar\omega_1 + \hbar\omega_2$ associated with creation or annihilation of a pair of these spinons and so the experimental data show a continuum of excitations between eqn 8.2 and $\hbar\omega = 2\pi|\,\mathrm{J}\sin(qa/2)|$ (see Fig. 8.7). Neutron scattering experiments have confirmed that these excitations do exist in some one-dimensional antiferromagnetic chains (see Fig. 8.8). (For a derivation of eqn 8.2, see des Cloizeaux and Pearson 1962.)

8.4.3 Haldane chains

The previous section described how the excitations in spin- $\frac{1}{2}$ antiferromagnetic Heisenberg spin chains are believed to be spinons which are gapless excitations. This result is believed to be true also for half-integer spin chains (i.e. chains of spins with $S = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots$). Haldane conjectured that something different would happen for integer spin chains (i.e. chains of spins with $S = 1, 2, 3, \ldots$), namely that there would be a gap in the excitation spectrum which occurs because of nonlinear quantum fluctuations in the ground state. A one-dimensional chain of integer spins is therefore known as a Haldane chain and the gap in the excitation spectrum is known as a Haldane gap. This fundamental difference between half-integer and integer spin chains is related to the difference between fermions and bosons under exchange;

this different exchange symmetry has a topological origin and has a dramatic effect on the nature of the excitations.

Most tests of Haldane's conjecture have been carried out on materials with chains of $Ni^{2+}(S = 1)$ ions, including $CsNiCl_3$, $Ni(C_2H_8N_2)_2NO_2ClO_4$ and Y_2BaNiO_5. These $S = 1$ spin chains all seem to possess gaps in their excitation spectra as predicted. Half-integer antiferromagnetic chains are gapless, unless magnetoelastic coupling opens up a spin-Peierls gap, as discussed in the following section.

8.4.4 Spin-Peierls transition

Although spin- $\frac{1}{2}$ antiferromagnetic chains are gapless, they are susceptible to an analogous kind of instability that afflicts one-dimensional metals (see Section 7.9) which can open up a gap. This occurs at the spin-Peierls transition.

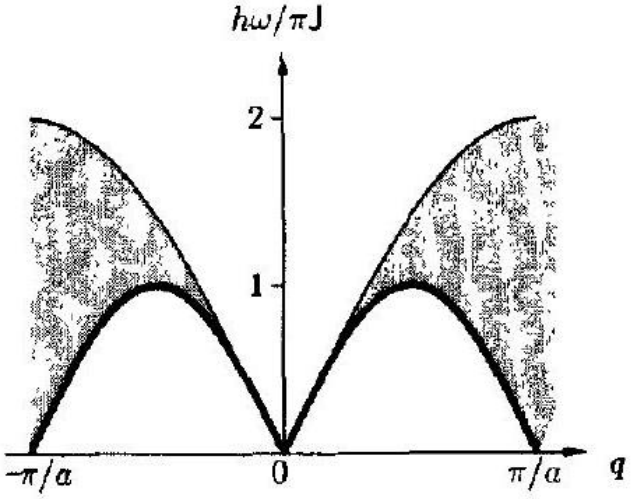


Fig. 8.7 Dispersion relation for spinon excitations in a one-dimensional antiferromagnetic Heisenberg spin chain (bold line). The shaded region shows the continuum of excitations measured in a neutron scattering experiment.

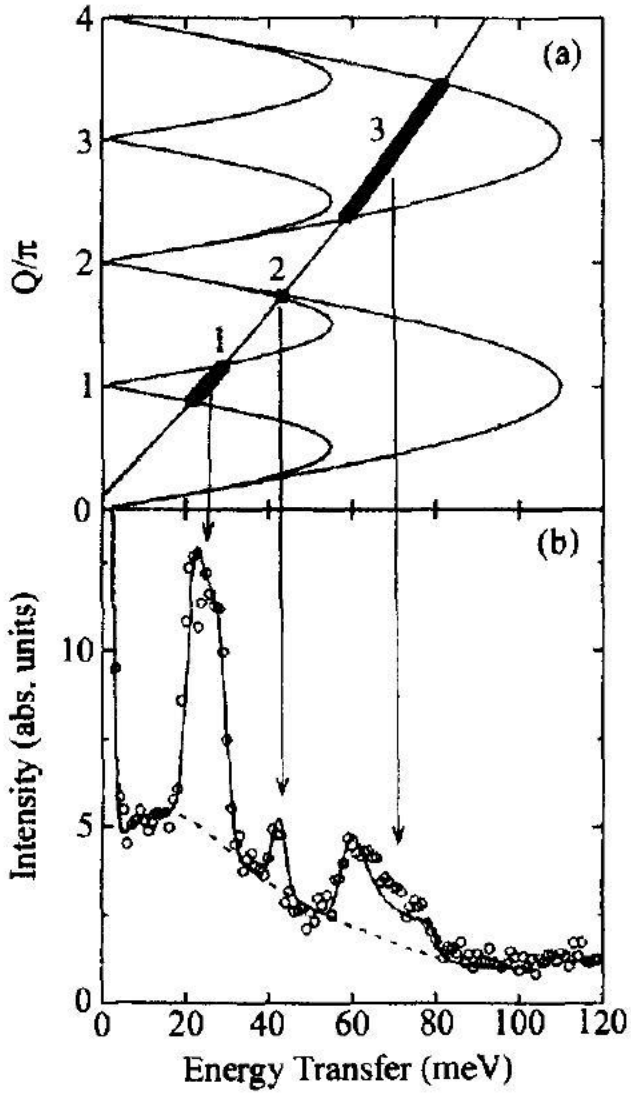
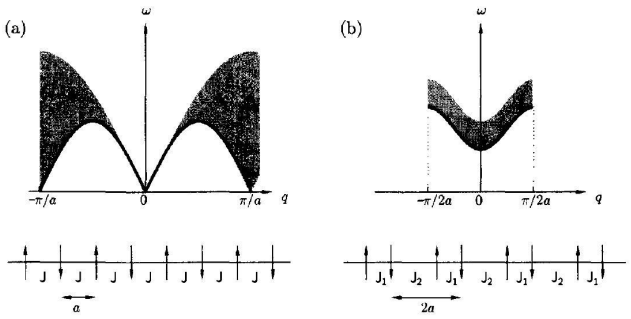


Fig. 8.8 (a) The dispersion relation for spinons in $\mathrm{KCuF _{3}$. The experiment was performed using a time-of-flight technique so the line shows the scattering trajectory for neutrons with incident energy 148.9 meV and an incident momentum aligned 8° away from the c^* direction in $KCuF_3$. Scattering results when the trajectory intersects with the continuum states. (b) Observed scattering at 20 K . The non-magnetic background is indicated by the dashed line. After Tennant et al. 1993.}

[01] MEM, TCNQ and TTF are organic molecules with lengthy chemical names.

Fig. 8.9 Schematic representation of the elementary excitations in (a) a uniform Heisenberg antiferromagnetic chain and (b) an alternating chain (for which the ground state is a singlet state at $q = 0$), and for which the unit cell is doubled. Adapted from Bray et al. 1983.



The driving force of this intrinsic lattice instability is the magnetoelastic coupling between the one-dimensional electronic structure and the threedimensional lattice vibrations (phonons). This coupling arises because the exchange energy of the chains is a function of the separation between adjacent lattice sites. A distortion of the lattice influences the magnetic energy (see Fig. 8.9). The name spin-Peierls reflects the similarity with the Peierls distortion (discussed in Section 7.9).

Above the transition temperature T_{SP}, there is a uniform antiferromagnetic next-neighbour exchange in each chain; below T_{SP} there is an elastic distortion resulting in dimerization, and hence two, unequal alternating exchange constants. The dimerization increases progressively as the temperature is lowered and reaches a maximum at zero temperature. The alternating chain possesses an energy gap between the singlet ground state and the lowest lying band of triplet excited states. The magnitude of the gap is related to the degree of dimerization and hence to the degree of lattice distortion, becoming zero for the uniform chain (zero dimerization) so that one returns to the gapless spinon case. Thus the magnetic susceptibility $\chi(T)$ shows a knee at T_{SP}, with a rather abrupt fall of χ below T_{SP}, corresponding to the opening of the gap (see Fig. 8.10). Whereas the normal Peierls distortion (the electronic analogue of the spin-Peierls transition, see Section 7.9) occurs at a temperature T_{P} of the order of $k_{\mathrm{B}}T_{\mathrm{P}} \sim E_F \exp\left(-1/\lambda_{\mathrm{el-ph}}\right)$, where $\lambda_{\mathrm{el-ph}}$ is the electronphonon coupling constant, the spin-Peierls transition will occur at $k_{\mathrm{B}}T_{\mathrm{SP}} \sim |J| \exp\left(-1/\lambda_{\text{sp-ph}}\right)$, where J is the exchange interaction between adjacent spins and $\lambda_{\text{sp-ph}}$ is the spin-phonon coupling constant. Since $J \ll E_{\mathbf{F}}$ (e.g. J is typically 50 K, E_{F} is typically $500 - 5000$ K), T_{SP} is always small in comparison with T_{P}.

There are only very few materials which show a spin-Peierls transition. This is because antiferromagnetic chains often become three-dimensionally ordered at low temperature due to interchain coupling. Only in very few materials is the spin-phonon coupling able to dominate the interchain spin-spin coupling and allow the formation of a spin-Peierls ground state. Examples of such materials include $CuGeO_3$ ($T_{\mathrm{SP}} = 14$ K) and a number of organic systems such as $MEM(TCNQ)_2$ ($T_{\mathrm{SP}} = 18$ K) and TTF-CuS $C_4(CF_3)_4$ ($T_{\mathrm{SP}} = 12$ K).[1]

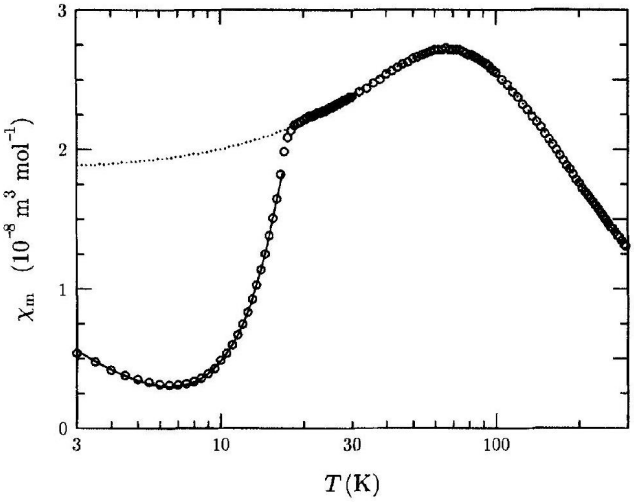
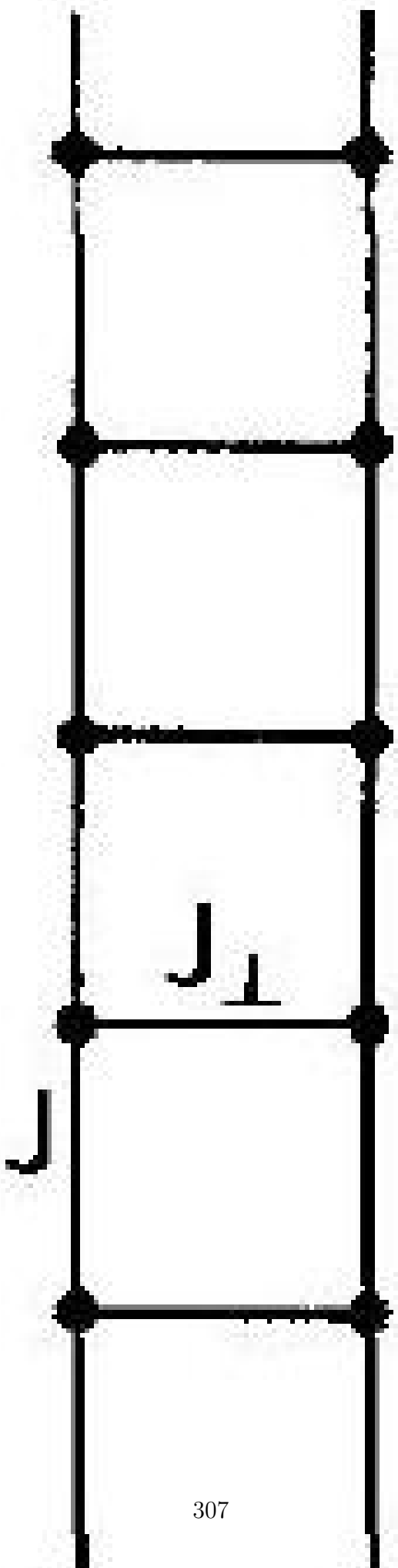


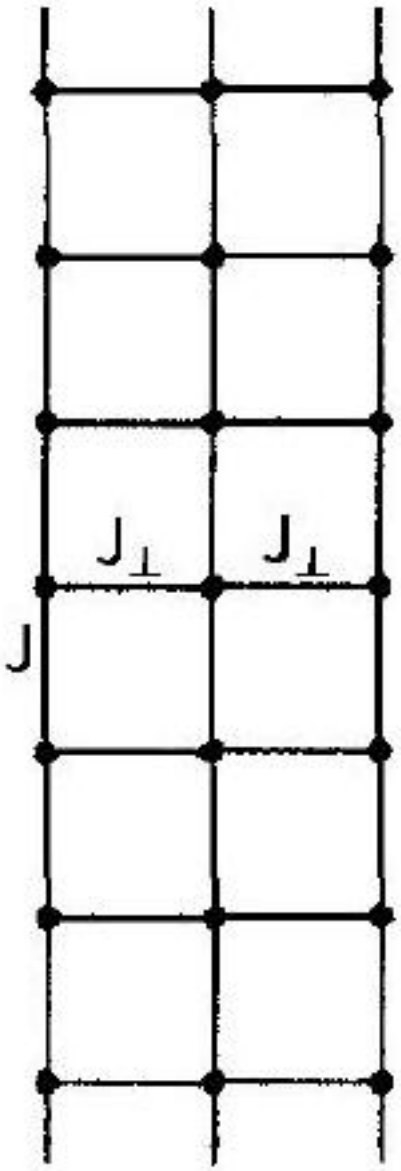
Fig. 8.10 The molar magnetic susceptibility of the organic spin-Peierls material MEM(TCNQ) $\{ _{2}$ which consists of stacks of the organic molecule MEM and stacks of the organic molecule TCNQ. At high temperature the susceptibility fits to a model appropriate for a uniform Heisenberg antiferromagnetic chain (dotted line), but on cooling the susceptibility drops rapidly at the spinPeierls transition as a gap in the excitation spectrum opens up. The rise at very low temperatures (as T is lowered) is due to the Curie-like ($\sim T^{-1}$) susceptibility from defects. After Lovett et al. 2000.}

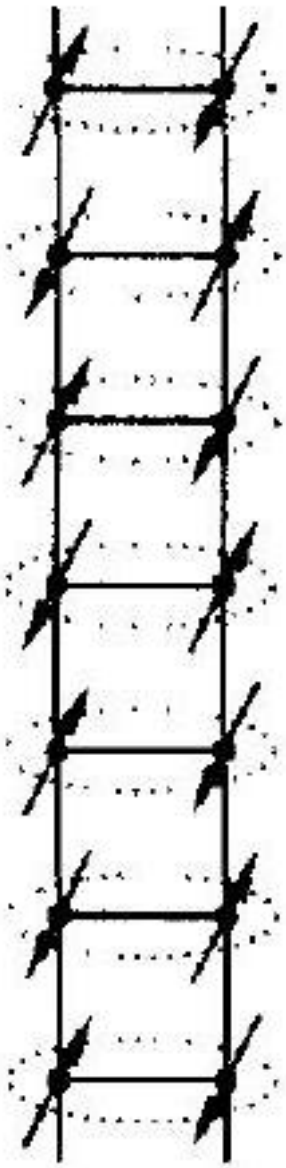
8.4.5 Spin ladders

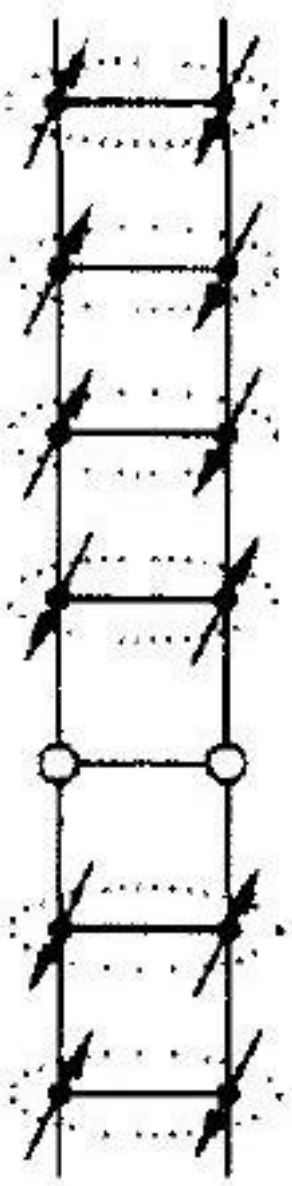
Before considering two-dimensional magnets, we can consider a system which is somewhere in between a one-dimensional magnet and a two-dimensional magnet. Consider two parallel spin chains with bonds between them such that the interchain coupling is of comparable strength to the intrachain coupling. Such a system is known as a two-leg spin ladder (see Fig. 8.11(a)). It is also possible to have three-leg (see Fig. 8.11(b)), four-leg, and in fact n-leg, spin ladders. Promising experimental systems include $SrCu_2O_3$ and $La_{1-x}Sr_xCuO_{2.5}$ (both two-leg spin- $\frac{1}{2}$), and $Sr_2Cu_3O_5$ (three-leg spin- $\frac{1}{2}$). In fact there is a general system $Sr_{n-1}Cu_{n+1}O_{2n}$ with n odd which consists of $(n+1)/2$-leg spin- $\frac{1}{2}$ because its structure has strips of a CuO_2 square lattice which have $(n+1)/2Cu^{2+}$ ions across their width.

The spin- $\frac{1}{2}$ two-leg ladder is known to have a finite gap in its excitation spectrum, which is easy to see in the 'strong-rung' limit in which the rung


J⊥
J


$J_\perp$
$J_\perp$
J





(b)
(c)
(d)
(e)

Fig. 8.11 (a) A two-leg ladder; (b) A threeleg ladder. (c) In a two-leg the ground state consists of spin singlets on each rung of the ladder and hence the two-leg ladder has a gap in its excitation spectrum. (d) Doping holes on to the ladder breaks up the singlets (e) but this energy cost can be minimized if the holes pair up.

coupling $J_\perp$ is larger than the coupling J along the legs. In this case the ground state consists simply of spin singlets along each rung of the ladder (see Fig. 8.11(c)). To create an excitation, you must promote a rung-singlet into a rung triplet (costing an energy $J_\perp$), hence the energy gap. If $J_\perp = 0$ the system is two isolated spin- $\frac{1}{2}$ chains which do not have a gap in their excitation spectrum. However it is believed that a gap appears as soon as $_\perp$ is non-zero, no matter how small it is. For the n-leg ladder, the situation is identical if n is even. However if n is odd, then on a given rung the spins will pair up into singlets leaving one left over. At large $J_\perp$ the system can be mapped into a spin- $\frac{1}{2}$ chain which is gapless. So an even-leg ladder has a gap in its excitation spectrum, while an odd-leg ladder is gapless. This appears to be supported by the results of neutron, μSR and transport experiments.

Doping the spin ladder with holes breaks up singlets in a spin ladder (see Fig. 8.11 (d)). For two-leg ladders there is an energetic advantage for the holes to pair up since they can then share a common rung, reducing the number of 'damaged' singlets from two to one (Fig. 8.11(e)). Hence it is favourable for holes to pair on two-leg ladders and this shows how it might be possible to engineer superconductivity in two-leg spin ladders. Superconductivity ($T_c \sim$ 14 K) has in fact been discovered in $Sr_{14-x}Ca_xCu_{24}O_{21}$ (sometimes called [14-24-41] or the 'phone number' compound) for $x = 13.6$ at 5 GPa . The interest in spin ladders derives from the fact they can have a gap in their excitation spectrum, they can become superconducting, and yet are simple well defined systems which theorists (who like working in one-dimension) can try to model. Hence these magnetic systems may shed light on the problem of high- T_c superconductivity (see Section 8.5 below).

9.8.5 8.5 Two-dimensional magnets

Two-dimensional magnetism is often studied in systems with formula A_2BX_4, where as before A is a non-magnetic cation of a single charge, B is a doubly charged magnetic cation and X is a halide anion. The crystal structure is tetragonal and the magnetic ions sit on a square lattice in two-dimensions. A typical material is K_2NiF_4 and very often these systems are said to have the K_2NiF_4 structure. The Mermin-Wagner-Berezinskii theorem demonstrates that in dimensions $d \leq 2$, thermal fluctuations prohibit the existence of long range magnetic order at non-zero temperature in an isotropic system. However it says nothing about the $T = 0$ ground state. In one-dimension, it turns out that for the spin- $\frac{1}{2}$ Heisenberg antiferromagnet no long range order exists

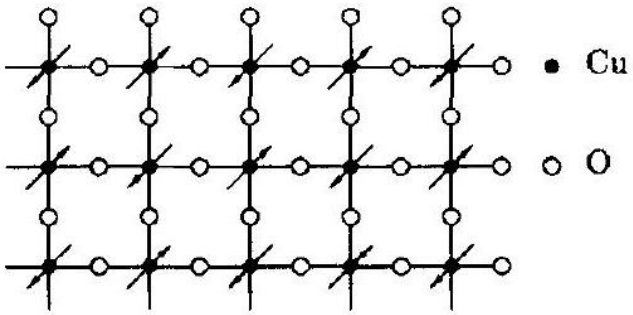


Fig. 8.12 Copper-oxygen planes in $\mathrm{La
{2} *mathrm{CuO}*{4}$.}

even at $T = 0$. The case for the two-dimensional Heisenberg antiferromagnet is not so clear cut, and there is currently evidence that this does show long range magnetic order at $T = 0$. This system is of great interest because twodimensional Heisenberg models appear to be important in understanding the high- T_c cuprate superconductors. These systems contain planes of copper and oxygen (see Fig. 8.12) which appear to be crucial for the superconducting properties.

For example, pure La_2CuO_4 is an antiferromagnetic insulator and has the K_2NiF_4 structure. The Cu^{2+} ions $\left(3\ d^9\right)$ have one hole in the d band and thus have spin- $\frac{1}{2}$. They sit on a square lattice, each Cu^{2+} ion separated from its
neighbour by oxygen ions which mediate an antiferromagnetic superexchange interaction between the Cu^{2+} spins (see Fig. 8.12). With one electron per site you would expect the system to be metallic, but the holes are localized because of the correlations. Doping the material with extra holes (by replacing some La^{3+} ions with Sr^{2+}) leads to high temperature superconductivity at around 40 K for the optimally doped samples (about 20% of the La^{3+} replaced by Sr^{2+}).

You can make isostructural materials by replacing Cu^{2+} with Ni^{2+} (which has $S = 1$) but these do not become superconducting, even when doped. There is therefore something special about Cu^{2+}. It is particularly surprising that these magnetic ions should be helpful in promoting superconductivity because magnetism normally destroys superconductivity. This is because local magnetic fields due to magnetic ions act as pair-breakers, splitting up the Cooper pairs that are responsible for the supercurrents. The magnetic properties of the two-dimensional $S = \frac{1}{2}$ square lattice Heisenberg antiferromagnet are therefore probably pertinent to the problem of high- T_c superconductivity, although this contention is unproven since, at the time of writing, no definitive theory of high- T_c superconductivity is universally believed.

The two-dimensional $S = \frac{1}{2}$ square lattice Heisenberg antiferromagnet is therefore of considerable interest and has attracted a number of theoretical and experimental investigations. As the temperature is reduced, the system does not order but the correlated regions of short range order increase in size. This size is known as the spin-spin correlation length, ξ, and it diverges exponentially at low temperature leading to true long range order at $T = 0$. The behaviour of ξ as a function of temperature can be calculated by mapping the two-dimensional $S = \frac{1}{2}$ square lattice Heisenberg antiferromagnet on to a continuum model known as the two-dimensional quantum nonlinear sigma model. This procedure is outside the scope of this book but the experimental results, which have given much support to this approach, are shown in Fig. 8.13. The results are obtained by using inelastic neutron scattering from $Sr_2CuO_2Cl_2$ which is an extremely good approximation to the ideal two-dimensional $S = \frac{1}{2}$ square lattice Heisenberg antiferromagnet. It has a larger interplane spacing than La_2CuO_4 and remains tetragonal down to low temperatures. Interplane interactions lead to long range three-dimensional antiferromagnetic order at $T_N = 256.5$ K, but well above this temperature the spin fluctuations are dominated by the two-dimensional fluctuations. Threedimensional order leads to sharp Bragg peaks in the inelastic neutron scattering data, but short range order with a correlation length ξ leads to broad peaks with width proportional to ξ^{-1} (for long range order $\xi \to \infty$ and the peaks become very sharp). The experimental data show that the peaks narrow as the material is cooled and the data for different incident energies are shown in Fig. 8.13 with a calculated line based on the predictions of the two-dimensional quantum nonlinear sigma model. There are no fitting parameters in this theory apart from J which has been obtained from Raman scattering measurements; the agreement is therefore extremely impressive and demonstrates that this system is beginning to be understood in some detail.

Nevertheless the application of these results to the problem of high- T_c superconductivity remains controversial. Recent tantalizing results have shown that in some of these systems the doped holes are not randomly arranged but segregate into stripes, the arrangement of the stripes depending rather subtlely

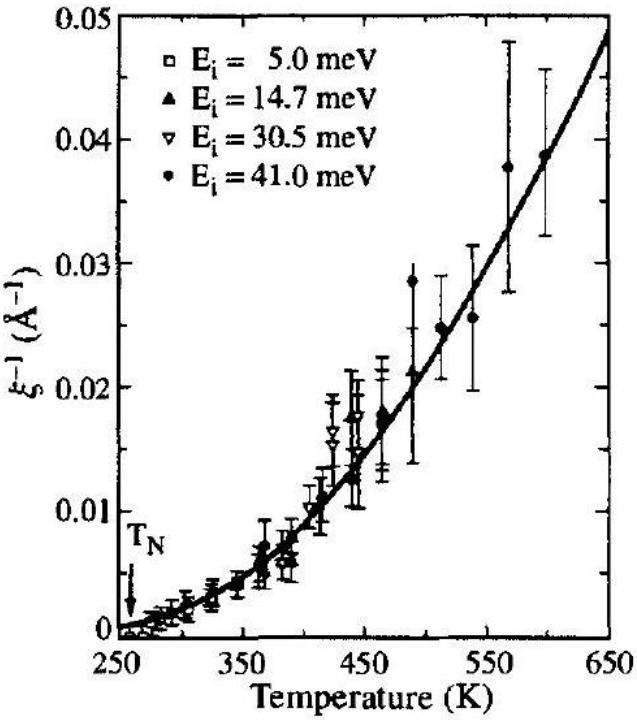


Fig. 8.13 The inverse correlation length ξ^{-1} in $Sr_2CuO_2Cl_2$ derived from inelastic neutron scattering. The data are for different values of the initial neutron energy E_i and the theoretical curve is based on the predictions of the two-dimensional quantum nonlinear sigma model. After Greven et al. 1994.}

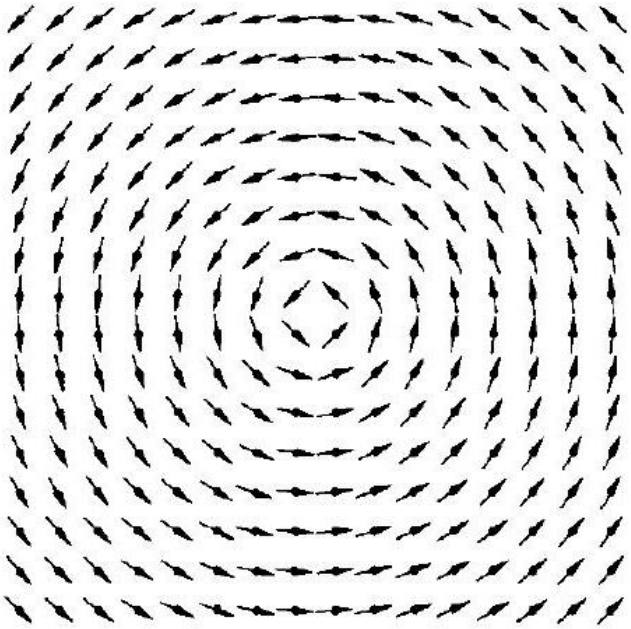


Fig. 8.14 A vortex state in the XY model.

on the doping level. This intriguing cooperative behaviour appears to be part of the puzzle, but whether it is the crucial piece of the jigsaw or another red herring remains to be seen.

The situation dealt with so far in this section has been Heisenberg spins which have dimensionality $D = 3$, i.e. even though they lie on a $d = 2$ dimensional lattice, the spins can point

in any three-dimensional direction. If the magnetic moments are constrained to lie in a plane, we have the twodimensional XY model ($D = 2$). In this case there is no long range order at $T \neq 0$, but nevertheless there is an interesting phase transition which can occur. The spin-spin correlation function only decays away algebraically at low temperatures, not exponentially as they do at high temperatures. There is a transition temperature between these two regimes which is associated with the thermal stability of vortices.

The energy to create a vortex in an XY system, such as the one drawn in Fig. 8.14 is given approximately by $\pi J \ln(R/a)$ where R is the size of the system and a is a lattice constant. This energy cost diverges as $R \to \infty$ so it would seem that vortices are very costly to produce. However, the centre of the vortex could be at any one of the $(R/a)^2$ sites of the system, so the entropy of the vortex is $S = k_{\mathrm{B}} \ln[(R/a)^2]$. Hence the free energy of a vortex is

$$F = (\pi\rfloor - 2k_{\mathrm{B}}T) \ln(R/a) \tag{8.3}$$

which becomes negative above a temperature T_{KT} given by

$$T_{\mathrm{KT}} = \frac{\pi J}{2k_{\mathrm{B}}}, \tag{8.4}$$

at the Kosterlitz-Thouless transition where vortices can be spontaneously thermally produced. This is a special phase transition because it occurs between two completely disordered phases, one at low temperature and one at high temperature. It is accompanied by a change of rigidity; the low temperature state has an elastic rigidity and the high temperature state does not. The transition can therefore also be thought of as a vortex unbinding transition since any vortices in the system below T_{KT} are strongly bound. It is called a topological phase transition because there is no symmetry breaking (in contrast with a ferromagnetic-paramagnetic transition).

If the spins are constrained to lie in a line, we have a two-dimensional Ising model ($d = 2, D = 1$) which does have a magnetic phase transition. Materials such as K_2CoF_4 and Rb_2CoF_4 are approximate realizations of the two-dimensional Ising model and have been studied in detail using neutron scattering and other techniques.

9.8.6 8.6 Quantum phase transitions

The phase transitions which we considered in chapter 6 are all driven by temperature. In such phase transitions, it is the thermal fluctuations (whose energy scale is controlled by $k_{\mathrm{B}}T$) which destroy the order as the sample is warmed through its transition. However, if one has a transition which is controlled by some other variable (such as pressure, magnetic field, or doping level) then at some critical value of this variable one can have a transition which can, in principle, occur at absolute zero. Such a zero temperature phase transition is called a quantum phase transition and the point at which it occurs is a quantum critical point. The relevant fluctuations are no longer thermal but the quantum mechanical fluctuations determined by Heisenberg's uncertainty principle. Quantum fluctuations have quite different properties from thermal fluctuations, and require completely different theories to describe them. A quantum system is described by a complex-valued wave function and this introduces features which are not present in classical systems.

In a classical phase transition driven by temperature, the correlation length and correlation time diverge as you approach the transition. The order parameter therefore fluctuates more slowly and over an increasing scale as the transition is approached. Therefore there is some frequency, say ω^*, associated with the thermal fluctuations, which tends to zero at the critical temperature T_{c}. As long as $k_{\mathrm{B}}T_{\mathrm{c}} \gg \hbar\omega^*$ close to the phase transition, the critical fluctuations

will behave classically. Hence for a quantum phase transition, where $T_c = 0$, quantum fluctuations cannot be ignored. Systems with quantum critical points are expected to show unusual dynamics which is controlled by quantum fluctuations and can be probed using such techniques as NMR and neutron scattering.

An example of a quantum phase transition is the Ising magnet $LiHoF_4$ in which the ferromagnetic order can be destroyed at absolute zero by applying a magnetic field perpendicular to the easy-axis of the Ising spins (see Fig. 8.15). This counterintuitive behaviour occurs because the magnetic field facilitates quantum tunnelling between the up and down spin states. Above a critical magnetic field the quantum fluctuations are sufficient to destroy the ferromagnetic order. The paramagnetic state is different from the usual high temperature paramagnet in which the spins constantly, but classically, fluctuate between the up and down states. In $LiHoF_4$ there is a unique phase-coherent wave function for the ground state which is a quantum superposition of up and down spin states.

There are a number of materials which show quantum critical points involving the proximity of magnetic and superconducting regions of the phase diagram. These include heavy-fermion [2] compounds such as $CePd_2Si_2$ (see Fig. 8.16(a)) which is an antiferromagnet at ambient pressure but upon application of hydrostatic pressure [3] can be made to superconduct. The phase transition at 20 kbar and $T = 0$ is a quantum critical point.

A second example is a family of organic superconductors based upon the organic molecule ET. [4] One can make very pure crystals of charge transfer salts of ET in which a pair of ET molecules jointly donate an electron to an inorganic anion, such as $Cu(NCS)_2^-$. The crystal structure consists of alternating layers of ET molecules and layers of the inorganic anion. The conductivity is high in the plane of the organic layers because of the molecular overlaps of the ET molecules, but is low between the planes. Organic materials are much softer than inorganic materials and so hydrostatic pressure has a much larger effect than in $CePd_2Si_2$. Also, by making changes to the anion, for example by making it bulkier, one can apply negative pressure by chemical means. Figure 8.16(b) shows that one obtains a similar phase diagram using a combination of chemical pressure and applied pressure. Again, an antiferromagnetic phase lies close to a superconducting phase.

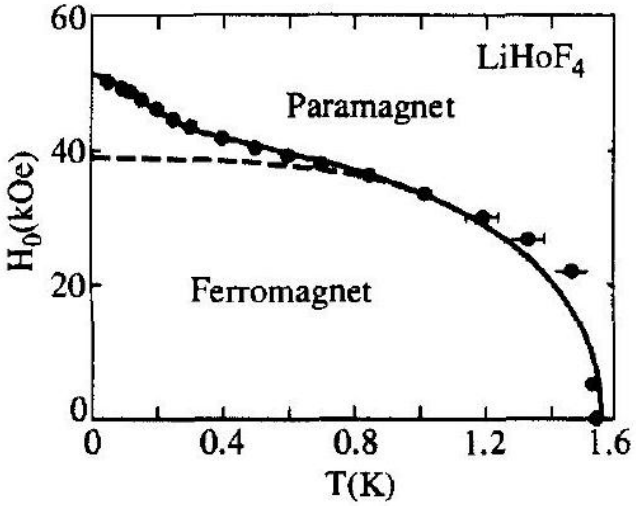


Fig. 8.15 The phase diagram of $\mathrm{LiHoF
_{4}$. The dashed line is a mean-field theory including only electronic spin degrees of freedom, while the solid line also incorporates the nuclear hyperfine interaction. After Bitko et al. 1996.}

[2] Heavy-fermion compounds have very large effective masses, hence the name. They are

often compounds of Ce or U .
[3] Hydrostatic pressure is pressure applied isotropically, the same way in all directions.
[4] ET is short for BEDT-TTF, which is in turn short for bis-ethylenedithiotetrathiafulvalene.
[5] In UGe_2 the superconducting region of the phase diagram is adjacent to a ferromagnetic region (Saxena et al. 2000).
[6] In the BCS theory of superconductivity, the attractive pairing interaction between electrons is mediated by phonons. This appears to be correct for conventional superconductors, but is not the right description of superconductivity in heavy-fermion, organic and high temperature superconductors.

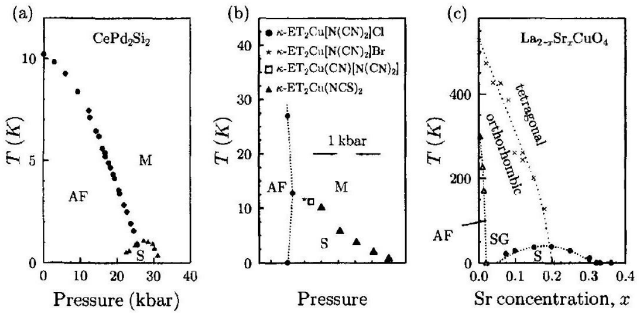


Fig. 8.16 (a) Temperature-pressure phase diagram of the heavy-fermion superconductor $\mathrm{CePd
*{2} \mathrm{Si}{2}$ (after Mathur et al. 1998); AF = antiferromagnetism, S = superconducitivity, M = metal. (b) Temperature-pressure phase diagram of the organic superconductors $\kappa-(\mathrm{ET})_2\mathrm{Cu}(\mathrm{NCS})_2, \kappa\ (\mathrm{ET})_2\mathrm{Cu}\left[\mathrm{N}(\mathrm{CN})_2\right]\mathrm{Br}, \kappa-(\mathrm{ET})_2\mathrm{Cu}(\mathrm{CN})\left[\mathrm{N}(\mathrm{CN})_2\right]$ and $\kappa-(\mathrm{ET})_2\mathrm{Cu}\left[\mathrm{N}(\mathrm{CN})_2\right]\mathrm{Cl}$. The pressure axis includes the effect of 'chemical pressure' caused by chemically varying the unit cell size (see text) as well as conventionally applied hydrostatic pressure (adapted from Kanoda 1997, Lubczynski et al. 1996). (c) Phase diagram of the cuprate superconductors $La_{2-x}Sr_xCuO_4$. (adapted from Kastner et al. 1998); SG = spin glass.}

A third example is one we have already described, namely the antiferromagnet La_2CuO_4 which becomes superconducting when it is doped with Sr . Here the variable which controls the properties is the level of Sr doping, x. At a critical value of x the system begins to switch from antiferromagnetic to superconducting.

This proximity between magnetic [5] and superconducting regions of the phase diagram is evidence that the spin fluctuations, rather than phonons, [6] are important in mediating the superconductivity in some of these materials. This story is however, at the time of writing, still far from being wrapped up.

9.8.7 8.7 Thin films and multilayers

So far we have considered two-dimensional magnets which occur in crystals. But it is also possible to directly engineer such systems by growing a very thin film of magnetic material using surface-science techniques. By using molecular-beam epitaxy, a thin magnetic film can be grown on to a flat substrate with a thickness of a single atomic layer. The growth is carried out under ultra-high vacuum conditions so that contamination and oxidation is reduced as far as possible. The lattice constant and symmetry of the substrate are chosen so as to closely match the material being grown. Not only films but mulitlayers can be grown, so that sandwich structures can be prepared and magnetic coupling studied through non-magnetic spacer layers, as will be considered in a later section.

The magnetic properties of a monolayer film are expected to be greatly different from the bulk material from which it is grown. An atom at the surface of a film has a smaller number of nearest neighbours in comparison with an atom in the bulk. The reduction in nearest neighbours reduces the electronic
bandwidth and hence increases the density of states at the Fermi energy level, $g(E_{\mathrm{F}})$. This increases the chance of the Stoner criterion (see Section 7.3) being satisfied and hence increases the propensity to magnetic order. Magnetic thin films are therefore expected to have enhanced magnetic moments. In addition, a monolayer of Pd or V (normally non-magnetic metals) is predicted to be ferromagnetic because of this enhancement.

Another trick which can be achieved with these techniques is to grow metastable phases of magnetic elements by appropriately choosing a substrate on which to grow the film. Thin film growth of magnetic materials on crystalline substrates allows the forces present at the interface to drive the film into a different crystallographic phase. This alternative phase may already be known as a high temperature or high pressure phase, or alternatively as a phase which is completely unknown in the bulk. At room temperature Ni is normally face-centred cubic (fcc), Fe is normally body-centred cubic (bcc) and Co is usually hexagonally close-packed (hcp). However, by using particular substrates that lattice-match to each other, metastable crystallographic phases can be grown. For example metastable phases of fcc Co can be stabilized on fcc Cu surfaces and bcc Co can be stabilized on GaAs(110) surfaces. Since the energies associated with a change in crystal structure ($\sim$ 0.1eV/ atom) are of the same order of magnitude as those associated with a change of magnetic structure (e.g. ferromagnetic to antiferromagnetic), often the magnetic properties of thin films dramatically depend on the growth conditions, as well as on the structure of the substrate.

The change in symmetry of atoms at the surface of a thin film also has an impact on the magnetic anisotropy and the easy direction of magnetization. To lowest order, the anisotropy energy of a ferromagnetic layer may be written as

$$E_{\mathrm{an}} = K \sin^2 \theta \tag{8.5}$$

where K is an effective anisotropy constant which is the sum of three terms,

$$K = \frac{2K_s}{t} + K_v - \mu_0 M^2 \tag{8.6}$$

and θ is the angle between the magnetization and the surface normal. The first term in eqn 8.6 represents the surface anisotropy. K_s is the surface or interface anisotropy constant and t is the thickness of the layer. The factor 2 appears because each layer has in general two faces. The second term is the volume anisotropy and may be due to lattice strains, or may appear because we have a uniaxial single crystal with its axis perpendicular to the plane of the film; K_v is the volume anisotropy constant. The third term is the shape anisotropy (see Section 6.7.2) of the film. It is a dipolar contribution and is calculated by assuming a uniform distribution of magnetic poles on the plane surfaces, and can therefore in practice be significantly reduced from this calculated value by the presence of interface roughness.

In thick films, the dipolar term dominates and the magnetization lies in the plane of the film. In thin films, the surface or interface term may dominate (because of the t^{-1} factor) and the spontaneous magnetization may become perpendicular to the plane of the film. This perpendicular anisotropy can be particularly useful in recording applications. [7] K_s arises because the surface states have lower symmetry which enhances the magnetocrystalline anisotropy

[7] In recording applications, small regions on the film store bits of information. The dipolar field experienced by one of these small regions due to its neighbours is lower if the magnetization is perpendicular to the plane. Therefore the bits can, in principle, be stored at a higher density. However, this explanation glosses over a large number of practical issues which can sometimes

lead manufacturers to prefer systems in which the magnetization lies in the plane.
at the surface. It is affected by the roughness of the interface and also by the lattice mismatch at each interface. The preparation of systems with large perpendicular magnetization is of great importance in the development of magnetic recording technology.

A number of experimental techniques have been used to study the magnetization of thin films. Mössbauer techniques have already been described (see Section 3.2.3) and magneto-optic methods will be introduced in the next section. Polarized neutron reflection (PNR) is a technique which involves spin-polarized neutrons reflected at grazing incidence from a magnetic multilayer. The reflectivity is different for spin-up and spin-down neutrons and can be used to infer the depth-dependent magnetization (direction and magnitude). In magnetic force microscopy (MFM), a magnetic moment on a cantilever is scanned across the surface of a sample. There is a force on the tip whenever there is a gradient in the magnetic field above the sample, so the technique is sensitive to domain walls (see Section 6.7.6). Another microscopy technique is SEMPA (Scanning electron microscopy with polarization analysis) which is a conventional scanning electron microscope which studies the ejection of secondary electrons (these retain their spin polarization and give information on the magnetization of the surface). Other techniques include Kerr microscopy (in which a polarized laser beam is scanned across a surface to image the magnetization) and scanning SQUID microscopy.

Michael Faraday (1791-1867)

9.8.8 8.8 Magneto-optics

John Kerr (1824-1907)
Magneto-optical effects were first studied by Michael Faraday in 1845 who showed that when polarized light passed through a piece of glass that was placed in a magnetic field, the light emerged with its plane of polarization rotated. This effect is now known as the Faraday effect. A related phenomenon was found by John Kerr in 1877; while Faraday's effect occurs in transmission, the magneto-optic Kerr effect occurs in reflection. John Kerr reflected light from the polished pole of an electromagnet and noticed that its plane of polarization was rotated.

Both effects are related to the spin-orbit interaction, and a rigorous treatment of both effects requires the use of perturbation theory. Nevertheless the general principles can be understood rather more simply as follows. Right-handed and left-handed circularly polarized light cause charges in a material to rotate in opposite senses, and each polarization therefore produces a contribution to the orbital angular momentum having an opposite sign. A magnetic field gives rise to a spin-polarization along the magnetic field direction and the spin-orbit interaction then leads to an energy contribution for the two circular polarizations having the same magnitude but with opposite sign. This leads to right-handed and left-handed polarizations having different refractive indices in the material. If plane polarized light is incident on a magnetic material, it should be considered as a sum of right-handed and left-handed circularly polarized beams which propagate through the material at different speeds. When they emerge, these two beams recombine but the phase-lag between them implies that the emerging beam has a rotated plane of polarization.

In an isotropic medium the dielectric tensor becomes

$$\epsilon = \epsilon \begin{pmatrix} 1 & \mathrm{i}Q_z & -\mathrm{i}Q_y \\ -\mathrm{i}Q_z & 1 & \mathrm{i}Q_x \\ \mathrm{i}Q_y & -\mathrm{i}Q_x & 1 \end{pmatrix}, \tag{8.7}$$

where $\mathbf{Q} = (Q_x, Q_y, Q_z)$ is known as the Voigt vector which is aligned with the magnetic field and has a magnitude which depends on the material. It is the off-diagonal terms in the dielectric tensor which contribute to magnetooptic effets. The tensor leads to the two circularly

polarized normal modes which have dielectric constants $\epsilon_{\pm} = \epsilon(1 \pm \mathbf{Q} \cdot \hat{\mathbf{k}})$ where $\hat{\mathbf{k}}$ is the direction of propagation of the light. The circular modes travel with different velocities and attenuate differently. The emerging waves therefore combine to yield a rotated axis of polarization and also an ellipticity (the different attenuations lead to the polarization of the light that emerges being slightly elliptical).

Studies of metallic magnets use the Kerr effect rather than the Faraday effect because metals absorb light and so cannot be studied in transmission, unless a sample is very thin. The Kerr effect is particularly useful for thin magnetic films and surfaces. In this case the technique is called SMOKE (which stands for the surface magneto-optic Kerr effect). The technique can be used during film growth inside an ultra-high vacuum (UHV) chamber by passing a laser beam through a window of the chamber such that it reflects from the sample and emerges from the chamber through another window. The change of polarization of the light can be monitored using crossed polarizers.

The Kerr effect can be performed in various geometries (see Fig. 8.17): in the polar Kerr effect the magnetization direction $\mathbf{M}$ is perpendicular to the plane of the film. If $\mathbf{M}$ is in the plane of the film, then one can measure the longitudinal Kerr effect if $\mathbf{M}$ is in the scattering plane of the light or the transverse Kerr effect if $\mathbf{M}$ is perpendicular to the plane of incidence. Because it can be performed during film growth, SMOKE is useful in studies of surface magnetism, magnetism of films as a function of film thickness, and surface anisotropies (for example it can be used to show at what thickness of film the magnetization changes from in-plane to out-of-plane).

Magneto-optical effects are extremely useful in magnetic recording since one can scan a polarized laser beam across a disk very rapidly and measure the magnetic information encoded on the disk by the changes in polarization of the reflected light. Although magneto-optical effects are used to read the disk, data are written by using another laser to locally heat a tiny region of the disk above T_C and applying a magnetic field to the disk; only the heated region switches - the information on the rest of the disk is left unaltered. The laser beam then moves on to the next region to write the next bit. At the turn of the 21st century, magneto-optical recording disks have started to be of enormous importance, providing a cheap and portable method of storage of relatively large amounts of data.

9.8.9 8.9 Magnetoresistance

The change in resistance, R, of a material under an applied magnetic field H is known as magnetoresistance. The magnetoresistance $\Delta\rho/\rho$ is usually defined

Woldemar Voigt (1850-1919)

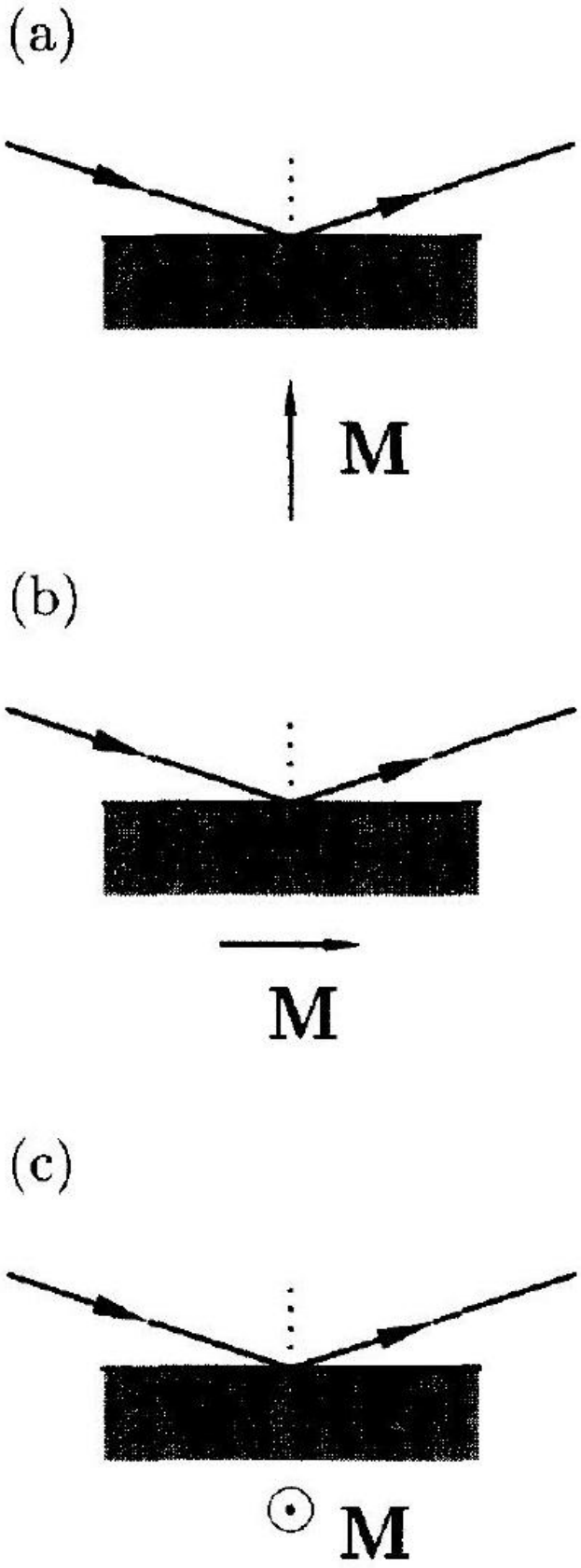


Fig. 8.17 (a) The polar Kerr effect; (b) the longitudinal Kerr effect; (c) the transverse Kerr effect.

by

$$\frac{\Delta\rho}{\rho} = \frac{R(H) - R(0)}{R(0)} \tag{8.8}$$

As we shall see, many interesting magnetoresistances are large and negative, and you can make the magnetoresistance of your new 'wonder magnetic material' appear deceptively larger by defining the magnetoresistance as $\{R(H) - R(0)\}/R(H)$, i.c. by dividing by the smaller number. This trick has been used more than once in the research literature!

William Thomson (Lord Kelvin) (18241907)

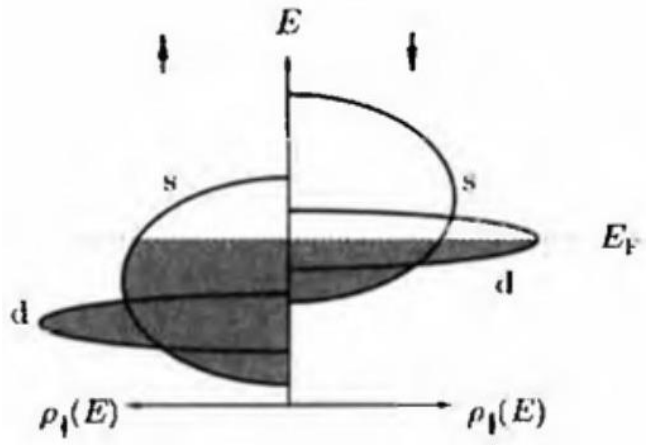


Fig. 8.18 The spin-split bands in a ferromagnet

It is a technologically useful quantity because magnetoresistive sensors are extensively used in applications (e.g. for measuring the magnetic field from the magnetic strip on a credit card). A free electron gas shows no magnetoresistance, even with an anisotropic effective mass; magnetoresistance only appears in models with more than one carrier, and its high field behaviour depends on the topology of the Fermi surface. The effect was first discovered in 1856 by Lord Kelvin (then William Thomson) who was examining the resistance of an iron sample. He found a 0.2% increase in the resistance of the iron when a magnetic field was applied longitudinally, and a 0.4% decrease when the field was applied in the transverse direction.

8.9.1 Magnetoresistance of ferromagnets

The observation of negative magnetoresistance in ferromagnets is a very puzzling one. When a metal carries a current, the displacement of electrons to different parts of the Fermi surface is such that scattering is minimized; the electrons find the path of least dissipation to cross the sample. Hence, if electrons are forced to take a different path, because of the presence of an applied magnetic field for example, they would take a path which leads to more scattering. Thus, in general, a positive magnetoresistance is expected. A negative magnetoresistance can sometimes be observed at low temperatures in samples which arc thin compared to the mean free path. If a magnetic field is applied in the plane and perpendicular to the current direction, the electron paths describe orbits with smaller diameters and therefore surface scattering is reduced. However, in ferromagnets, the explanation of the negative magnetoresistance must be entirely different.

A very important insight into this problem was provided by Mott who considered the transport properties of Ni (see Fig. 8.18 for a schematic diagram of the spin-split bands), in which only a few eV are needed to change the configuration from $(3d^8 4s^2)$ to $(3d^9 4s^1)$ or $(3d^{10})$. In general, Ni is considered to be $\left(3\ \mathrm{d}^{9.4} 4\ \mathrm{s}^{0.6}\right)$. The d band is very narrow (which is a necessary

condition for ferromagnetism in transition metals so that $g(E_{\mathrm{F}})$ is large and the Stoner criterion is satisfied), and hence $m_{\mathrm{d}}^* \gg m_{\mathrm{e}}$. As the s bands are nearly free. $m_s^* \sim m_e$. Hence, the conductivity σ, which is given by

$$\sigma = \frac{n_{\mathrm{s}} e^2 \tau_{\mathrm{s}}}{m_{\mathrm{s}}^*} + \frac{n_{\mathrm{d}} e^2 \tau_{\mathrm{d}}}{m_{\mathrm{d}}^*}, \tag{8.9}$$

is dominated by the first term and conduction is mainly due to the selectrons. In eqn 8.9, n_{s} and n_{d} are the number of electrons in the s and d bands respectively, and the scattering times are $\boldsymbol{\tau}_{\mathrm{s}} \sim \boldsymbol{\tau}_{\mathrm{d}}$.

The transition probability is mainly due to s $\rightarrow$ d transitions. At low temperatures, $T \ll T_{\mathrm{c}}$, all the unoccupied d states are antiparallel, so only half of the s electrons can make transitions. For $T > T_{\mathrm{c}}$, all the s electrons can make transitions, and so there is more scattering. Hence, a decrease in the resistivity below T_c is expected. Now, consider applying a magnetic field to

Ni. A magnetic field may increase the spin polarization and allow fewer $s \rightarrow d$ transitions; therefore a negative magnetoresistance is observed.

Most elastic collision processes are such that the electron conserves its spin. These collisions are characterized by a relatively short relaxation time τ. However, even a weak spin-orbit coupling can allow a weak spin-flip scattering which has a much longer relaxation time, τ_{sf}. In the absence of any external forces, a perturbation created in the equilibrium distribution of spinup and spin-down electrons, will decay first into a uniform distribution in each spin in a time τ, and after that it will reach the equilibrium distribution in a time τ_{sf}. For $T \ll T_{\mathrm{c}}$, spin-flip scattering is not expected to be dominant, and a ferromagnet can be well approximated by the two-current model in which the $\uparrow$ and $\downarrow$ electron currents are considered independently. This has been particularly successful in describing the properties of alloys in which a small quantity of one transition metal (the impurity) is dissolved in another transition metal (the host). The scattering due to certain transition-metal impurities is strongly spin-dependent. This is due to the combined effects of the spinsplitting of the host d band, the spin-splitting of the impurity d levels and the different hybridization between the host and impurity states for the spin $\uparrow$ and spin $\downarrow$ directions. For example, Cr impurities in Fe scatter the spin $\uparrow$ electrons much more strongly, resulting in a ratio of the resistivities for each spin-state of $\rho_{\uparrow}/\rho_{\downarrow} \sim 6$. At high temperatures, the spin-flip scattering due to collisions with spin waves leads to spin-mixing, i.e. the blurring of the distinction between the two spin channels. This concept should be kept in mind since it will return in the discussion of the giant magnetoresistance in sandwich structures.

8.9.2 Anisotropic magnetoresistance

In ferromagnets, the measured magnetoresistance depends on the orientation of the magnetization with respect to the direction of the electric current in the material. This effect is known as anisotropic magnetoresistance. Its origin is connected with the spin-orbit interaction and its influence on sd scattering. The symmetry of atomic wave functions is lowered by the spin-orbit interaction which also mixes states. The crystal axes determine the direction of $\mathbf{L}$ and the magnetization determines the direction of $\mathbf{S}$, so that the mixing of states leads to anisotropic scattering. Using symmetry arguments, it is possible to predict the general form of the dependence of the magnetoresistance on the direction of the magnetization and the current density in particular crystalline or polycrystalline materials.

8.9.3 Giant magnetoresistance

Anisotropic magnetoresistive effects are rather small, and so something of a revolution occurred in the late 1980s with the discovery of a very large effect (given the name giant magnetoresis-

tance, or GMR for short) in Fe/Cr/Fe multilayers. A large negative magnetoresistance of more than 50% was discovered at high magnetic field at low temperatures (see Fig. 8.19). The effect is associated with samples with magnetic Fe layers which were antiferromagnetically coupled. It was found that the coupling between magnetic layers through a spacer layer oscillates in sign as the spacer thickness increases. For

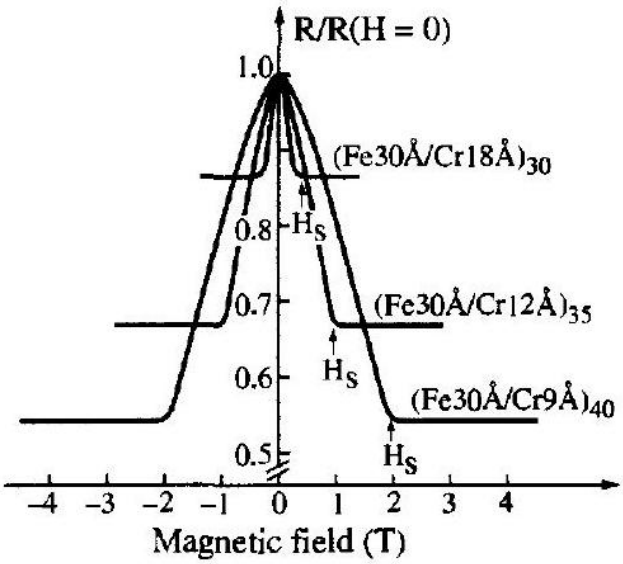


Fig. 8.19 Giant magnetoresistance in $\mathrm{Fe / \mathrm{Cr} / \mathrm{Fe}$ multilayers (Baibich et al. 1988).}

certain thicknesses it is ferromagnetic, and then is antiferromagnetic at larger thicknesses, and then returns to antiferromagnetic. It appears that the period of the oscillation is of the order of ten atomic spacings (usually between 9 and 18Å) and its value is mainly determined by the spacer metal, not by the ferromagnetic metal. Good matching between the spacer and ferromagnetic lattices favours large couplings. In the best cases, like Co/Ru superlattices, the amplitude of the oscillation decays with spacer thickness t as $1/t^2$.

The origin of these oscillations is connected with the RKKY interaction $\mathrm{J}(r) \sim \cos\left(2k_{\mathrm{F}}r\right)/r^3$ between two localized spins separated by a distance r in a bulk metal. When summed over all the spins on the interfaces, the coupling becomes $\mathrm{J} \sim \cos\left(2k_{\mathrm{F}}t\right)/t^2$ where t is the separation of the two ferromagnetic layers. The oscillations in the coupling can be directly related to the topology of the Fermi surface of the spacer layer. If the ferromagnetic metal is assumed to have a full majority spin d band, then in the case of ferromagnetic coupling, minority spin holes will be confined in the spacer layer, but in the case of antiferromagnetic coupling, there will be no confinement. The energy difference between these different couplings, with the number of particles conserved, is then entirely determined by the size quantization of the energy of minority spin holes in the spacer layer (which has thickness t). This arises because the energy levels in the spacer layer are discrete; therefore the density of states consists of a series of steps, the width of each being proportional to $1/t^2$. As t increases, the step width decreases, and one of the steps must eventually cross the Fermi level. The formalism which is used to calculate this effect can be understood as a one-dimensional analogue of the de Haasvan Alphen effect; in the case considered here, the size quantization is due to one-dimensional (rather than two-dimensional) confinement in a direction perpendicular to the layer planes

The type of magnetic coupling in a sandwich structure can directly influence the observed magnetotransport behaviour since this is very sensitive to the alignment of the magnetic layers,

with the GMR effects being largest for antiferromagnetic coupling. The first explanation for GMR was given in terms of the two-current model (see above) which separately considers the individual currents of ↑ and ↓ electrons (↑ means parallel to the majority spin band). In this discussion, I will initially assume that the mean free path λ is much greater than the Cr interlayer thickness t_{Cr}. Imagine an Fe/Cr/Fe structure in which the magnetization in each of the two Fe layers are aligned antiparallel in zero applied field, and suppose further that $\rho_{\downarrow} \ll \rho_{\uparrow}$. Then there are two cases to consider:
(1) $H > H_{\mathrm{s}}$ (where H_{s} is the saturation field). Here the magnetic moments in the Fe layers are aligned (see Fig. 8.20(a)) so that the resistivity ρ is given by

$$\frac{1}{\rho} = \frac{1}{\rho_{\uparrow}} + \frac{1}{\rho_{\downarrow}} \quad \Rightarrow \quad \rho \sim \rho_{\downarrow} \tag{8.10}$$

There is an effective short circuit by the less scattered electrons.
(2) $H = 0$. The magnetic moments in the two iron layers are now antiparallel (see Fig. 8.20(b)). In this case the electrons are alternately ↑ and ↓ in each of the layers with respect to the local magnetization, and the spin ↑ and ↓ channels are effectively 'mixed' (cf. the spin mixing due to spin waves in alloys at high temperature, considered above). so that $\rho_{\uparrow} \rightarrow \rho_{\mathrm{av}}$ and $\rho_{\downarrow} \rightarrow \rho_{\mathrm{av}}$ where $\rho_{\mathrm{av}} = (\rho_{\uparrow} + \rho_{\downarrow})/2$ so that the total resistivity ρ is given by

$$\rho = \frac{\rho_{\uparrow} + \rho_{\downarrow}}{4} \gg \rho_{\downarrow}. \tag{8.11}$$

This again predicts the negative magnetoresistance which is observed. The large effect is due to the inequality of ρ_i and $\rho_{\downarrow}$, which can be very large, and the ease in which the 'spin-mixing' can be switched on and off, simply by the application of the magnetic field.

For thicker Cr layers. the spin-dependent scattering at the interface affects the electron distribution function near the interface within a layer whose thickness is of the order of the mean free path λ. If $f_{\mathrm{Cr}} \gg \lambda$, the magnetoresistance is expected to fall roughly as $\exp(-t_{\mathrm{Cr}}/\lambda)$. However, there will be no giant magnetoresistance effect if there is no antiferromagnetic coupling, and since this appears periodically in t_{Cr}, the magnetoresistance is also expected to oscillate with increasing interlayer thickness, as well as die away exponentially.

Magnetic coupling through multilayers can be measured using many of the experimental techniques that we described in Section 8.7 on magnetic thin films. Useful additions to these are ferromagnetic resonance (the analoguc of NMR or ESR for ferromagnets) and Brillouin light scattering techniques (which study excitations using the inelastic scattering of light). These can both be used to study spin waves and magnetostatic modes in magnetic thin films. The spins in the multilayers are coupled together through exchange. dipolar and anisotropic interactions. Hence the spin-wave excitations, which are the eigenmodes of this magnetic system, have a dispersion relation which can depend quite sensitively on the exchange coupling and anisotropies and magnetoelastic effects.

The preparation of multilayers is not the only way to achieve GMR. It has also been observed in heterogencous $\mathrm{Cu-Co}$ alloy films. The relative oricentation of the magnetic moments inside the Co-rich grains inside the Curich matrix determines the magnetoresistance and this can be varied by an applied field. GMR is however not observed in homogeneous alloys which do not possess isolated, large, magnetically-rich grains of the appropriate size. Alloys are easier to prepare than multilayers and therefore this offers exciting prospects for sensor applications.

8.9.4 Exchange anisotropy

Exchange anisotropy (or unidirectional anisotropy) is the interfacial exchange that can be observed between a ferromagnet and an antiferromagnet. If the Curie temperature T_{C} of a ferromagnet is greater than the Néel temperature T_{N} of an antiferromagnet, then by depositing

one over the other (for an appropriate choice of the ferromagnet and antiferromagnet, and for a sufficiently thick ferromagnetic layer), and cooling them in an applied ficld through the Néel temperature, the measured magnetic hysteresis loop observed at $T \ll T_{\mathrm{N}}$ appears to be shifted as if another magnetic field was present in addition to the applied magnetic field. It appears that is is energetically

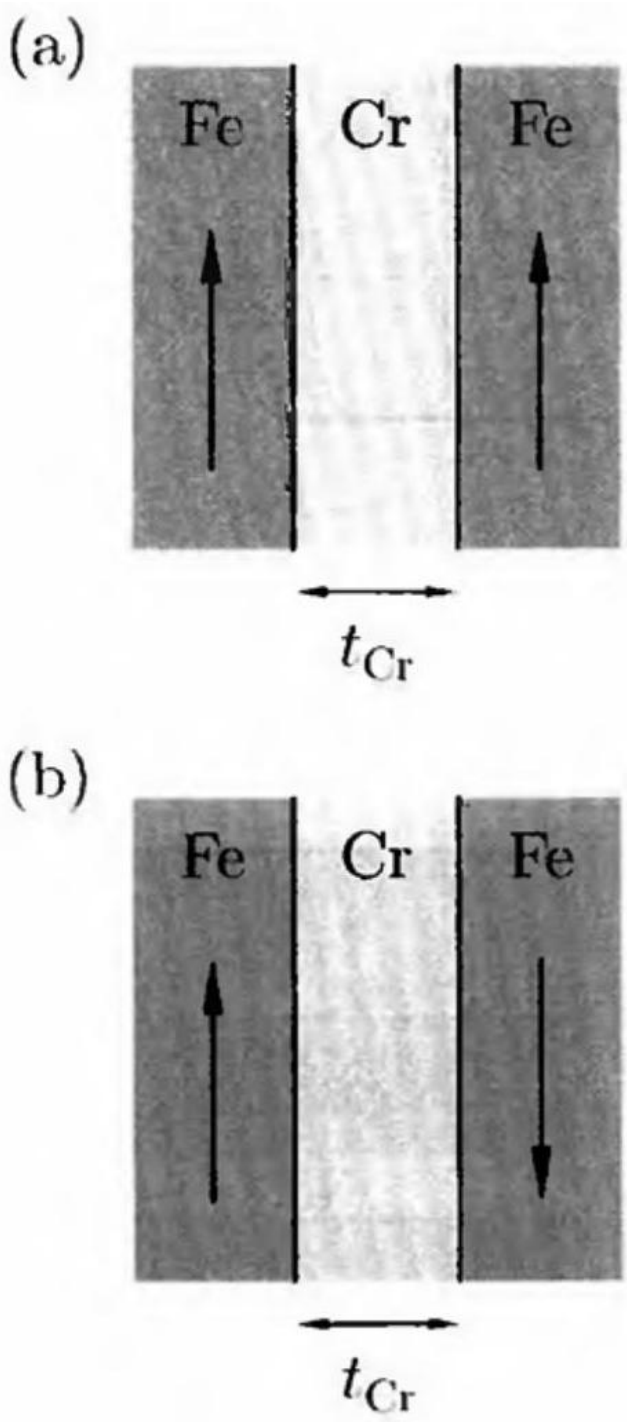


Fig. 8.20 (a) The magnetic moments in the Fe layers are parallet when $H > H$. (a) The magnetic moments in the Fe layers are antiparallel when $H = 0$ if the thickness of the chromium layer. $I_{\mathrm{Cr

}$. is chosen correctly.}

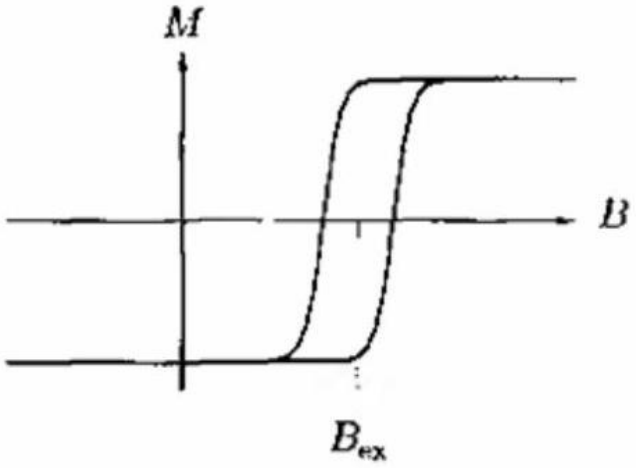


Fig. 8.21 A hysteresis kop shifted by an exchange anisotropy. Without the exchange anisotropy the loop would be centred at zero magnetic field.

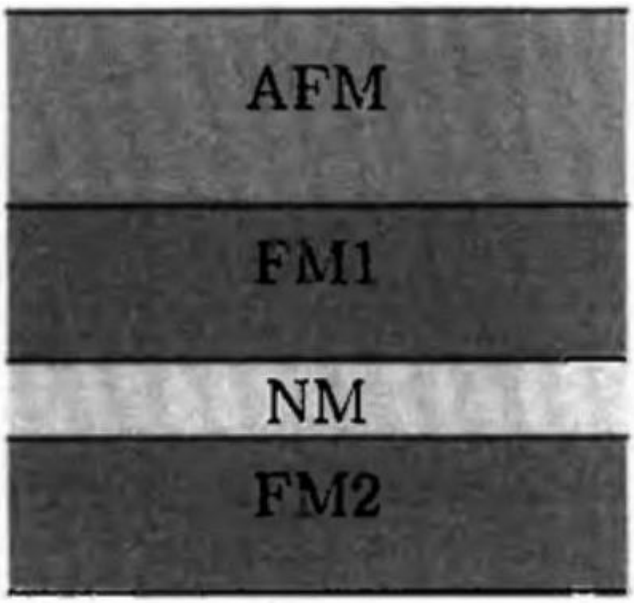


Fig. 8.22 A spin valve. The antiferromagnet (AFM) 'clamps' the magnetic moment in the upper ferromagnetic layer (FMI) along a particular direction. The magnetic moment in the lower magnetic layer (FM2) is free to rotate and can be aligned with an applied magnetic field. The angle between the magnetic moments in the two ferromagnetic layers controls the resistance of the device. In zero applied magnetic field, the relative alignment of the layers is controlled by the thickness of the non-magnetic spacer layer (NM).

favourable for the ferromagnetic film to be magnetized one way (in the direction in which it was cooled) rather than the other. It is sometimes said to be exchange biased by the antiferromagnetic layer. The effect of the exchange biasing is to produce a unidirectional exchange field $\mathbf{B}_{ex}$ which can act in competation with the applied field $\mathbf{B}$ (the total energy of the magnetic layer is given by $-\mathbf{M}\cdot(\mathbf{B}+\mathbf{B}_{\rm ex})$). If $\mathbf{B}$ and $\mathbf{B}_{\rm ex}$ are in the same direction, the effect is simply to shift the hysteresis loop (see Fig. 8.21). If they are at right angles, a hard axis response is obtained. In both cases, there is a unique angle of the magnetization which minimizes the energy at each value of the magnetic field.

This technique can be used to add an 'exchange bias' field to magnetoresistive sensor heads in order to bias them into their linear regions. It is also used to constrain the direction of the magnetization in one soft ferromagnetic layer in a sandwich structurc. The magnetization in the other layer can then be rotated by an applied field, allowing magnetoresistance and coupling energies to be measured in a carefully controlled way. This is also the hasis of the spinvalve (see Fig. 8.22) which is a giant magnetoresistive sensor consisting of two magnetic layers with a non-magnetic spacer, with one of the magnetic layers adjacent to an antiferromagnetic layer. The resistance across the sandwich structure is then sensitive to the value of the applied magnetic field along a particular direction.

8.9.5 Colossal magnetoresistance

The transport properties of Mn oxides have recently generated enormous interest. $LaMnO_3$ contains Mn in the Mn^{3+} state which is a Jahn-Teller ion. $LaMnO_3$ shows A-type antiferromagnetic ordering (see Section 5.2.4). If a fraction x of the trivalent La^{3+} ions are replaced by divalent Sr^{2+}, Ca^{2+} or Ba^{2+} ions, holes are introduced on the Mn sites. This results in a fraction $1-x$ of the Mn ions remaining as $Mn^{3+}\left(3\ \mathrm{d}^4, \mathrm{t}_{2g}^3\mathrm{e}_g^1\right)$ and a fraction x becoming $Mn^{4+}\left(3\ \mathrm{d}^4, \mathrm{t}_{2g}^3\mathrm{e}_g^0\right)$. When $x = 0.175$ the Jahn-Teller distortion vanishes and the system becomes ferromagnetic with a Curie temperature around room temperature. Above T_C the material is insulating and non-magnetic, but below T_C it is metallic and ferromagnetic. Particularly near T_C the material shows an extremely large magnetoresistive effect (see Fig. 8.23) which has been called colossal magnetoresistance (abbreviated to CMR), the term 'giant' having been already taken! The origin of the CMR is connected with the presence of a metal-insulator transition.

The origin of the behaviour described above is partly connected with the phenomenon of double exchange (see Section 4.2.5). In a Mn^{3+} ion, the t_{2g} electrons are tightly bound to the ion but the e_g electron is itinerant. Because of the double exchange interaction, the hopping of c_g electrons between Mn sites is only permitted if the two Mn core spins are aligned (in fact the hopping probability is proportional to $|\cos(\theta/2)|$ where θ is the angle between the two Mn core spins). The magnetic field aligns the core spins and therefore increases the conductivity, especially near T_C.

The situation is actually more complicated because the carriers interact with phonons because of the Jahn-Teller effect. The strong electron-phonon coupling in these systems implies that the carriers are actually polarons above T_C, i.e. electrons accompanied by a large lattice distortion. These polarons are

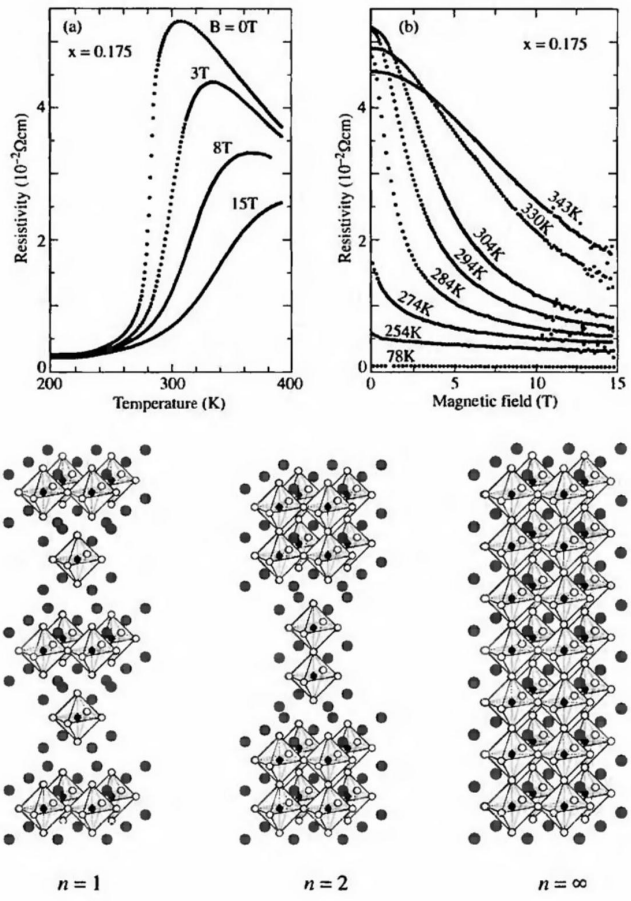


Fig. 8.23 Colossal magnetoresistance in $\mathrm{La
{1-x} |mathrm{Sr}{x} \mathrm{MnO}_{3}$ for $x = 0.175$. (a) Termperature dependence of the resistivity. (b) Isothermal magnetoresistance. After Tokura et al. 1994.}

magnetic and are self-trapped in the lattice. The transition to the magnetic state can be regarded as an unbinding of the trapped polarons. There are other signatures of the electron-phonon couplings, including magnetic-field dependent structural transitions and charge ordering. The charge-ordered ground state can compete with ferromagnetism and is enhanced near commensurate values of the doping (e.g. at $x = \frac{1}{2}$ where there is one hole for every two Mn sites). These effects are not fully understood and are under active current study.

The manganese perovskites are just one type of oxide material which shows colossal magnetoresistance. Perovskites are a member of a large family of crystals known as Ruddlesden-Popper phases (see Fig. 8.24). These phases

Fig. 8.24 The Ruddlesden-Popper phases of the manganites.

[8] Perovskite blocks are essentially stacks of corner-sharing MnO_6 octahedra.

[9] Usually this is simply called the Hall effect

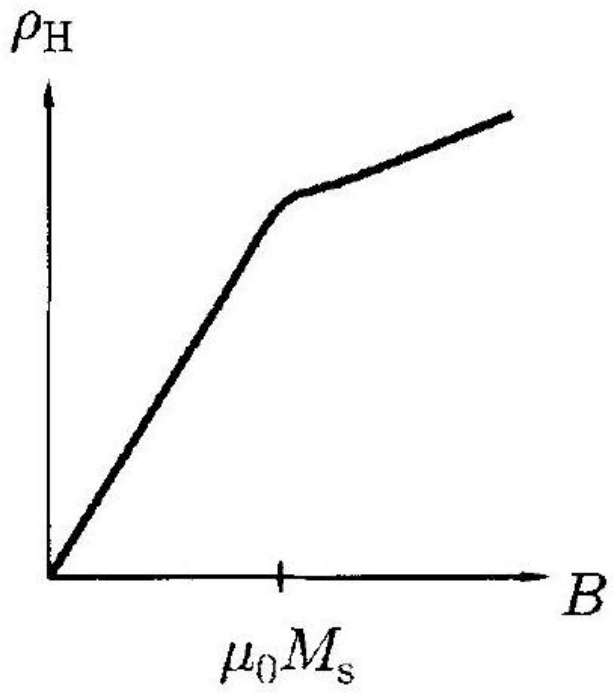


Fig. 8.25 The Hall resistivity in a ferromagnet.

have a general formula $X_{n+1}Mn_nO_{3n+1}$ (where X is a lanthanide, strontium, or mixture) and can be thought of as stacks of perovskite blocks [8] n layers thick with each block separated by a rock-salt like $(Sr, Ln)_2O_2$ layer which tends to decouple the blocks electrically and magnetically. The perovskite compounds are realized when $n = \infty$, whereas the case $n = 1$ is equivalent to the K_2NiF_4 structure, adopted by the high- T_c cuprates. Manganese compounds can be made which adopt a range of structures (e.g. $n = 1, 2, 3$ and ∞ have all been prepared at the time of writing). In the perovskites $n = \infty$, each MnO_6 octahedron is surrounded by six others. In the $n = 2$ phases, this 'coordination number' drops to five, and it is four for the $n = 1$ phases. The reduction in the number of nearest-neighbours is expected to produce an anisotropic reduction in the width of the energy bands which are derived (largely) from the Mn 3d orbitals and this modifies the electrical conductivity and magnetic behaviour of these materials. These materials can thus be carefully controlled by using different dopants, different doping levels, and adjusting the dimensionality.

8.9.6 Hall effect

A magnetic field applied normal to the current direction in a conducting material produces a transverse force on the conduction electrons in the film. This force on the conduction electrons gives rise to a transverse Hall voltage. This is known as the ordinary Hall effect [9] and is proportional to the applied magnetic field $\mathbf{B}$, because the Lorentz force on conduction electrons is $\mathbf{F} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B})$. In ferromagnets an additional effect occurs, known as the extraordinary Hall effect which depends on the magnetization. Note that it is not just a Lorentz force due to $\mathbf{M}$ rather than $\mathbf{B}$, because this is already included in the ordinary Hall effect since $\mathbf{B} = \mu_0(\mathbf{H}+\mathbf{M})$. Empirically the transverse resistivity $\rho_{\mathbf{H}}$ is given

$$\rho_H = R_0 B + \mu_0 R_e M \tag{8.12}$$

where R_o is the ordinary Hall coefficient, and R_e is the extraordinary Hall coefficient (sometimes the spontaneous Hall coefficient). The ordinary Hall coefficient R_o tends to be fairly

temperature independent, whereas R_e is usually very temperature dependent. The effect is not only seen in ferromagnets, but also in strong antiferromagnets or paramagnets. The effect only requires the presence of localized moments. (For example, it is seen in Tb in both its ferromagnetic and paramagnetic states.) If you measure the Hall resistivity as a function of field, a straight line graph results but there is a sudden change of gradient when the saturation field is reached (see Fig. 8.25). This is because at low magnetic field both B and M increase, so both ordinary and extraordinary Hall effects are seen and the gradient is due to R_o and R_e. Above the saturation field M can no longer show a further increase and the gradient results from R_o alone.
The mechanism for the extraordinary Hall effect is associated with the spinorbit interaction between the conduction electrons and localized moments. The carriers have an orbital angular momentum about the scattering centre whose sign depends on whether they pass the centre on one side or the other. This orbital angular momentum **L** couples to the spin angular momentum **S** of the scattering centre because of the spin-orbit interaction which is proportional to L. S. The electrons thus find it energetically favourable to pass to one side
rather than the other. This results in asymmetric scattering and hence to a transverse current which is responsible for the extraordinary Hall effect.

9.8.10 8.10 Organic and molecular magnets

Magnetic materials have been conventionally prepared from substances which are essentially the products of inorganic chemistry. However there has been recent interest in organic magnetic materials which are based upon organic chemistry. The tunability, resulting from the rich structure of carbon chemistry which allows many small adjustments to be made to each molecule, means that in principle materials can be tailor-made to exhibit desired properties. This is only 'in principle' however, since the structure-property relations can turn out to be remarkably intricate. The impetus to study organic materials in condensed matter physics stems particularly from the recent remarkable discoveries of three types of new material: first, conducting polymers which can have electrical conductivities as high as conventional metals such as copper and which have been used in the fabrication of polymer transistors and light-emitting-diode devices; second, superconducting charge transfer salts which can exhibit superconducting transition temperatures as high as 15 K ; third, carbon 'bucky-balls' C_{60}, a new allotrope of carbon which when appropriately doped may exhibit superconductivity or even unusual magnetic properties.

Ferromagnets are rather rare even among the elements and are exclusively found in the d or f block. To achieve a magnetic moment on an organic material, it is natural to search for an organic free radical which has unpaired spins. Many organic radicals exist, but few are stable enough to be assembled into crystalline structures. Moreover, even when that is possible, aligning these spins ferromagnetically is usually impossible. Thus the discovery of ferromagnetism, albeit at rather low temperatures, in certain nitronyl nitroxide molecular crystals in the early 1990s was particularly remarkable. The first material of this sort to be found was para-nitrophenyl nitronyl nitroxide (pNPNN) (see Fig. 8.26) which shows ferromagnetism up to $T_C \sim 0.65$ K only in one of its crystal phases. Nitronyl nitroxides contain only the elements C, H, N , and O and are therefore fully organic magnets. On each nitronyl nitroxide molecule there is an unpaired spin associated with the two $N-O$ groups. Small chemical changes to the rest of the molecule lead to significant changes in crystal structure, thereby altering the intermolecular overlaps and thus the magnetic interactions between unpaired spins on neighbouring molecules. Thus, different compounds have greatly different magnetic ground states. These materials mostly have transition temperatures below 1 K .

Some of the most technologically promising materials are molecular magnets in which a

transition metal ion provides the localized moment and organic bridges act as exchange pathways. Progress has also been achieved using materials with unpaired electrons on both the metal ions and the organic molecule. These have much higher transition temperatures than purely organic magnets, sometimes above room temperature. One of the promising aspects of these materials is their optical properties. Some of them are transparent and change colour when they become magnetic, often when this is associated with a spin transition. They therefore have a number of potential applications in cases

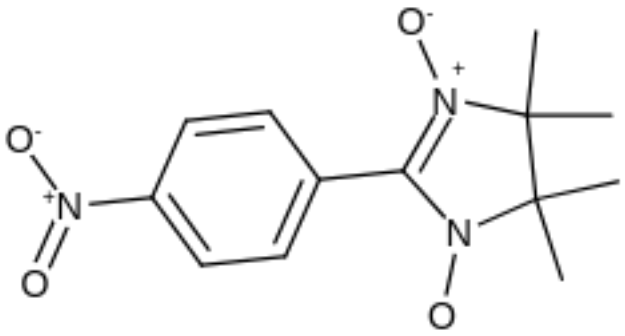


Fig. 8.26 The molecular structure of the organic ferromagnet para-nitrophenyl nitronyl nitroxide (p-NPNN).
when it is necessary to see by eye that a magnetic change has taken place (e.g. on telephone cards or display applications).

9.8.11 8.11 Spin electronics

The ability to manipulate and amplify currents of different spin-type is the basis of the emerging field of spin electronics (also known as spintronics) which is an attempt to fabricate devices out of combinations of metallic ferromagnets, non-magnetic metals and semiconductors. The ideas at the basis of this field are all connected with the fact that spin-up and spin-down electrons have different mobilities in a ferromagnet. A crucial feature is the degree of spin polarization in a given ferromagnet. The spin-polarization P is defined by

$$P = \left| \frac{n_\uparrow - n_\downarrow}{n_\uparrow + n_\downarrow} \right|, \tag{8.13}$$

and is less than unity for elemental metallic ferromagnets (0.44 for Fe, 0.34 for Co and 0.11 for Ni). However in some materials $P = 1$, implying that there is one spin-split band which is completely empty. Such materials are called halfmetallic magnets (because the electrons of one spin are metallic and those of the other spin are insulating) and CrO_2, Fe_3O_4 and some of the manganites are believed to fall into this class. Such materials have a very long spin-diffusion length, and may be exciting materials for use in spin electronics applications, for example in magnetic tunnel junctions in which spin polarized electrons are used to tunnel through an insulating barrier.

A few different designs for a spin transistor have recently appeared which attempt to use the tunnelling of electrons of different spins from ferromagnets into semiconducting layers. Conventional electronics is based only on electronic charge and ignores electronic spin. There is much current optimism that the use of magnetic materials to control the spin of electronic currents could be a promising new development for electronic devices in this century.

9.8.12 Further reading

- J. A. Mydosh, Spin glasses: an experimental introduction, Taylor and Francis 1993 and K. H. Fischer and J. A. Hertz, Spin glasses, CUP 1991, both contain useful accounts of theoretical and experimental work on spin glasses.

- E. Dagotto, Reports on Progress in Physics 62, 1525 (1999) is a readable review on spin ladders.

- A. M. Tsvelik, Quantum field theory in condensed matter physics, CUP 1995, is an advanced text which contains a description of current theories used in describing lowdimensional magnetism.

- A. Auerbach, Interacting electrons and quantum magnetism, Springer-Verlag 1994, is a lively treatment of the quantum mechanics of low-dimensional magnets and the

Haldane gap. These topics are also reviewed in I. Affleck, J. Phys.: Condensed Matter, 1, 3047 (1989).

- M. A. Kastner, R. J. Birgeneau, G. Shirane and Y. Endoh, Rev. Mod. Phys., 70, 897 (1998) contains a review of magnetic properties of single layer copper oxides and the high-T_c problem.

- E. Manousakis, Rev. Mod. Phys., 63, 1 (1991) reviews the theory of the two-dimensional Heisenberg antiferromagnet on a square lattice.

- P. M. Chaikin and T. C. Lubensky, Principles of condensed matter physics, CUP 1995, is an excellent source of information concerning the statistical mechanics of systems in condensed matter.

- Quantum phase transitions are reviewd in S. L. Sondhi, S. M. Girvin, J. P. Carini and D. Shahar, Rev. Mod. Phys., 69, 315 (1997).

- A review of magnetic multilayers with an emphasis on spin-wave excitations is R. E. Camley and R. L. Stamps. J. Phys.: Condensed Matter, 5, 3727 (1993).

- Organic magnetism is reviewed in O. Kahn, Molecular magnetism, VCH (1993).

- K. De'Bell, A. B. MacIsaac and J. P. Whitehead, Rev. Mod. Phys., 72, 225 (2000) contains an account of
dipolar effects in thin film magnets.

- The current status of various topics in magnetism research may be discerned by reading the conference proceedings of major magnetism conferences which are published from time to time in the Journal of Applied Physics, the Journal of Magnetism and Magnetic Materials and Physica B. Recent views of various aspects of magnetism research are published in the Journal of Magnetism and Magnetic Materials volume $100 \times n$ where n is an integer.

9.8.13 References

- M. N. Baibich, J. M. Broto, A. Fert, F. Nguyen van Dau, F. Petroff, P. Etienne, G. Creuzet, A. Friederich, and J. Chazelas, Phys. Rev. Lett., 61, 2472 (1988).
- D. Bitko, T. F. Rosenbaum and G. Aeppli, Phys. Rev. Lett. 77, 940 (1996).
- J. W. Bray, L. V. Interrante, I. S. Jacobs, J. C. Bonner, p. 353 of Extended linear chain compounds 3, edited by J. S. Miller, New York Plenum Press (1983).
- B. R. Coles, B. V. B. Sarkissian and R. H. Taylor, Phil. Mag. 37, 489 (1978).
- J. des Cloizeaux and J. J. Pearson, Phys. Rev. 128, 2131 (1962).
- M. Greven, R. J. Birgeneau, Y. Endoh, M. A. Kastner, B. Keimer, M. Matsuda, G. Shirane and T. R. Thurston, Phys. Rev. Lett. 72, 1096 (1994).
- K. Kanoda, Hyp. Int., 104, 235 (1997).
- M. A. Kastner, R. J. Birgeneau, G. Shirane and Y. Endoh, Rev. Mod. Phys. 70, 897 (1998).
- B. W. Lovett, S. J. Blundell, F. L. Pratt, Th. Jestdt, W. Hayes, S. Tagaki and M. Kurmoo, Phys. Rev. B, 61, 12241, (2000).
- W. Lubczynski, S. V. Demishev, J. Singleton, J. M. Caulfield, L. du Croo de Jongh, C. J. Kepert, S. J. Blundell, W. Hayes, M. Kurmoo and P. Day, J. Phys.: Condens. Matter, 8, 6005, (1996)
- N. D. Mathur, F. M. Grosche, S. R. Julian, I. R. Walker, D. M. Freye, R. K. W. Haselwinner and G. G. Lonzarich, Nature, 394, 39 (1998).
- S. Nagata, P. H. Keesom and H. R. Harrison, Phys. Rev. B 19, 1633 (1979).
- S. S. Saxena, P. Agarwal, K. Ahilan, F. M. Grosche, R. K. W. Haselwimmer, M. J. Steiner, E. Pugh, I. R. Walker, S. R. Julian, P. Monthoux, G. G. Lonzarich, A. Huxley, I. Sheikin, D. Braithwaite and J. Flouquet, Nature, 406, 587 (2000).
- D. A. Tennant, T. G. Perring, R. A. Cowley and S. E. Nagler, Phys. Rev. Lett. 70, 4003 (1993).
- Y. Tokura, A. Urushibara, Y. Moritomo, T. Arima, A. Asamitsu, G. Kido and N. Furukawa, J. Phys. Soc. Jpn., 63, 3391 (1994).

9.9 Units in electromagnetism

Units in electromagnetism represent a potential minefield of complication. SI units are used throughout this book, but much of the twentieth century literature uses non-SI notation. Therefore this appendix attempts to unravel some of the mysteries of the old system and it may be used to interpret older literature for readers who have been brought up on the new system.

From the end of the nineteenth century until just after the second world war, the cgs system was the system of units preferred by the international committees who are concerned with such things. In the cgs system, distance is measured in eentimetres, mass in grams, and time in seconds, hence the name cgs. In the Sl system, we use metres (100 times as big as centimetres), kilograms (1000 times as big as grams) and seconds. This difference in choice of

units of distance and mass appears to be rather innocent, but ends up causing great changes in other units. For example, force has the dimensions of mass × length / time ^{2}, so in cgs the unit of force, the dyne, is $1000 \times 100 = 10^5$ times smaller than a Newton, the SI unit of force. The cgs unit of energy is the erg, a dyne cm , so this is $10^5 \times 10^2 = 10^7$ times smaller than the Joule, the SI unit of energy.

Unfortunately, that is not the end of it, and when it comes to the definitions of magnetic field things get decidedly worse. This arises because the choice of definitions of the fields $\mathbf{B}, \mathbf{H}$ and $\mathbf{M}$ are different in the two systems:

$$\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M}) \quad \text{(SI)} \quad \mathbf{B} = \mathbf{H} + 4\pi\mathbf{M} \quad \text{(cgs)} \tag{A.1}$$

Thus B and H have the same dimensions in the cgs system but confusingly have different units (the former in Gauss (G), the latter in Oersted (Oe)). The difference in relation between H and M arises because the two systems make different choices in the equations defining the magnitude of the magnetic field, H, at a distance r from a magnetic monopole q :

$$H = \frac{q}{4\pi r^2} \quad \text{(SI)} \quad H = \frac{q}{r^2} \quad \text{(cgs)} \tag{A.2}$$

The omission of the factor of 4π makes this equation, and similar ones relating the magnetic field from a magnetic dipole, rather simpler in the cgs system. However the factor of 4π is not removed entirely from the theory but pops up in a more unwanted place in eqn A.1. The cgs system defined H according to non-existent monopoles from eqn A.2, so that a unit monopole produces a field of one Oersted a distance of one centimetre away (and if another unit monopole is lurking there it will experience a force of one dyne). In SI, H is more helpfully defined in terms of current (since people normally use a current source in a laboratory to produce a magnetic field, rather than devising
a suitable arrangement of unit magnetic monopoles). Thus a long solenoid with n turns per metre and carrying a current of I Amps produces a field given by $H = nI$ in the SI system.

The logic in the SI system is that the 4π appears in equations like eqn A. 2 to remind us that the field from a point magnetic monopole has spherical symmetry (the 4π is related to the surface area of a sphere, $4\pi r^2$). The same is true of the electric field E a distance r from a point charge,

$$E = \frac{q}{4\pi\epsilon_0 r^2} \quad \text{(SI)} \quad E = \frac{q}{r^2} \quad \text{(cgs)} \tag{A.3}$$

in which similar arguments apply. With SI, we have to live with a few extra factors, which fit from symmetry arguments, in equations defining the field from point charges and dipoles. The rest of the equations come out much cleaner and with fewer arbitrary factors.

The one thing to be said in favour of the cgs system is that electric and magnetic fields come out with the same dimensions, so that certain relativistic equations can look better. Also in free space $\mathbf{B} = \mathbf{H}$ which has a certain symmetry (although this can create an opportunity for confusion). But extra factors of c get littered all over the place in the cgs system, so that for example the equation for the Lorentz force on a charge moving with velocity $\mathbf{v}$ in electric and magnetic fields becomes

$$\mathbf{F} = q(\boldsymbol{\epsilon} + \mathbf{v} \times \mathbf{B}) \quad \text{(SI)} \quad \mathbf{F} = q\left(\boldsymbol{\varepsilon} + \frac{\mathbf{v} \times \mathbf{B}}{\boldsymbol{c}}\right) \tag{A.4}$$

But there is much worse to come. It turns out that there are two ways to set up electromagnetism in cgs, with either the esu (electrostatic units) or the emu (electromagnetic units) with conversion factors needed to go between them. In magnetism the emu system is usually adopted. The emu unit of magnetic moment, usually abbreviated just to the 'emu', is a unit

with dimensions of volume (which in egs is cm^3). Magnetic susceptibility, $\chi = M/H$, is dimensionless in both systems, but in cgs is often written as emu cm^{-3}, which is still dimensionless even if it doesn't appear so at first glance. There is always the troublesome factor of 4π to remember when converting between cgs susceptibility and SI susceptibility which arises because $\mathbf{B} = \mathbf{H} + 4\pi\mathbf{M}$ in the cgs system (see above). For example, the demagnetization factor, N , is dimensionless and is $0 < \mathrm{N} < 1$ in SI but $0 < \mathrm{N} < 4\pi$ in cgs. Unfortunately the emu is still widely used in some parts of magnetism research. Table A summarizes the most useful results for converting SI to and from cgs.

Though not an SI unit, the Bohr magneton $\mu_\mathrm{B} = 9.274 \times 10^{-24}$ J T^{-1} is a useful measure of magnetic moment since it corresponds to the magnetic moment of a 1 s electron in hydrogen. For a paramagnet, the molar susceptibility χ_m is given by Curie's law (adapting eqn 2.44) which is in SI units

$$\chi_\mathrm{m} = \frac{\mu_0 N_\mathrm{A} \mu_\mathrm{eff}^2 \mu_\mathrm{B}^2}{3 k_\mathrm{B} T} \tag{A.5}$$

where N_A is Avogadro's number. Hence $\chi_\mathrm{m} T$ is independent of temperature (see Fig. 2.10) and this can be related to the effective moment. Hence by rearranging eqn A.5, one has

$$\mu_\mathrm{eff} = \left[3 k_\mathrm{B} / \mu_0 N_\mathrm{A} \mu_\mathrm{B}^2\right]^{1/2} \sqrt{\chi_\mathrm{m} T} \tag{A.6}$$

Table A. 1 Units in the SI system and the cgs system. The abbreviations are $\mathrm{m =}$ metre, g = gramme, N = Newton, J = Joule, T = Tesla, G = Gauss, A = Amp, Oe = Oersted, Wb = Weber, Mx = Maxwell. The term emu is short for electromagnetic unit. Note that magnetic susceptibility is dimensionless in SI units.}

Quantity	symbol		SI unit		cgs unit
Length	x	10^{-2}	m	$= 1$	cm
Mass	m	10^{-3}	kg	$= 1$	g
Force	F	10^{-5}	N	$= 1$	dyne
Energy	E	10^{-7}	J	$= 1$	erg
Magnetic induction	B	10^{-4}	T	$= 1$	G
Magnetic field strength	H	$10^3/4\pi$	Am^{-1}	$= 1$	De
Magnetic moment	μ	10^{-3}	JT^{-1} or Am^2	$= 1$	erg $\mathbf{G}^{-1}$
Magnetization (= moment per volume)	M	10^3	Am^{-1} or JT^{-1} m^{-3}	$= 1$	Oe
Magnetic susceptibility	χ	$\mathbf{4\pi}$		$= 1$	emu cm^{-3}
Molar susceptibility	χ_m	$4\pi \times 10^{-6}$	m^3 mol^{-1}	$= 1$	emu mol^{-1}
Mass susceptibility	χ_g	$4\pi \times 10^{-3}$	m^3 kg^{-1}	$= 1$	emu g^{-1}
Magnetic flux	ϕ	10^{-8}	Tm^2 or Wb	$= 1$	Gcm^{-2} or M
Demagnetization factor	N		$0 < \mathrm{N} < 1$		$0 < N < 4\pi$

so that

$$\mu_\mathrm{eff} = 797.8 \sqrt{\chi_\mathrm{m}^\mathrm{SI} T} \tag{A.7}$$

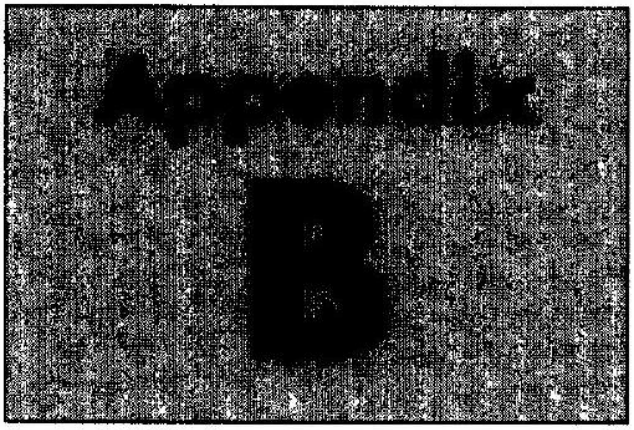
$$\mu_\mathrm{eff} = 2.827 \sqrt{\chi_\mathrm{m}^\mathrm{cgs} T} \tag{A.8}$$

where μ_eff is measured in Bohr magnetons per formula unit, χ_m^SI is measured in m^3 mol^{-1}, and $\chi_\mathrm{m}^\mathrm{cgs}$ is measured in emu mol^{-1}. These numerical relationships can be useful for extracting effective moments from graphs of $\chi_\mathrm{m} T$ against T.

9.9.1 Further reading

- B. I. Bleaney and B. Bleaney, Electricity and magnetism, OUP 1989.
- J. D. Jackson, Classical electrodynamics, Wiley 1962.

9.10 Electromagnetism



In this appendix several key results in electromagnetism are briefly reviewed. The further reading should be consulted for detailed derivations of these results.

9.10.1 B. 1 Magnetic moments

The Lorentz force d$\mathbf{F}$ on a straight segment of wire of length dr carrying current I in a magnetic field of flux density $\mathbf{B}$ is given by $\mathrm{d}\mathbf{F} = I\ \mathrm{d}\mathbf{r} \times \mathbf{B}$ (see Fig. B.1).

This result can now be applied to calculate the couple on the elementary loop of wire in Fig. B.2. The couple is given by

$$\begin{aligned} \mathrm{d}G_x &= -I\ \mathrm{d}x\ \mathrm{d}yB_y \\ \mathrm{d}G_y &= I\ \mathrm{d}x\ \mathrm{d}yB_x, \end{aligned} \tag{B.1}$$

so that defining the magnetic moment of the loop by $\mathrm{d}\mu_z = I\ \mathrm{d}x\ \mathrm{d}y$, and repeating the argument for arbitrary directions of the loop, the couple dG is given by

$$\mathrm{d}\mathbf{G} = \mathrm{d}\boldsymbol{\mu} \times \mathbf{B}. \tag{B.2}$$

For a loop of finite size, this yields by integration

$$\mathbf{G} = \boldsymbol{\mu} \times \mathbf{B}. \tag{B.3}$$

There will be a net force on the loop only if the magnetic field is non-uniform. Thus the force in the x direction on the elementary loop in Fig. B. 2 is

$$\mathrm{d}F_x = I\ \mathrm{d}y \left(B_z + \frac{\partial B_z}{\partial x}\ \mathrm{d}x \right) - I\ \mathrm{d}yB_z = \mathrm{d}\mu_z \frac{\partial B_z}{\partial x}, \tag{B.4}$$

so that

$$\mathrm{d}\mathbf{F} = (\mathrm{d}\boldsymbol{\mu} \cdot \nabla)\mathbf{B} \tag{B.5}$$

and by integrating for a magnetic moment of finite size,

$$\mathbf{F} = (\boldsymbol{\mu} \cdot \nabla)\mathbf{B}. \tag{B.6}$$

It is therefore only a non-uniform magnetic field which can exert a force on a magnetic moment.
A magnetic moment $\boldsymbol{\mu}$ aligned with a magnetic field $\mathbf{B}$ has zero torque on it (by eqn B.3, $\mathbf{G} = 0$ if $\boldsymbol{\mu}$ and $\mathbf{B}$ are parallel). Rotating $\boldsymbol{\mu}$ with respect to $\mathbf{B}$
B. 1 Magnetic moments 198
B. 2 Maxwell's equations in free space 199
B. 3 Free and bound currents 200
B. 4 Maxwell's equations in matter 201
B. 5 Boundary conditions 201

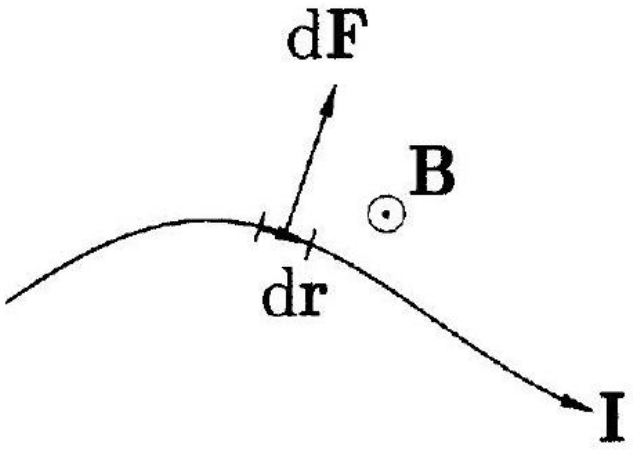


Fig. B. 1 The force on a current element is $\mathrm{dF =I \mathrm{dr} \times \mathrm{B}}$.}

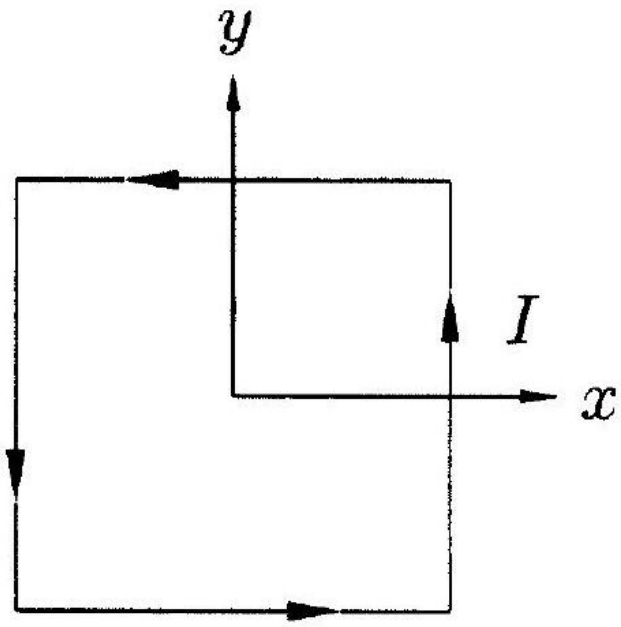


Fig. B. 2 A square elementary current loop (with sides of length $\mathrm{d x}$ and dy) oriented perpendicular to the z-direction.}

This argument can be generalized to show that

$$\delta W = -\boldsymbol{\mu} \cdot \delta \mathbf{B}.$$

James Clerk Maxwell (1831-1879)

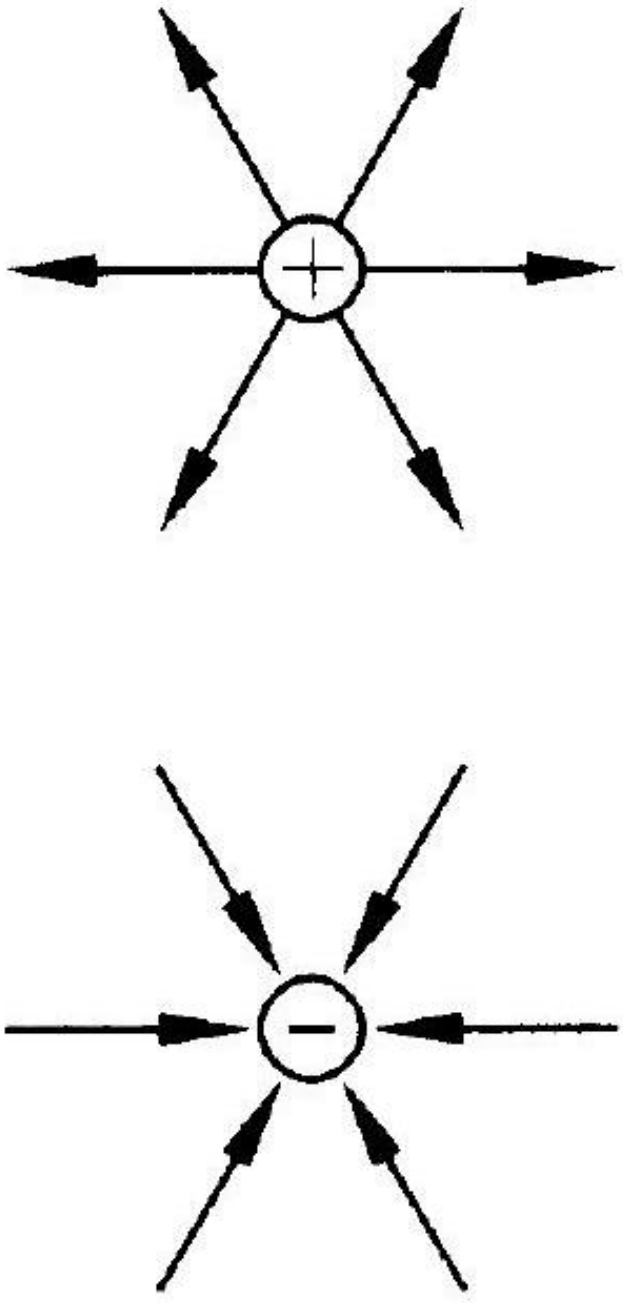


Fig. B. 3 Electric field diverges away from positive charges and converges into negative charges.

therefore can only be accomplished by applying a torque. The work done W by the torque to rotate μ to an angle θ with respect to the magnetic field is

$$W = \int_0^\theta G \, \mathrm{d}\theta' = \int_0^\theta \mu B \sin\theta' \mathrm{d}\theta' = \mu B(1 - \cos\theta). \tag{B.7}$$

The potential energy U of the magnetic moment can therefore be written as $-\mu B \cos\theta$ (ignoring the constant term). Hence we can write

$$U = -\boldsymbol{\mu} \cdot \mathbf{B} \tag{B.8}$$

A magnetic moment μ in a non-uniform field $\mathbf{B}$ has a force on it in the x direction equal to $\mu\partial B/\partial x$ by eqn B.6. To stop it moving in the x direction, one would have to apply a restoring force $-\mu\partial B/\partial x$. If it were allowed to move a distance δx into a region of stronger field, the work done by this force would be

$$\mathrm{d}W = -\mu\frac{\partial B}{\partial x}\delta x = -\mu\delta B \tag{B.9}$$

and this is an increment to the sample's free energy (in agreement with the results of Exercise 2.10).

The magnetic field due to a magnetic moment at the origin is

$$\mathbf{B}(\mathbf{r}) = -\mu_0\nabla\left(\frac{\mu\cdot\mathbf{r}}{4\pi r^3}\right) \tag{B.10}$$

and the magnetic vector potential (in the most convenient choice of gauge) is

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi}\frac{\boldsymbol{\mu}\times\mathbf{r}}{r^3}. \tag{B.11}$$

9.10.2 B. 2 Maxwell's equations in free space

Maxwell's equations in free space are

$$\nabla\cdot\boldsymbol{\varepsilon} = \rho/\epsilon_0 \tag{B.12}$$

$$\nabla\cdot\mathbf{B} = 0 \tag{B.13}$$

$$\nabla\times\boldsymbol{\varepsilon} = -\frac{\partial\mathbf{B}}{\partial t} \tag{B.14}$$

$$\nabla\times\mathbf{B} = \mu_0\mathbf{J} + \epsilon_0\mu_0\frac{\partial\boldsymbol{\varepsilon}}{\partial t}, \tag{B.15}$$

and describe the interrelationships between the electric field $\boldsymbol{\varepsilon}$, the magnetic induction $\mathbf{B}$, the charge density ρ and the current density $\mathbf{J}$. Equation B. 12 shows that electric field diverges away from positive charges and converges into negative charges; charge density therefore acts as a source or a sink of electric field (see Fig. B.3). Equation B. 13 shows that magnetic fields have no such divergence; thus there are no magnetic charges (monopoles) and lines of $\mathbf{B}$ field must just exist in loops; they can never start or stop anywhere. Equation B. 14 shows that you only get loops of electric field around regions in space in which there is a changing magnetic field. This leads to Faraday's law of electromagnetic induction. Equation B.15, in the absence of a changing electric field, shows that loops of magnetic induction are found around electric currents. Note that everything in this section has been for electromagnetic fields in free space.

9.10.3 B. 3 Free and bound currents

Once matter is brought into the picture, things get a little more complicated and it is worth trying to understand why. The treatment here will just consider the effects for magnetism, although analogous thinking lies behind the corresponding derivations for electric fields. Real matter contains a different type of electric current from the familiar currents we are used to thinking about flowing down wires. Atoms contain microscopic bound currents due to the electrons circulating around the nucleus. Thus the total current $\mathbf{J}$ in eqn B. 15 can be separated into two components

$$\mathbf{J} = \mathbf{J}_{\text{free}} + \mathbf{J}_{\text{bound}} \tag{B.16}$$

due to the free currents (10 amps flowing down a copper wire) and bound currents (electrons orbiting an atom).

If $\mathbf{M}$ is uniform in a sample, the only net bound current that flows is around the edge of the material. If $\mathbf{M}$ is non-uniform, bound currents will be generated inside the material. We need a general relation between the bound currents and the magnetization, and this is that

$$\mathbf{J}_{\text{bound}} = \nabla \times \mathbf{M} \tag{B.17}$$

A derivation of eqn B. 17 will now be sketched (for more details see the further reading). The vector potential $\mathbf{A}$ due to a single point dipole at the origin is given in eqn B.11. Now consider a magnetized specimen, contained in volume V, with magnetization $\mathbf{M}(\mathbf{r}')$ at position $\mathbf{r}'$. The vector potential at position $\mathbf{r}$ is

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int_V \mathrm{d}^3 r' \frac{\mathbf{M}(\mathbf{r}') \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3} \tag{B.18}$$

This equation can be simplified using $\nabla_r(1/r) = -\mathbf{r}/r^3$, the vector identity

$$\nabla \times (f\mathbf{g}) = f\nabla \times \mathbf{g} - \mathbf{g} \times \nabla f \tag{B.19}$$

where f and $\mathbf{g}$ are scalar and vector functions respectively, and transforming the integral of $\mathbf{r}'$ to one over $\mathbf{r}$. The result is that the vector potential at $\mathbf{r}$ can be written as

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int_V \frac{\nabla \times \mathbf{M} \mathrm{d}\tau}{r} + \frac{\mu_0}{4\pi} \int_S \nabla \times \left(\frac{\mathbf{M}}{\mathbf{r}}\right) d\tau \tag{B.20}$$

$$= \frac{\mu_0}{4\pi} \int_V \frac{\nabla \times \mathbf{M} \mathrm{d}\tau}{r} + \frac{\mu_0}{4\pi} \int_V \frac{\mathbf{M} \times \mathrm{d}\mathbf{S}}{r} \tag{B.21}$$

$$= \frac{\mu_0}{4\pi} \int_V \frac{\mathbf{J}_{\text{bound}}\, \mathrm{d}\tau}{r} + \frac{\mu_0}{4\pi} \int_S \frac{\mathbf{K}_{\text{bound}}\, \mathrm{d}S}{r} \tag{B.22}$$

where the first integral is over the volume of the sample, and the second integral is over the surface of the sample. The bound current in eqn B. 22 agrees with eqn B.17. The surface current is given by $\mathbf{K}_{\text{bound}} = \mathbf{M} \times \hat{\mathbf{n}}$ and the vector $\hat{\mathbf{n}}$ is perpendicular to the surface.

The problem with eqn B. 15 is that the curl of $\mathbf{B}$ is related not only to the free current density but also to the more inconvenient bound current density. Therefore we are motivated to define a new 'corrected' field which only 'sees'

[1] These are traditionally called 'pill-boxes' in the physics literature, although pills have long since stopped being sold in such shaped boxes!

[2] And we can make sure this is the case by making sure that we let the box thickness tend to zero before its radius tends to zero.

the free current density, although as shown below this will be only partially successful. Using the result for the bound volume currents, we define a new field $\mathbf{H}$ (called the magnetic field strength, or just the H field) by

$$\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M}) \tag{B.23}$$

and hence we can write a new version of eqn B. 15 which includes only $\mathbf{J}_{\text{free}}$. Thus in the absence of changing electric fields

$$\nabla \times \mathbf{H} = \mathbf{J}_{\text{free}}\,. \tag{B.24}$$

This is useful because it relates the **H**-field to a quantity that can be read on an ammeter (which J_{bound} cannot be). Unfortunately this procedure, having simplified one of Maxwell's equation, complicates another. Equation B. 13 changes from the simple $\nabla \cdot \mathbf{B} = 0$ to the slightly more tortuous

$$\nabla \cdot \mathbf{H} = -\nabla \cdot \mathbf{M}. \tag{B.25}$$

Thus **H**, unlike **B**, is not divergence-free (or solenoidal, to use the mathematical jargon). **H** behaves as if magnetic monopoles (i.e. the magnetic equivalent of charges) exist.

9.10.4 B. 4 Maxwell's equations in matter

Maxwell's equations in the presence of matter therefore become

$$\nabla \cdot \mathbf{D} = \rho_{\text{free}} \tag{B.26}$$
$$\nabla \cdot \mathbf{B} = 0 \tag{B.27}$$
$$\nabla \times \boldsymbol{\varepsilon} = -\frac{\partial \mathbf{B}}{\partial t} \tag{B.28}$$
$$\nabla \times \mathbf{H} = \mathbf{J}_{\text{free}} + \frac{\partial \mathbf{D}}{\partial t}, \tag{B.29}$$

where a similar distinction is made between free and bound charges and where the electric displacement, **D**, is given by

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \tag{B.30}$$

and **P** is the electric polarization.

9.10.5 B. 5 Boundary conditions

A significant difference between the B and H fields is revealed when they meet a boundary between two different regions. The boundary conditions between two regions 1 and 2 (see Fig. B.4) can be derived as follows. Consider the shoe-polish-tin-shaped box [1] straddling the surface. $\nabla \cdot \mathbf{B} = 0$ implies that $\int_S \mathbf{B} \cdot \, \mathrm{d}\mathbf{S} = 0$ using the divergence theorem. Since the surface area of the box is mostly represented by the top and bottom surfaces of the box, [2] we have that the perpendicular component of **B** is continuous, i.e. that

$$B_{\perp 1} = B_{\perp 2}. \tag{B.31}$$

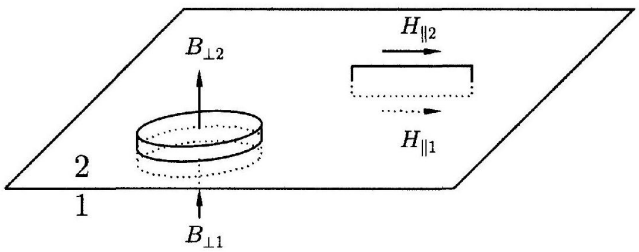


Fig. B. 4 Electromagnetic boundary conditions between regions 1 and 2.

The argument won't work for **H** because $\nabla \cdot \mathbf{H} = -\nabla \cdot \mathbf{M}$ which may be nonzero. The perpendicular component may be discontinuous if the magnetization diverges inside the box

(which it may well do if region 1 is magnetized and region 2 is not). The argument for the parallel components comes from the equation $\nabla \times \mathbf{H} = \mathbf{J}_{\text{free}}$ which implies that $\oint \mathbf{H} \cdot \mathrm{d}\mathbf{l} = \int \mathbf{J}_{\text{free}} \cdot \mathrm{dS}$. As the Amperean loop shown on the right-hand side of Fig. B. 4 shrinks to zero, so does the free current enclosed by the loop and so the parallel component of $\mathbf{H}$ is continuous, i.e. that

$$H_{\parallel 1} = H_{\parallel 2}. \tag{B.32}$$

This argument won't work for $\mathbf{B}$ because $\nabla \times \mathbf{B} = \mu_0 (\mathbf{J}_{\text{free}} + \mathbf{J}_{\text{bound}})$ and so the Amperean loop, although enclosing no free current as it shrinks to zero, can enclose some bound surface current in the limit that the loop shrinks to zero.

9.10.6 Further reading

- B. I. Bleaney and B. Bleaney, Electricity and magnetism, OUP 1989.
- D. J. Griffiths, Introduction to electrodynamics, Prentice-

  Hall 1989.

- J. D. Jackson, Classical electrodynamics, John-Wiley 1962.
  C. 1 Quantum mechanics 203
  C. 2 Dirac bra and ket notation 204
  C. 3 The Bohr model 205
  C. 4 Orbital angular momentum 206
  C. 5 The hydrogen atom 207
  C. 6 The g-factor 208
  C. 7 d orbitals 210
  C. 8 The spin-orbit interaction 211
  C. 9 Landé g-factor 211
  C. 10 Perturbation theory 212

9.11 Quantum and atomic physics

9.11.1 C. 1 Quantum mechanics

In this section we briefly review some basic features of quantum mechanics. The fundamental object in quantum mechanics is the wave function (also known as the state function) ψ which contains all the knowledge one obtains about a system from observation. The wave function ψ is a complex function and $|\psi|^2 = \psi\psi^*$ is the probability density. [1] To every observable quantity, there corresponds a Hermitian operator. When you make a measurement, the result you obtain is one of the eigenvalues of that operator. Hermitian operators have real eigenvalues and eigenfunctions which are orthogonal to each other. If the i^{th} eigenfunction of the operator $\hat{A}$ (operators are given 'hats') is ϕ_i (assumed normalized) and has an eigenvalue a_i, then

$$\hat{A}\phi_i = a_i\phi_i. \tag{C.1}$$

The expected value that you obtain after a measurement of operator $\hat{A}$ is given by

$$\langle \hat{A} \rangle = \int \mathrm{d}\tau \psi^* \hat{A} \psi \tag{C.2}$$

[01] Here the symbol * denotes complex conjugation.

where $d\tau$ is a volume element. The wave function ψ can be expanded in terms of the eigenfunctions of $\hat{A}$, i.e.

$$\psi = \sum_i c_i \phi_i, \tag{C.3}$$

so that $\langle \hat{A} \rangle$ is then given by

$$\langle \hat{A} \rangle = \sum_i |c_i|^2 a_i, \tag{C.4}$$

using the orthonormality of ϕ_i. The probability of obtaining the i^{th} eigenvalue, a_i, is $|c_i|^2$. Measurement is a drastic process. If you take a wave function ψ, measure the physical quantity which is associated with operator $\hat{A}$ and obtain the result a_n (the n^{th} eigenvalue of $\hat{A}$), the system is forced into state function ϕ_n. This, at least, is the Copenhagen interpretation.

The commutator of two operators $\hat{A}$ and $\hat{B}$ is defined by $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$. An expression defining a commutator is known as a commutation relation.

Example C. 1

The commutator $[\hat{x}, \hat{p}] = \mathrm{i}\hbar$ where $\hat{x}$ and $\hat{p}$ are the position and momentum operators respectively.

If the two operators $\hat{A}$ and $\hat{B}$ commute, ($[\hat{A}, \hat{B}] = 0$), they are said to be compatible and measurement of one does not affect the value of the other in any way. If they do not commute, there exists an uncertainty relation between them so that

$$\Delta A \Delta B \geq \frac{1}{2} |\langle i[\hat{A}, \hat{B}] \rangle|, \tag{C.5}$$

where $\Delta A \equiv \sqrt{\left\langle (\hat{A} - \langle \hat{A} \rangle)^2 \right\rangle}$.

The time dependence of ψ in the absence of a measurement is given by the Schrödinger equation

$$\hat{\mathcal{H}}\psi = \mathrm{i}\hbar \frac{\mathrm{d}\psi}{\mathrm{d}t} \tag{C.6}$$

where $\hat{\mathcal{H}}$ is the Hamiltonian. The time dependence of the expected value of an operator is given by

$$\frac{\mathrm{d}\langle \hat{A} \rangle}{\mathrm{d}t} = \frac{1}{\mathrm{i}\hbar} \langle [\hat{A}, \hat{\mathcal{H}}] \rangle, \tag{C.7}$$

if the operator $\hat{A}$ itself is not time dependent. An observable whose operator commutes with the Hamiltonian is a conserved quantity and is known as a constant of the motion or a good quantum number.

9.11.2 C. 2 Dirac bra and ket notation

In this section we review some notation, due to Paul Dirac, which is used sporadically in this text and often in the magnetism literature. The scalar product between two vectors $\mathbf{a}$ and $\mathbf{b}$ is given by $\mathbf{a} \cdot \mathbf{b} = \sum_i a_i b_i$. The length of $\mathbf{a}$ is given by $\sqrt{\mathbf{a} \cdot \mathbf{a}}$ which must be real and positive. The notation makes the inner product look completely symmetric between $\mathbf{a}$ and $\mathbf{b}$, but this hides some subtlety. If the vectors are complex, the scalar product should be written

$$\mathbf{a}^\dagger \cdot \mathbf{b} = \sum_i a_i^* b_i = \begin{pmatrix} a_1^* & a_2^* & a_3^* \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}. \tag{C.8}$$

This definition ensures that $\mathbf{a}^\dagger \cdot \mathbf{a} = \sum_i |a_i|^2$ is real and positive, and so the length of the complex vector is well defined. Equation C. 8 emphasizes that the two vectors in the scalar product should not really be considered on an equal footing: one is a row vector, the other a column vector; one is complex conjugated, the other is not. Mathematicians say that actually the two vectors in a scalar product 'live' in different spaces, so that here $\mathbf{b}$ lives in a vector space and a lives in its 'dual space'. To turn a vector into its dual, you need

Paul A. M. Dirac (1902-1984)

to take the adjoint, signified by a † sign (it means 'complex conjugate and transpose'), so that

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}^\dagger = \begin{pmatrix} a_1^* & a_2^* & a_3^* \end{pmatrix} \tag{C.9}$$

However in quantum mechanics, the state function is sometimes a vector, like $\mathbf{a} = \binom{a_1}{a_2}$, and sometimes a continuous function $a(x)$. If a function, then the scalar product between $a(x)$ and $b(x)$ is

$$a^\dagger \cdot b = \int \mathrm{d}x a^*(x) b(x) \tag{C.10}$$

Dirac circumvented this notational problem by writing the state function as a ket: $|a\rangle$, whether he was dealing with finite-component vectors or functions (which can be regarded as infinite dimensional vectors). This is just notation, so that $|a\rangle + |b\rangle = |c\rangle$ is simply another way of writing $\mathbf{a} + \mathbf{b} = \mathbf{c}$, or equivalently $\vec{a} + \vec{b} = \vec{c}$ or $a(x) + b(x) = c(x)$. One can turn the ket $|a\rangle$ into a bra, $\langle a|$, by using the adjoint

$$|a\rangle^\dagger = \langle a|. \tag{C.11}$$

The scalar product between $\langle a|$ and $|b\rangle$ is then written as

$$\langle a \mid b \rangle = \sum_i a_i^* b_i = \int \mathrm{d}x a^*(x) b(x) \tag{C.12}$$

a bra-(c)-ket (Dirac here demonstrating evidence of humour). The expression for an expectation value of an operator can be written

$$\langle \hat{A} \rangle = \langle \psi | \hat{A} | \psi \rangle \tag{C.13}$$

The notation conveys the asymmetry of the scalar product, is widely used and greatly simplifies many computations. It is particularly useful in labelling particular states, as in the following:

$$\mid \text{ Schrödinger's cat } \rangle = \frac{\mid \text{ alive } \rangle + \mid \text{ dead } \rangle}{\sqrt{2}}. \tag{C.14}$$

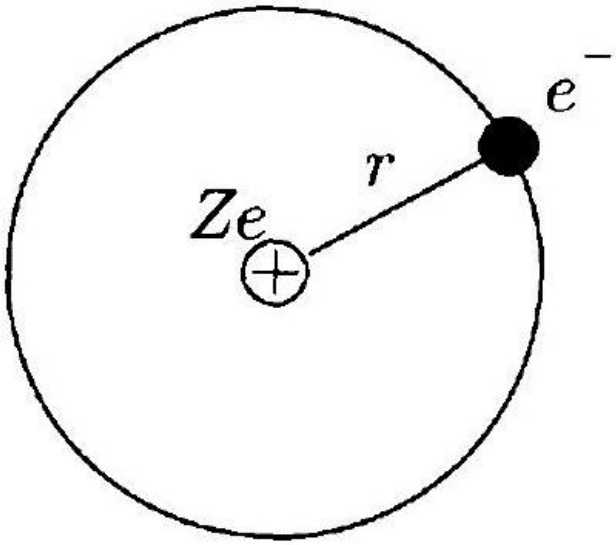


Fig. C. 1 The Bohr model.

9.11.3 C. 3 The Bohr model

The Bohr model is a semiclassical model describing the motion of an electron in a hydrogen atom. It gives the right results, but for the wrong reason. Nevertheless, it is simple and therefore useful for quickly deriving results such as the dependence of the radius of an atomic orbit on nuclear charge Z and effective mass.

Consider an electron orbiting a nucleus of charge $+Ze$ (see Fig. C.1). Equating electrostatic and centripetal forces yields

$$\frac{Ze^2}{4\pi\epsilon_0 r^2} = \frac{m_e v^2}{r} \tag{C.15}$$

The angular momentum of the electron $m_c vr$ is quantized in units of $\hbar$ so we can write

$$m_{\mathrm{e}} vr = n\hbar, \tag{C.16}$$

where n is an integer (the principal quantum number). Hence the radius of the orbit can be written

$$r = \frac{4\pi\epsilon_0\hbar^2 n^2}{Ze^2 m_e} = \left(\frac{4\pi\epsilon_0\hbar^2}{m_e e^2}\right)\frac{n^2}{Z} = \frac{a_0 n^2}{Z}, \tag{C.17}$$

where a_0 is the Bohr radius defined by

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_{\mathrm{e}} e^2} \tag{C.18}$$

The energy E is then given by

$$E = \frac{1}{2}mv^2 - \frac{Ze^2}{4\pi\epsilon_0 r} \tag{C.19}$$

$$= -\frac{1}{2}\left(\frac{Ze^2}{4\pi\epsilon_0 r}\right) \tag{C.20}$$

$$= -\frac{\hbar^2}{2m_\mathrm{e}}\frac{1}{a_0^2}\frac{Z^2}{n^2} \tag{C.21}$$

$$= -\frac{m_\mathrm{e}e^4}{(4\pi\epsilon_0\hbar)^2}\frac{Z^2}{n^2} \tag{C.22}$$

9.11.4 C. 4 Orbital angular momentum

The angular momentum operator is given by

$$\hbar\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}} = -\mathrm{i}\hbar\hat{\mathbf{r}} \times \nabla \tag{C.23}$$

The z component of the angular momentum thus has an operator

$$\hbar\hat{L}_z = \mathrm{i}\hbar\left[y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right] \tag{C.24}$$

$$= -\mathrm{i}\hbar\frac{\partial}{\partial\phi} \tag{C.25}$$

which has eigenfunctions given by $\mathrm{e}^{\mathrm{i}m_l\phi}$ and eigenvalues $m_l\hbar$, where m_l is the magnetic quantum number. The following commutation relations can be easily proved:

$$\left[\hat{L}_i, \hat{L^2}\right] = 0 \tag{C.26}$$

and

$$\left[\hat{L}_i, \hat{L}_j\right] = \mathrm{i}\epsilon_{ijk}\hat{L_k} \tag{C.27}$$

where the alternating tensor ϵ_{ijk} is defined by

$$\epsilon_{ijk} = \begin{cases} 1 \text{ if } ijk \text{ is an even permutation of } 123 \\ -1 \text{ if } ijk \text{ is an odd permutation of } 123 \\ 0 \text{ if any two of } i, j \text{ or } k \text{ are equal.} \end{cases} \tag{C.28}$$

These results can be written as $a_0 = \frac{h}{m\alpha c}$ and $E = -\frac{1}{2}m(\alpha c)^2\frac{Z^2}{n^2}$ where α is the fine structure constant.

Equation C. 27 is shorthand for

$$|\ \hat{L}_x \cdot \hat{L}_y\Big] = \mathrm{i}\hat{L}_z$$

and cyclic permutations. The equation uses Einstein's summation convention in which any twice-repeated index is assumed summed. For example a_1b_i is short for $\sum_i a_ib_i$. The alternating tensor is useful in expressions such as

$$(\mathbf{a} \times \mathbf{b})_i = \epsilon_{ijk}a_jb_k$$

In this equation j and k are twice-repeated and thus assumed summed. Hence for example

$$(\mathbf{a}\times\mathbf{b})_1 = \epsilon_{123}a_2b_3 + \epsilon_{132}a_3b_2$$
$$= a_2b_3 - a_3b_2$$

Table C.1 $P_{l

ˆ{\left|m_{l}\right|}(\cos \theta)$.}

	$m_l = 0$	$m_l = 1$	$m_l = 2$	$m_l = 3$
$l = 0$	1	-	-	-
$l = 1$	$\cos\theta$	$\sin\theta$	-	-
$l = 2$	$\frac{1}{2}(3\cos^2\theta - 1)$	$\frac{1}{3}\sin\theta\cos\theta$	$\frac{1}{3}\sin^2\theta$	-
$l = 3$	$\frac{1}{2}(5\cos^3\theta - 3\cos\theta)$	$\frac{2}{3}\sin\theta(5\cos^2\theta - 1)$	$\frac{1}{15}\sin^2\theta\cos\theta$	$\frac{1}{15}\sin^3\theta$

The operator $\hat{L^2}$ has eigenfunctions

$$|l, m_l\rangle = Y_{lm_l}(\theta,\phi) \propto P_l^{m_l}(\cos\theta)\mathrm{e}^{\mathrm{i}m_l\phi}, \tag{C.29}$$

known as spherical harmonics, with eigenvalues $l(l+1)$ where l is the angular momentum quantum number and $P_l^{m_l}(\cos\theta)$ is an associated Legendre polynomial (see Table C.1).

Thus

$$\hat{L^2}\,|l, m_l\rangle = l(l+1)\,|l, m_l\rangle\,. \tag{C.30}$$

Similarly

$$\hat{L_z}\,|l, m_l\rangle = m_l\,|l, m_l\rangle\,. \tag{C.31}$$

The raising and lowering operators $\hat{L_\pm}$are defined by

$$\hat{L_\pm} = \hat{L_x} \pm \mathrm{i}\hat{L_y}. \tag{C.32}$$

Hence one can show that

$$\hat{L_\pm}\,|l, m_l\rangle = \sqrt{l(l+1) - m_l\,(m_l \pm 1)}\,|l, m_l \pm 1\rangle\,. \tag{C.33}$$

Certain states have no orbital angular momentum. Singlet states ($S = 0$) are a good example, and this arises because they have real wave functions, as can be proved straightforwardly. In the absence of a magnetic field the Hamiltonian is real so that if ψ is an eigenfunction of the Hamiltonian $\hat{\mathcal{H}}$ with energy E, then so is ψ^*. ($\hat{\mathcal{H}}\psi = E\psi$ so $\hat{\mathcal{H}}\psi^* = E\psi^*$.) But since the state is by definition a singlet, $\psi = \psi^*$ and ψ is then real. The operators for all components of $\hat{\mathbf{L}}$ contain i so that the expectation values $\langle\psi|\hat{L}_\alpha|\psi\rangle$ for $\alpha = x, y, z$ are all purely imaginary if ψ is real. But expectation values are measurable quantities and therefore must be pure real; hence they must all be zero and the state has no orbital angular momentum.

9.11.5 C. 5 The hydrogen atom

Most of the atoms considered in this book are not hydrogen, but the Schrödinger equation can only be solved exactly for the hydrogen atom. In this section, some results concerning the hydrogen atom are tabulated for easy reference.

The Schrödinger equation for a spherically symmetrical potential can be solved by separation of variables and the eigenfunction Ψ can be written as a

Table C. 2 Radial wave functions.

n	l	$L_{n+l}^{2l+1}(x)$	$R_{nl}(\rho)$
1	0	$L_1^1(x) = -1$	$2\left(Z/a_0\right)^{3/2} \mathrm{e}^{-\rho/2}$
2	0	$L_2^1(x) = 2x - 4$	$(1/2\sqrt{2})\left(Z/a_0\right)^{3/2} (2-\rho)\mathrm{e}^{-\rho/2}$
2	1	$L_3^3(x) = -6$	$(1/2\sqrt{6})\left(Z/a_0\right)^{3/2} \rho\mathrm{e}^{-\rho/2}$
3	0	$L_3^1(x) = -3x^2 + 18x - 18$	$(1/9\sqrt{3})\left(Z/a_0\right)^{3/2} (6 - 6\rho + \rho^2)\,\mathrm{e}^{-\rho/2}$
3	1	$L_4^3(x) = 24x - 96$	$(1/9\sqrt{6})\left(Z/a_0\right)^{3/2} (4-\rho)\rho\mathrm{e}^{-\rho/2}$
3	2	$L_5^5(x) = -120$	$(1/9\sqrt{30})\left(Z/a_0\right)^{3/2} \rho^2\mathrm{e}^{-\rho/2}$

product of radial and angular parts:

$$\Psi = R(r)\Theta(\theta)\Phi(\phi) \tag{C.34}$$

The equation for $\Phi(\phi)$ is

$$\frac{\mathrm{d}^2\Phi}{\mathrm{d}\phi^2} + m_l^2\Phi = 0 \tag{C.35}$$

which can be readily solved to yield solutions of the form

$$\Phi = A\mathrm{e}^{\mathrm{i}m_l\phi}. \tag{C.36}$$

Solutions to the radial part of the wave function can be found and are of the form

$$R_{nl}(r) = -2\left(\frac{Z}{na_0}\right)^{\frac{3}{2}} \left[\frac{(n-l-1)!}{n[(n+l)!]^3}\right]^{\frac{1}{2}} \rho^l \mathrm{e}^{-\rho/2} L_{n+l}^{2l+1}(\rho) \tag{C.37}$$

where n is the principal quantum number,

$$\rho = \frac{2Zr}{na_0}, \tag{C.38}$$

and the associated Laguerre polynomials are given by

$$L_\alpha^\beta(x) = \frac{\mathrm{d}^\beta}{\mathrm{d}x^\beta}\left(\mathrm{e}^x \frac{\mathrm{d}^\alpha}{\mathrm{d}x^\alpha}\left[\mathrm{e}^{-x}x^\alpha\right]\right) = \sum_{j=0}^{\alpha-\beta} \frac{\alpha!^2(-1)^{j+\beta}}{(j+\beta)!j!(\alpha-\beta-j)!}x^j. \tag{C.39}$$

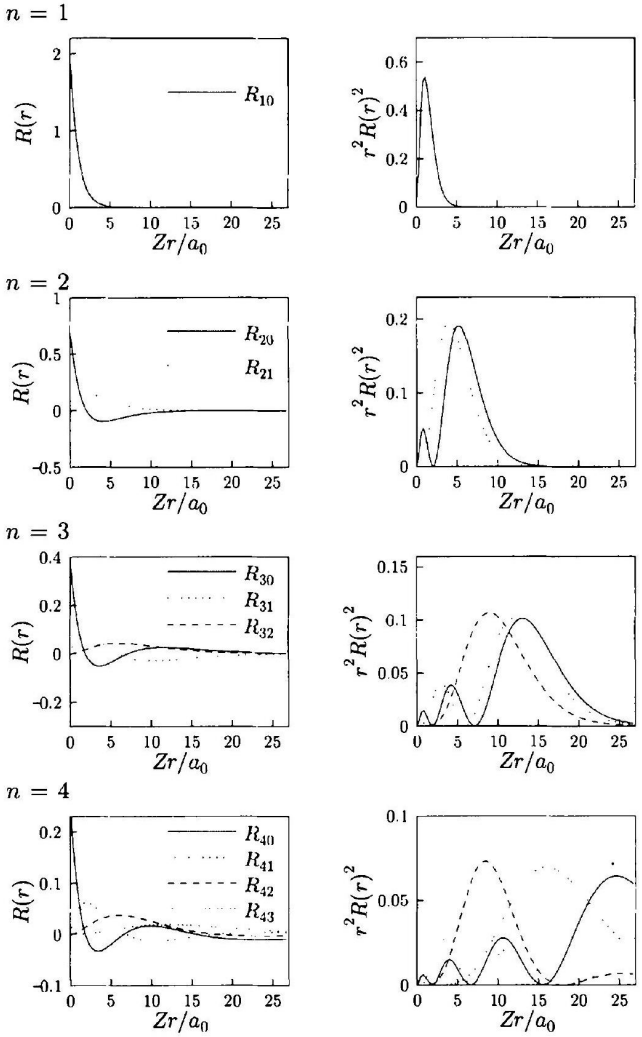
For atomic states the quantum numbers are in the range $n \geq 1, 0 \leq l \leq n-1$ and $-l \leq m_l \leq l$. Table C. 2 lists some radial wave functions for $n = 1, 2, 3$. Some radial wave functions $R(r)$ for the hydrogen atom are plotted in Fig. C.2. Also shown is the radial probability density $r^2R(r)^2$.

9.11.6 C. 6 The g-factor

The energy of the electron in a magnetic field B is

$$E = g\mu_\mathrm{B} m_s B \tag{C.40}$$

Fig. C. 2 The radial wave functions $R(r)$ (left) and the corresponding probability density $r^{2}R(r)^{2}$ (right) for $n = 1, 2, 3$ and 4.



where g is known as the g-factor. The energy levels therefore split by an amount $g\mu_{\mathrm{B}}B$. A natural consequence of Dirac's theory of the electron (outside the scope of this book) is that g is precisely equal to 2 . Actually the g -factor is not quite 2 but takes the value

$$g = 2\left(1 + \frac{\alpha}{2\pi} + \cdots\right) = 2.0023\ldots \tag{C.41}$$

where α is the fine structure constant, a dimensionless quantity given by

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} = \frac{1}{137.04} \tag{C.42}$$

This theoretical value obtained from quantum electrodynamics (QED) agrees with experiment to an astonishing degree of precision. The discrepancy from $g = 2$ can be explained as follows: electromagnetic interactions are due to the action of virtual photons. An electron can in fact emit a virtual photon, only to later reabsorb it. If you measure the magnetic moment of the electron during the life of one of these virtual photons, you will actually be measuring the magnetic moment of the electron-photon pair which will include an extra orbital component. You may also measure the system during a time in which more than one virtual photon has been created, though this sort of process becomes less probable as the number of virtual photons increases. On average then, the system has a magnetic moment slightly higher than what

you might first expect, and this accounts for the non-zero value of $g-2$. The expression is a power series in α with each successive term reflecting the contributions of progressively more convoluted creations and absorptions of virtual photons.

9.11.7 C. 7 d orbitals

The angular parts of the d wave functions ($l = 2$), in the form most often considered in condensed matter physics, can be constructed as linear combinations of functions of the form $Y_{2m}\mathrm{e}^{\mathrm{i}m_l\phi}$ where the functions Y_{2m} are those listed in the $l = 2$ row of Table C.1:

$$Y_{xy} = \frac{Y_{22} - Y_{2-2}}{\mathrm{i}\sqrt{2}} \tag{C.43}$$

$$Y_{x^2-y^2} = \frac{Y_{22} + Y_{2-2}}{\sqrt{2}} \tag{C.44}$$

$$Y_{yz} = \frac{-Y_{21} - Y_{2-1}}{\mathrm{i}\sqrt{2}} \tag{C.45}$$

$$Y_{zx} = \frac{-Y_{21} + Y_{2-1}}{\sqrt{2}} \tag{C.46}$$

$$Y_{z^2} = Y_{20} \tag{C.47}$$

which yields the following results (with the prefactors included for normalization)

$$Y_{xy} = \sqrt{\frac{15}{16\pi}}\sin^2\theta\sin 2\phi = \sqrt{\frac{15}{4\pi}}\frac{xy}{r^2} \tag{C.48}$$

$$Y_{x^2-y^2} = \sqrt{\frac{15}{16\pi}}\sin^2\theta\cos 2\phi = \sqrt{\frac{15}{16\pi}}\frac{x^2-y^2}{r^2} \tag{C.49}$$

$$Y_{yz} = \sqrt{\frac{15}{4\pi}}\sin\theta\cos\theta\sin\phi = \sqrt{\frac{15}{4\pi}}\frac{yz}{r^2} \tag{C.50}$$

$$Y_{zx} = \sqrt{\frac{15}{4\pi}}\sin\theta\cos\theta\cos\phi = \sqrt{\frac{15}{4\pi}}\frac{zx}{r^2} \tag{C.51}$$

$$Y_{z^2} = \sqrt{\frac{15}{16\pi}}\left(3\cos^2\theta - 1\right) = \sqrt{\frac{15}{16\pi}}\frac{2z^2-x^2-y^2}{r^2} \tag{C.52}$$

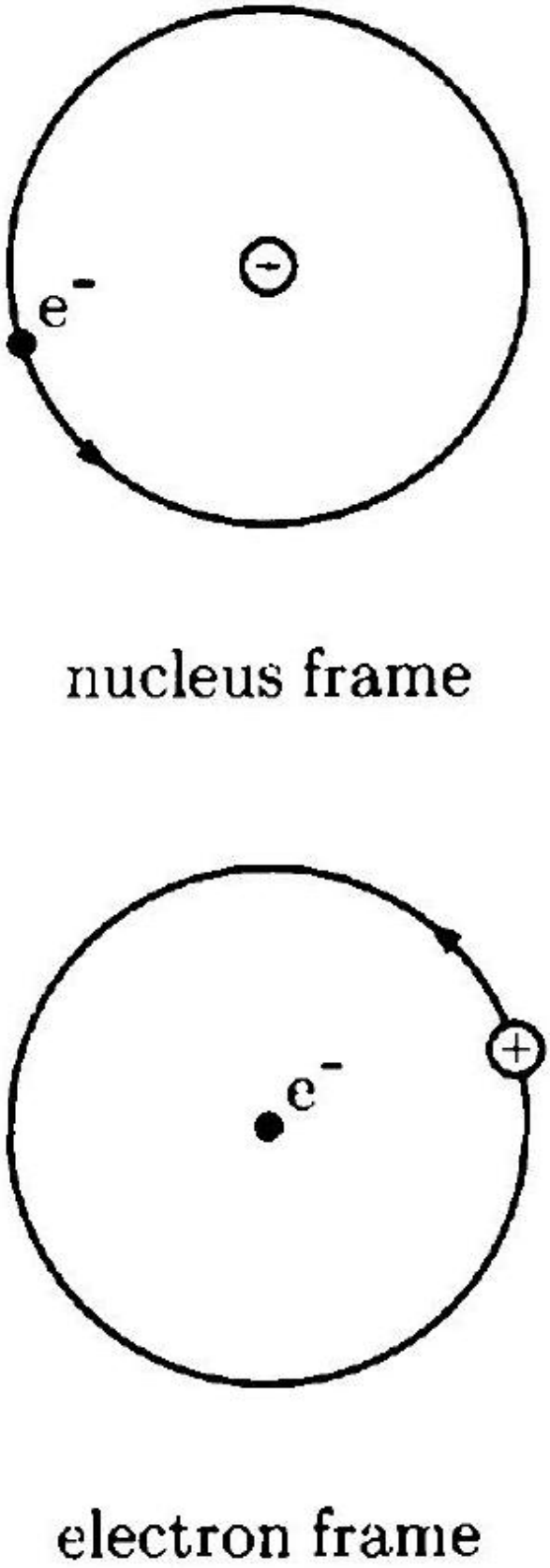


Fig. C. 3 The intrinsic spin-orbit interaction. In the frame of the nucleus the electron orbits the nucleus. but an inertial frame comoving with the electron it is the nucleus which orbits.

[2] There are an infinite set of Lorentz transformations corresponding to the various

instantaneous rest frames of the electron around its orbit. The Lorentz transformations do not commute when the direction of velocity changes and this is something our derivation has ignored. The net Lorentz transformation contains a rotation and it turns out that this Thomas precession yields a factor of precisely one-half.
L. H. Thomas (1903-1992)

These orbitals, together with the s and p orbitals, are shown in Fig. 3.1.

9.11.8 C. 8 The spin-orbit interaction

The spin-orbit interaction in an atom arises as follows. Consider an electron orbiting an atom. This is depicted at the top of Fig. C. 3 in the nucleus rest frame. In the lower diagram, the atom is shown in an inertial frame comoving with the electron in which the nucleus appears to be orbiting the electron. The orbiting nucleus constitutes a current which gives rise to a magnetic field at the origin equal to

$$\mathbf{B} = \frac{\boldsymbol{\varepsilon} \times \mathbf{v}}{c^2}. \tag{C.53}$$

where

$$\varepsilon = -\nabla V(r) = -\frac{\mathbf{r}}{r}\frac{\mathrm{d}V(r)}{\mathrm{d}r} \tag{C.54}$$

is the electric field at the electron due to the nucleus and $V(r)$ is the corresponding potential energy. Equation C. 53 comes from the transformation of electric and magnetic fields in special relativity. This magnetic field interacts with the spin of the electron to give a term in the Hamiltonian

$$H_{\mathrm{so}} = -\frac{1}{2}\mathbf{m}\cdot\mathbf{B} \tag{C.55}$$

$$= \frac{e\hbar^2}{2m_{\mathrm{e}}c^2 r}\frac{\mathrm{d}V(r)}{\mathrm{d}r}\mathbf{S}\cdot\mathbf{L} \tag{C.56}$$

where the orbital angular momentum is given by $\hbar\mathbf{L} = m_{\mathrm{e}}\mathbf{r}\times\mathbf{v}$ and the magnetic moment $\mathbf{m} = (geh/2m)\mathbf{S}$, and where the factor of $\frac{1}{2}$ in eqn C. 55 is the relativistic Thomas factor. This result can be obtained extremely elegantly with the relativistic correction automatically included using the Dirac equation. [2] This effect is known as the intrinsic spin-orbit interaction and is an interaction between the spin and the orbital part of an electron's wave function in an atom. For the Coulomb field in a hydrogen-like atom

$$\frac{1}{r}\frac{\mathrm{d}V(r)}{\mathrm{d}r} = \frac{Ze}{4\pi\epsilon_0 r^3} \tag{C.57}$$

and for electronic states with quantum numbers l and n one has

$$\left\langle r^{-3}\right\rangle = \frac{Z^3}{a_0^3 n^3 l\left(l+\frac{1}{2}\right)(l+1)} \tag{C.58}$$

so that the spin-orbit splitting is given by

$$\frac{Z^4 e^2\hbar^2\langle\mathbf{S}\cdot\mathbf{L}\rangle}{4\pi\epsilon_0 a_0^3 n^3 l\left(l+\frac{1}{2}\right)(l+1)}. \tag{C.59}$$

9.11.9 C. 9 Landé g-factor

The Landé g -factor g_J is given by

$$g_J = \frac{3}{2} + \frac{S(S+1) - L(L+1)}{2J(J+1)} \tag{C.60}$$

The expression for g_J in eqn C. 60 is obtained by starting with an expression for the magnetic moment operator:

$$\hat{\boldsymbol{\mu}} = \mu_{\mathbf{B}} \left(g_L \hat{\mathbf{L}} + g_S \hat{\mathbf{S}} \right). \tag{C.61}$$

In this expression $g_L = 1$ and $g_S = 2$ are the g-factors for orbital and spin angular momentum respectively. Now $\mathbf{S}$ and $\mathbf{L}$ may not be good quantum numbers in many atoms, but $\mathbf{J}$ is. Therefore the component of the magnetic moment which is parallel to $\mathbf{J}$ will be a conserved quantity but the component perpendicular to $\mathbf{J}$ will not be. We therefore write

$$\hat{\boldsymbol{\mu}} = g_J \mu_{\mathrm{B}} \hat{\mathbf{J}} \tag{C.62}$$

where g_{J} is a constant to be determined. Thus eqn C. 62 implies that g_J is the projection of $\hat{\mathbf{L}} + 2\hat{\mathbf{S}}$ on to $\hat{\mathbf{J}}$. Multiplying both sides of eqn C .61 by $\hat{\mathbf{J}}$ yields

$$\hat{\boldsymbol{\mu}} \cdot \mathbf{J} = \mu_{\mathbf{B}} \left(g_L \hat{\mathbf{L}} \cdot \hat{\mathbf{J}} + g_S \hat{\mathbf{S}} \cdot \hat{\mathbf{J}} \right) \tag{C.63}$$

Multiplying both sides of eqn C .62 by $\hat{\mathbf{J}}$ yields [3]

$$\hat{\boldsymbol{\mu}} \cdot \mathbf{J} = g_J \mu_{\mathrm{B}} \hat{\mathbf{J}}^2 = g_J \mu_{\mathrm{B}} J(J+1) \tag{C.64}$$

Also, $\mathbf{L}^2 = (\mathbf{J} - \mathbf{S})^2 = \mathbf{J}^2 + \mathbf{S}^2 - 2\mathbf{S} \cdot \mathbf{J}$ so that

$$\mathbf{S} \cdot \mathbf{J} = \frac{1}{2} \left(\mathbf{J}^2 - \mathbf{L}^2 + \mathbf{S}^2 \right) \tag{C.65}$$

and $\mathbf{S}^2 = (\mathbf{J} - \mathbf{L})^2 = \mathbf{J}^2 + \mathbf{L}^2 - 2\mathbf{L} \cdot \mathbf{J}$ so that

$$\mathbf{L} \cdot \mathbf{J} = \frac{1}{2} \left(\mathbf{J}^2 + \mathbf{L}^2 - \mathbf{S}^2 \right) \tag{C.66}$$

Equating eqns C. 63 and C.64, and inserting the results from eqns C. 65 and C. 66 yields

$$\begin{aligned} g_J = & g_L \left(\frac{J(J+1) + L(L+1) - S(S+1)}{2J(J+1)} \right) \\ & + g_S \left(\frac{J(J+1) - L(L+1) + S(S+1)}{2J(J+1)} \right) \end{aligned} \tag{C.67}$$

which reduces to eqn C. 60 if $g_L = 1$ and $g_s = 2$.

9.11.10 C. 10 Perturbation theory

Consider a Hamiltonian $\hat{\mathcal{H}}_0$ which has known eigenfunctions $|\phi_i\rangle$ with known eigenvalues E_i, so that

$$\hat{\mathcal{H}}_0 |\phi_i\rangle = E_i |\phi_i\rangle . \tag{C.68}$$

The eigenfunctions $|\phi_i\rangle$ are all orthogonal so that

$$\langle \phi_i \mid \phi_j \rangle = \int \mathrm{d}\tau \phi_i^* \phi_j = \delta_{ij} \tag{C.69}$$

[3] In this, and subsequent steps, we use the fact that the eigenvalue of S^2 is $S(S+1)$, of L^2 is $L(L+1)$ and of $\mathbf{J}^2$ is $J(J+1)$.

From now on, we shall assume the nondegenerate case, i.e. we will assume that none of the states are degenerate.

where $\mathrm{d}\tau$ is a volume element. Now suppose that a perturbation $\hat{V}$ is added to $\hat{\mathcal{H}}_0$ so that the new Hamiltonian is $\hat{\mathcal{H}}$ where

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_0 + \hat{V} \tag{C.70}$$

The system starts in a state $|\phi_k\rangle$ with energy E_k before the addition of the perturbation. The new eigenfunction of the system will be $|\psi\rangle$ and the eigenvalue E where

$$\hat{\mathcal{H}}|\psi\rangle = E|\psi\rangle, \tag{C.71}$$

but both $|\psi\rangle$ and E are unknown so far. $|\psi\rangle$ can be expanded in terms of the old eigenfunctions so that

$$|\psi\rangle = \sum_j a_j \, |\phi_j\rangle \tag{C.72}$$

Inserting this into eqn C.71, premultiplying by $\langle\phi_i|$ and integrating over the appropriate volume gives

$$\sum_j a_j \int \mathrm{d}\tau \phi_i^* \left(\hat{\mathcal{H}}_0 + \hat{V}\right) \phi_j = E \sum_j a_j \int \mathrm{d}\tau \phi_i^* \phi_j \tag{C.73}$$

so that using the orthogonality of the original eigenfunctions (eqn C.69) yields

$$\sum_j V_{ij} a_j = (E - E_i) \, a_i \tag{C.74}$$

which is known as the secular equation, where

$$V_{ij} = \langle \phi_i | \, \hat{V} \, | \phi_j \rangle = \int \mathrm{d}\tau \phi_i^* \hat{V} \phi_j \tag{C.75}$$

is the matrix element of the perturbation.

If the perturbation is small, $|\psi\rangle$ will be very close to the starting state $|\phi_k\rangle$ so that $a_k \approx 1$ and $|a_i| \ll 1$ for $i \neq k$. Since the wave function has changed only marginally, one might imagine that the new energy is given by

$$E \approx E_k \tag{C.76}$$

in other words that it remains completely unchanged. We can improve on this zeroth order approximation by using the results $a_k \approx 1$ and $a_i \ll 1$ for $i \neq k$. Substituting these into the secular equation can give some approximate expressions for the energy and for a_i. For $i = k$, the sum in the secular equation is dominated by the term for which $j = k$ giving

$$E \approx E_k + V_{kk} \tag{C.77}$$

the shift of the energy of the state is given simply by V_{kk}. This result, which contains the 'first-order perturbation theory' correction to the energy, is used very often in the text. The secular equation can also be evaluated for the case $i \neq k$, yielding

$$a_i = \frac{V_{ik}}{E_k - E_i} \tag{C.78}$$

(where E has been replaced by E_k). Putting eqn C. 78 back into eqn C. 74 yields the next level of approximation for the energy, namely

$$E \approx E_k + V_{kk} + \sum_{i \neq k} \frac{|V_{ik}|^2}{E_k - E_i} \tag{C.79}$$

which is the 'second-order perturbation theory' result. This process can be continued and will generate a power series expression for E, with each successive term in the series containing a higher power of the matrix elements of the perturbation.

9.11.11 Further reading

- A. I. Rae, Introduction to quantum mechanics, 3rd edition, JOP Publishing 1992.
- J. J. Sakurai, Modern quantum mechanics, 2nd edition, Addison-Wesley 1994.
- P. A. M. Dirac, The principles of quantum mechanics, 4th edition, OUP 1958.
- B. H. Bransden and C. J. Joachain, Physics of atoms and molecules, Longman 1983.
- P. W. Atkins, Molecular quantum mechanics, OUP 1983.
- G. K. Woodgate, Elementary atomic structure, OUP 1980.

9.12 Energy in magnetism and demagnetizing fields

9.12.1 D. 1 Energy

Calculating the energy of a magnetized medium in a magnetic field is a surprisingly subtle business. Various expressions can be obtained but they can refer to different things. In magnetism it depends on whether one is talking about the energy of the magnetic moment alone, or the magnetic moment plus whatever it is that is providing the magnetic field, and exactly how this energy is partitioned up.

For example, if you place a screwdriver close to the pole pieces of a large magnet, the screwdriver is strongly attracted into the region of maximum field. Replacing a bit of the air in the gap between the pole pieces with the magnetized screwdriver therefore seems to be an energetically favourable thing, suggesting that the presence of the magnetizable medium (the screwdriver) is lowering the energy. However, it takes more energy to establish a current in a magnet which contains an iron core than in one without an iron core, suggesting that the presence of the magnetizable medium (in this case the iron core) is increasing the energy. So does the energy decrease or increase? It depends on which energy you are considering. In the first example we are only measuring the energy saved by the force we feel on our arm holding the screwdriver, and we are ignoring the extra work needed to be done by the magnet power supply in maintaining the current at the level it had before we started waving the screwdriver around. This emphasizes that great care is needed in treating the energetics of magnetic materials (for a detailed treatment, see the article by Heine in the further reading).

In Appendix B, we found that $\delta W = -\boldsymbol{\mu} \cdot \delta \mathbf{B}$ is the appropriate free energy to consider for magnetized media. This expression will be useful in formulating magnetism in statistical mechanics (see Appendix E). In this appendix, we concentrate on the way in which the presence of a ferromagnetic material can change both $\mathbf{H}$ and $\mathbf{B}$ because of the introduction of demagnetizing fields which themselves cost energy.

9.12.2 D. 2 Demagnetizing factors

When the magnetization **M** inside a ferromagnetic body meets the surface, it has to suddenly stop. Hence there is a divergence of **M**. Using the equation

$$\nabla \cdot \mathbf{H} = -\nabla \cdot \mathbf{M}, \tag{D.1}$$

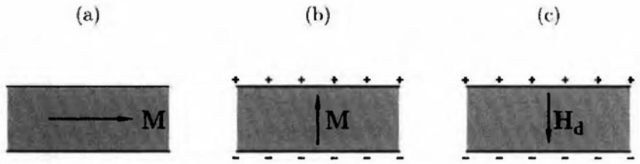


Fig. D. 1 Demagnetization in an infinite flat plate. viewed in cross-section. (a) If the magnetization lies in the plane of the plate, no magnetic poles are created on the surface of the plate fexcept for tiny ones at the ends). (b) If the magnetization is perpendicular to the plane of the plate, as shown, there is a negative (positive) divergence of M on the top (bottom) surface which produces a positive (negative) divergence of $\mathbf{H
$. This results in positive (negative) magnetic poles on the top (bottom) surface as shown. (c) This results in a demagnetization tield which runs from positive magnetic poles to negative magnetic poles.}

(which follows from $\nabla \cdot \mathbf{B} = 0$, see appendix B) we find that there is an equal and opposite divergence of **H**. The situation is as if magnetic monopoles have been left on the surface of the ferromagnet, and these monopoles act as sources of **H**. The resulting **H** field is known as a demognetizing field. This situation is illustrated for a simple case in Figure D. 1 which describes the magnetization in an infinite fiat plate or thin film of ferromagnetic material. If the magnetization lies in the plane of the plate, the only divergence of M is at the ends which. we suppose, are an infinite distance away. Therefore there is no demagnetizing field. If the magnetization is perpendicular to the plane of the plate, magnetic poles are created on top and bottom surfaces and give rise to a demagnetizing field $\mathbf{H}_\mathrm{d} = -\mathbf{M}$ inside the plate.

The demagnetizing field can be an extremely complicated function of position for a ferromagnet of arbitrary shape. However, it takes a relatively simple form in the case of an ellipsoidal ferromagnet. In this case it is uniform inside the ferromagnet with a value of $\mathbf{H_d}$ equal to

$$\mathbf{H_d} = -\mathbf{NM} \tag{D.2}$$

where N is the demagnetizing tensor. Thus in general we can write

$$(H_0)_i = -\sum_j N_{ij} M_j. \tag{D.3}$$

If M is along one of the principal axes of the ellipse, N_{ij} can be diagonalized so that

$$\mathrm{N} = \begin{pmatrix} N_x & 0 & 0 \\ 0 & N_y & 0 \\ 0 & 0 & N_z \end{pmatrix}. \tag{D.4}$$

The demagnetizing tensor satisfies

$$\mathrm{Tr\,N} = N_x + N_y + N_z = 1. \tag{D.5}$$

Example D. 1

We now examine some special cases.
(1) For a sphere, $N_x = N_y = N_z = \frac{1}{3}$, so that

$$\mathbf{H}_\mathrm{d} = -\frac{\mathbf{M}}{3} \tag{D.6}$$

(2) For a very long cylindrical rod parallel to $z, N_x = N_y = \frac{1}{2}, N_z = 0$. This is because if the magnetization lies exactly along the rod, the magnetic poles created will be at either end of the rod which can be assumed to be too far away to matter.
(3) For a flat plate perpendicular to $z, N_x = N_y = 0, N_z = 1$. This is the case considered above.

9.12.3 D. 3 A ferromagnet of arbitrary shape

In this section we will show how to compute the energy due to the demagnetizing energy for a ferromagnet of arbitrary shape in an applied magnetic field. For this problem, the magnetic fields $\mathbf{H}(\mathbf{r})$ and $\mathbf{B}(\mathbf{r})$ can be broken up into two components:

$$\mathbf{H}(\mathbf{r}) = \mathbf{H}_\mathrm{a}(\mathbf{r}) + \mathbf{H}_\mathrm{d}(\mathbf{r}) \tag{D.7}$$
$$\mathbf{B}(\mathbf{r}) = \mathbf{B}_\mathrm{a}(\mathbf{r}) + \mathbf{B}_\mathrm{d}(\mathbf{r}), \tag{D.8}$$

where the applied field is $\mathbf{H_a}(\mathbf{r}) = \mathbf{B_a}(\mathbf{r})/\mu_0$ and the demagnetizing field is $\mathbf{H}_\mathrm{d}(\mathbf{r})$ and we write $\mathbf{B}_\mathrm{d}(\mathbf{r})$ as a sum of the flux density due to the demagnetizing field $\mu_0\mathbf{H}_\mathrm{d}(\mathbf{r})$ and $\mu_0\mathbf{M}(\mathbf{r})$ inside the material, i.e. $\mathbf{B}_\mathrm{d}(\mathbf{r}) = \mu_0\left(\mathbf{H}_\mathrm{d}(\mathbf{r}) + \mathbf{M}(\mathbf{r})\right)$. The demagnetizing field arises from the magnetic poles which are produced on the surface of the magnetic body wherever $\nabla\cdot\mathbf{M}(\mathbf{r}) \neq 0$.

We now state an important result. The demagnetizing fields $\mathbf{B}_\mathrm{d}(\mathbf{r})$ and $\mathbf{H}_\mathrm{d}(\mathbf{r})$ satisfy

$$\int_{\text{all space}} \mathbf{B}_\mathrm{d}\cdot\mathbf{H}_\mathrm{d}\ \mathrm{d}\tau = 0. \tag{D.9}$$

The proof is as follows: since $\nabla\times\mathbf{H}_d = 0$, the demagnetizing field $\mathbf{H}_d$ can be written in terms of the gradient of a scalar function ϕ (e.g. $\mathbf{H}_d = -\nabla\phi$). Now

$$\nabla\cdot(\phi\mathbf{B}_\mathrm{d}) \equiv \phi\nabla\cdot\mathbf{B}_\mathrm{d} + (\nabla\phi)\cdot\mathbf{B}_\mathrm{d} = (\nabla\phi)\cdot\mathbf{B}_\mathrm{d} \tag{D.10}$$

where the last equality follows from $\nabla\cdot\mathbf{B_d} = 0$. Then

$$\int_{\text{all space}} \mathbf{B}_\mathrm{d}\cdot\mathbf{H}_\mathrm{d}\ \mathrm{d}\tau = -\int_{\text{als space}} \nabla\cdot(\phi\mathbf{B}_\mathrm{d})\,\mathrm{d}\tau = \lim_{R\to\infty}\int_{\text{surface}} \phi\mathbf{B}_\mathrm{d}\cdot\mathbf{dS} = 0, \tag{D.11}$$

where the surface integral is taken over a sphere whose radius R tends to infinity. The integral is zero if the ferromagnetic body is of finite extent, so that as $R\to\infty, \mathbf{B}_\mathrm{d}\sim R^{-2}$ and $\phi\sim R^{-1}$ at worst.

If the applied field is not present, then the energy E of the demagnetizing field is simply the integral over all space of the energy density $\frac{1}{2}\mu_0\mathbf{H_d^2}$. This
can be converted to an integral inside the volume V of the ferromagnetic body by using eqn D.9, so that the energy can be expressed as

$$E = \frac{\mu_0}{2} \int_{\text{as yoce}} \mathbf{H}_\text{d}^2 \, \text{d}\tau = -\frac{1}{2} \int_{\text{ar. war}} (\mathbf{B}_\text{d} - \mu_0 \mathbf{M}) \cdot \mathbf{H}_\text{d} \, \text{d}\tau$$
$$= -\frac{\mu_0}{2} \int_{a^2 \text{ yax}} \mathbf{M} \cdot \mathbf{H}_\text{d} \, \text{d}\tau$$
$$= -\frac{\mu_0}{2} \int_V \mathbf{M} \cdot \mathbf{H}_\text{d} \, \text{d}\tau \tag{D.12}$$

where the last line follows from the fact that $M = 0$ outside the body. It should be noted that even though this final integral is only over the body. it expresses the energy of all stray fields, including those outside the body.

If there is an applied field, then the energy of the ferromagnetic body may be written as the difference between the energy density in the total field $\frac{1}{2}\mu_0 \mathbf{H}^2$ and the energy density in the applied field $\frac{1}{2}\mu_0 \mathbf{H}_\text{a}^2$. This removes the infinite energy contribution from a uniform applied field integrated over all space. Thus the energy can be usefully written as

$$E = \frac{\mu_0}{2} \int_{\text{all quec}} \left(\mathbf{H}^2 - \mathbf{H}_\text{a}^2 \right) \text{d}\tau \tag{D.13}$$

Using the fact that $\nabla \times \mathbf{H}_a = 0$, we have

$$\int_{\text{dulP}_\text{Pax}} \mathbf{H}_\text{a} \cdot \mathbf{B}_\text{d} \, \text{d}\tau = 0 \tag{D.14}$$

Hence, by also using the following equations:

$$\mathbf{H}^2 = \mathbf{H}_\text{a}^2 + 2\mathbf{H}_\text{d} \cdot \mathbf{H}_\text{a} + \mathbf{H}_\text{d}^2. \tag{D.15}$$
$$\mathbf{H}_\text{d} \cdot \mathbf{H}_\text{a} = \mathbf{H}_\text{a} \cdot \left(\frac{\mathbf{B}_\text{d}}{\mu_0} - \mathbf{M} \right). \tag{D.16}$$

and by also using eqn D.9. the energy E can be written in the familiar form

$$\boldsymbol{E} = -\mu_0 \int_V \mathbf{M} \cdot \mathbf{H}_\text{a} \, \text{d}\tau - \frac{\mu_0}{2} \int_V \mathbf{M} \cdot \mathbf{H}_\text{d} \, \text{d}\tau \tag{D.17}$$

the sum of a Zeeman term and a demagnetizing term, with each integral being taken only over the volume V of the ferromagnetic body. The energy in eqn D. 17 nevertheless implicitly includes the energy associated with the demagnetizing fields outside the body.

Example D. 2

A ferromagnetic sphere of radius a with uniform magnetization $\mathbf{M}$ is used here as a detailed example of the ideas leading to eqn D.12. It is sketched in Fig. D.2. The demagnetizing fields are given by

$$\mathbf{H}_\text{d} = -\frac{\mathbf{M}}{3} \tag{D.J8}$$
$$\mathbf{B}_\text{d} = \frac{2\mu_0 \mathbf{M}}{3}. \tag{D.19}$$

The factor of one-half in the demagnetizing term can be understood in terms of an avoidance of double-counting. The 7eeman term is the interaction of the magnetization with the applied

field. The demagnetizing term is a scif-energy so the moments which cause the field interact with their own field. The factor of one-half ensures that you don't count this energy twice.

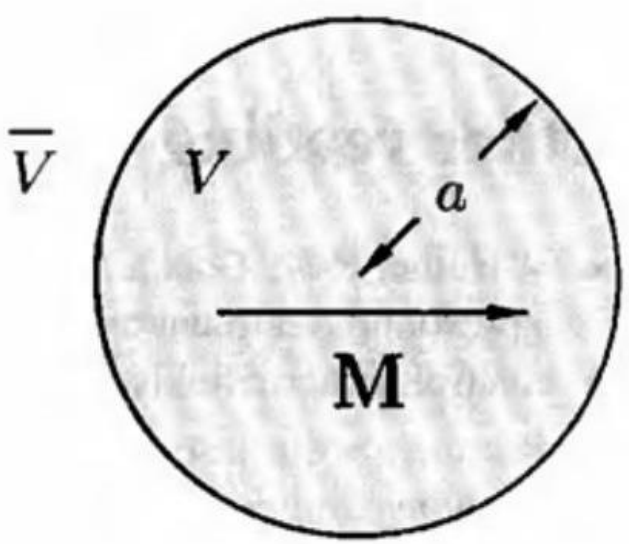


Fig. D. 2 A ferromagnetic sphere with uniform magnetization M. The region inside the sphere of volume $V_{\text {sphere }}=\frac{4}{3} \pi a^{3}$ is denoted V. the region outside the sphere is denoted $\bar{V}$.}

so that inside the sphere (a space of volume $V_{\text{sphere}} = \frac{4}{3}\pi a^3$, denoted by V) simple integration yields the following results:

$$-\frac{\mu_0}{2}\int_V \mathbf{M}\cdot\mathbf{H}_\mathrm{d}\ \mathrm{d}\tau = \frac{1}{6}\mu_0 M^2 V_{\text{sphere}} \tag{D.20}$$

$$\frac{\mu_0}{2}\int_V \mathbf{H}_\mathrm{d}^2\ \mathrm{d}\tau = \frac{1}{18}\mu_0 M^2 V_{\text{sphere}} \tag{D.21}$$

$$\frac{1}{2}\int_V \mathbf{B}_\mathrm{d}\cdot\mathbf{H}_\mathrm{d}\ \mathrm{d}\tau = -\frac{1}{9}\mu_0 M^2 V_{\text{sphere}} \tag{D.22}$$

Outside the sphere (a space denoted by $\bar{V}$), $\mathbf{B}_\mathrm{d} = \mu_0\mathbf{H}_\mathrm{d}$ (since $\mathbf{M} = 0$ outside the sphere), and the demagnetizing fields result from the magnetic dipole moment of the sphere. These have radial and polar components given by

$$\begin{aligned}(H_\mathrm{d})_r &= 2Ma^3\cos\theta/3r^3\\ (H_\mathrm{d})_\theta &= Ma^3\sin\theta/3r^3\end{aligned} \tag{D.23}$$

so that integration over this region outside the sphere gives:

$$\frac{\mu_0}{2}\int_{\bar{V}} \mathbf{H}_\mathrm{d}^2\ \mathrm{d}\tau = \frac{1}{2}\int_{\bar{V}} \mathbf{B}_\mathrm{d}\cdot\mathbf{H}_\mathrm{d}\ \mathrm{d}\tau = \frac{1}{2\mu_0}\int_{\bar{V}} \mathbf{B}_\mathrm{d}^2\ \mathrm{d}\tau = \frac{1}{9}\mu_0 M^2 V_{\text{sphere}} \tag{D.24}$$

Then integration over all space gives

$$\begin{aligned}-\frac{\mu_0}{2}\int_{\text{all space}} \mathbf{M}\cdot\mathbf{H}_\mathrm{d}\ \mathrm{d}\tau &= \frac{\mu_0}{2}\int_{\text{all space}} \mathbf{H}_\mathrm{d}^2\ \mathrm{d}\tau\\ &= \frac{1}{6}\mu_0 M^2 V_{\text{sphere}}\end{aligned} \tag{D.25}$$

in agreement with eqn D.12. By combining eqns D. 22 and D. 24 one has

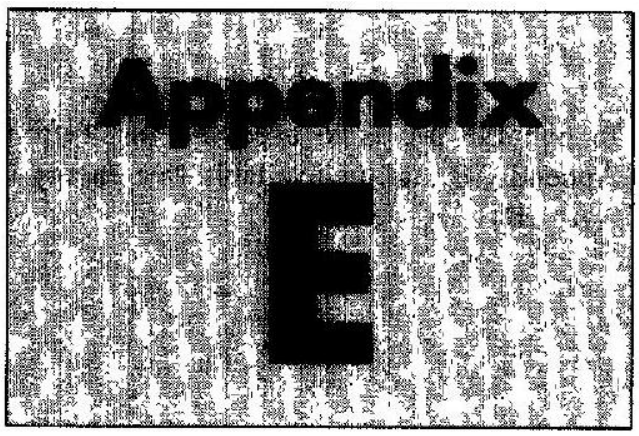
$$\frac{1}{2}\int_{\text{als space}} \mathbf{B}_\text{d} \cdot \mathbf{H}_\text{d}\ \text{d}\tau = 0 \tag{D.26}$$

in agreement with eqn D.9.

9.12.4 Further reading

- V. Heine, Proc. Camb. Phil. Soc. 52, 546 (1956) contains a profound treatment of the thermodynamics of bodies in electromagnetic fields.
- Treatments are also found in C. J. Adkins, Equilibrium thermodynamics, CUP 1983, and J. R. Waldram, The theory of thermodynamics, CUP 1985.
- Magnetostatic energy and demagnetizing factors are covered in depth in A. Aharoni, Introduction to the theory of ferromagnetism, OUP 1996.
- I. Brevik, J. Phys.: Condensed Matter 7, 8065 (1995) contains a helpful analysis of the problem of calculating the energy of a ferromagnetic body immersed in a ferromagnetic liquid.

9.13 Statistical mechanics



This appendix briefly sketches some results in statistical mechanics which will be useful in the main text.

9.13.1 E. 1 The partition function and thermodynamic functions

The probability p_i of occupying a state i with energy E_i at temperature T is

$$p_i = \frac{\text{e}^{-E_i/k_\text{B}T}}{Z} \tag{E.l}$$

where Z is the single-particle partition function which is used to normalize the probability such that

$$\sum_i p_i = 1. \tag{E.2}$$

This implies that Z is given by

$$Z = \sum_i \mathrm{e}^{-E_i/k_\mathrm{B}T}. \tag{E.3}$$

Defining $\beta = 1/k_\mathrm{B}T, \partial Z/\partial\beta$ can be evaluated, giving

$$\frac{\partial Z}{\partial \beta} = -\sum_i E_i \mathrm{e}^{-\beta E_i}. \tag{E.4}$$

Now since we have found the probability distribution, we can use it to calculate expected values of quantities. The expected value of the energy $\langle E \rangle$ is given by

$$\langle E \rangle = \sum_i E_i p_i = \frac{\sum_i E_i \mathrm{e}^{-\beta E_i}}{Z} = -\frac{1}{Z}\frac{\partial Z}{\partial \beta} = -\frac{\partial \ln Z}{\partial \beta}. \tag{E.5}$$

The entropy per particle is given by

$$S = -k_\mathrm{B} \sum_i p_i \ln p_i \tag{E.6}$$

and hence

$$S = -k_\mathrm{B} \sum_i \frac{\mathrm{e}^{-\beta E_i}}{Z} \left[-\beta E_i - \ln Z\right] = k_\mathrm{B} \ln Z + \frac{\langle E \rangle}{T}. \tag{E.7}$$

E. 1 The partition function and thermodynamic functions
E. 2 The equipartition theorem

In eqn E. 8 we have used the result from Appendix B that the available free energy is $-\mu \cdot \delta\mathbf{B}$.

For N independent distinguishable particles, the appropriate partition function becomes Z^N and so the total energy is $E = N\langle E \rangle$. The first law of thermodynamics can be expressed as

$$\mathrm{d}E = T\ \mathrm{d}S - p\ \mathrm{d}V - M\ \mathrm{d}B \tag{E.8}$$

The Helmholtz free energy F, also known as the Helmholtz function, is defined by

$$F = E - TS \tag{E.9}$$

and is related to the partition function Z by

$$F = -Nk_B T \ln Z \tag{E.10}$$

Using eqns E. 8 and E. 9, $\mathrm{d}F$ can be written

$$\mathrm{d}F = -S\ \mathrm{d}T - p\ \mathrm{d}V - M\ \mathrm{d}B \tag{E.11}$$

and so

$$M = -\left(\frac{\partial F}{\partial B}\right)_{T,V} = Nk_B T \left(\frac{\partial \ln Z}{\partial B}\right)_{T.V} \tag{E.12}$$

The entropy is given by

$$M = -\left(\frac{\partial F}{\partial T}\right)_{V.B} \tag{E.13}$$

9.13.2 E. 2 The equipartition theorem

Very often the energy of a system can be written as a sum of quadratic terms. For example, a mass m on a spring of spring constant K has an energy $E = \frac{1}{2}mv^2 + \frac{1}{2}Kx^2$, which contains two such quadratic terms. At low temperatures, where quantum effects dominate, we expect harmonic oscillator solutions and a ladder of energy levels. But at high temperatures, where $k_{\mathrm{B}}T$ is much larger than the energy scale set by the spacing between the levels, the system can be considered classically.

The equipartition theorem states that in the classical limit, when $k_{\mathrm{B}}T$ is much larger than the energy spacing between quantum levels, and the energy contains a number of independent quadratic terms, the expected value of the energy is equal to $\frac{1}{2}k_{\mathrm{B}}T$ multiplied by the number of independent quadratic terms in the energy.

The proof is straightforward: if $E(x_1, x_2, \ldots x_n) = \sum_{i=1}^{n} \alpha_i x_i^2$ where α_i are constants, then

$$\langle E \rangle = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \mathrm{d}x_1\ \mathrm{d}x_2 \ldots\ \mathrm{d}x_n \mathrm{e}^{-\beta E(x_1,x_2,\ldots x_n)} E}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \mathrm{d}x_1\ \mathrm{d}x_2 \ldots\ \mathrm{d}x_n \mathrm{e}^{-\beta E(x_1,x_2,\ldots x_n)}} \tag{E.14}$$

This looks formidable but the integrals for each variable can be easily separated. Hence

$$\begin{aligned}
\langle E \rangle &= \sum_{i=1}^{n} \alpha_i \left(\frac{\int_{-\infty}^{\infty} \mathrm{d}x_i x_i^2 \mathrm{e}^{-\beta\alpha_i x_i^2}}{\int_{-\infty}^{\infty} \mathrm{d}x_i \mathrm{e}^{-\beta\alpha_i x_i^2} \prod_{j\neq i} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}x_j \mathrm{e}^{-\beta\alpha_i x_j^2}} \right) \\
&= \sum_{i=1}^{n} \left(\frac{\frac{\alpha_i}{2}\sqrt{\frac{\pi}{\alpha_i^3\beta^3}}}{\sqrt{\frac{\pi}{\alpha_i\beta}}} \right) \\
&= \sum_{i=1}^{n} \frac{1}{2\beta} \\
&= \frac{n}{2} k_{\mathrm{B}}T
\end{aligned} \tag{E.15}$$

where use has been made of the standard integrals

$$\int_{-\infty}^{\infty} \mathrm{d}x \mathrm{e}^{-\gamma x^2} = \sqrt{\frac{\pi}{\gamma}} \quad \text{and} \quad \int_{-\infty}^{\infty} \mathrm{d}x x^2 \mathrm{e}^{-\gamma x^2} = \frac{1}{2}\sqrt{\frac{\pi}{\gamma^3}} \tag{E.16}$$

Example E. 1

Three examples of this result now follow:
(1) For a mass m on a spring of spring constant K, $E = \frac{1}{2}mv^2 + \frac{1}{2}Kx^2$, (two quadratic terms), so that $\langle E \rangle = k_{\mathrm{B}}T$.
(2) For a molecule of mass m in a gas, $E = \frac{1}{2}mv_x^2 + \frac{1}{2}mv_y^2 + \frac{1}{2}mv_z^2$, (three quadratic terms), so that $\langle E \rangle = \frac{3}{2}k_{\mathrm{B}}T$.
(3) For a solid containing N atoms (and therefore $3N$ springs, and hence the energy contains $6N$ quadratic terms), $\langle E \rangle = 3Nk_{\mathrm{B}}T$.

In addition to adding $\frac{1}{2}k_{\mathrm{B}}T$ to the energy, each quadratic degree of freedom adds $\frac{1}{2}k_{\mathrm{B}}$ to the heat capacity.

9.13.3 Further reading

- F. Reif, Fundamentals of statistical and thermal physics, Mc-Graw-Hill, 1965.
- J. R. Waldram, The theory of thermodynamics, CUP 1985.
- P. M. Chaikin and T. C. Lubensky, Principles of condensed matter physics, CUP 1995.
- A. M. Glazer and J. S. Wark, Statistical mechanics, OUP 2001.

9.14 Answers and hints to selected problems

(1.1) $9.27 \times 10^{-24}\mathrm{Am}^2; 8.40\mathrm{GHz}; 34.7\mu\mathrm{eV}; 8.40\mathrm{GHz}$. Here we are only considering spin angular momentum.
(1.4) This is the commutator of a scalar operator ($\hat{\mathbf{S}} \cdot \mathbf{X}$) and a vector operator ($\hat{\mathbf{S}}$). Work out one component of this, and so evaluate

$$\begin{aligned}
[\hat{\mathbf{S}} \cdot \mathbf{X}, \hat{\mathbf{S}} \cdot \mathbf{Y}] &= \frac{1}{4}((\sigma \cdot \mathbf{X})(\sigma \cdot \mathbf{Y}) - (\sigma \cdot \mathbf{y})(\sigma \cdot \mathbf{X})) \\
&= \frac{\mathrm{i}}{4}(\sigma \cdot (\mathbf{X} \times \mathbf{Y}) - \sigma \cdot (\mathbf{Y} \times \mathbf{X})) \\
&= \frac{\mathrm{i}}{2}\sigma \cdot (\mathbf{X} \times \mathbf{Y}) \\
&= \mathrm{i}\mathbf{S} \times \mathbf{X} \cdot \mathbf{Y} \qquad \text{(F.1)}
\end{aligned}$$

and the result follows.
(1.5) Using eqn 1.59 one can show that

$$\begin{aligned}
\hat{S}_z \left(\hat{S}_\pm \left|S, S_z\right\rangle\right) &= \left(\hat{S}_\pm \hat{S}_z \pm \hat{S}_\pm\right) \left|S, S_z\right\rangle \\
&= (S_z \pm 1) \left(\hat{S}_\pm \left|S, S_z\right\rangle\right) \qquad \text{(F.2)}
\end{aligned}$$

so that the z -component of the spin is raised or lowered. Equations 1.58 and 1.61 can be written as

$$\hat{S}_+ \hat{S}_- - \hat{S}_- \hat{S}_+ = 2\hat{S}_z \qquad \text{(F.3)}$$

$$\begin{aligned}
\hat{S}_+ \hat{S}_- + \hat{S}_- \hat{S}_+ &= 2\left(\hat{S}_x^2 + \hat{S}_y^2\right) \\
&= \hat{\mathbf{S}}^2 - \hat{S}_z^2 \qquad \text{(F.4)}
\end{aligned}$$

so that adding or subtracting leads to

$$\hat{S}_+ \hat{S}_- = \hat{\mathbf{S}}^2 - \hat{S}_z^2 + \hat{S}_z \qquad \text{(F.5)}$$

$$\hat{S}_- \hat{S}_+ = \hat{\mathbf{S}}^2 - \hat{S}_z^2 - \hat{S}_z \qquad \text{(F.6)}$$

Hence

$$\begin{aligned}
\left\langle S, S_z\right| \hat{S}_- \hat{S}_+ \left|S, S_z\right\rangle &= S(S+1) - S_z^2 - S_z \\
&= S(S+1) - S_z\left(S_z + 1\right) \qquad \text{(F.7)}
\end{aligned}$$

and

$$\begin{aligned}\langle S, S_z|\, \hat{S}_+\hat{S}_-\, |S, S_z\rangle &= S(S+1) - S_z^2 + S_z \\ &= S(S+1) - S_z\,(S_z - 1)\end{aligned} \tag{F.8}$$

and the required normalization is therefore proved.
(1.6) Without loss of generality, put $\mathbf{B} = (0, 0, B)$ so that $\mathbf{B} \times \mathbf{r} = B(-y, -x, 0)$ and $\frac{1}{2}\nabla \times (\mathbf{B} \times \mathbf{r}) = (0, 0, B)$. Other choices are possible, e.g. $\mathbf{A} = -By(1, 0, 0)$ or $\mathbf{A} = Bx(0, 1, 0)$. In fact, $\mathbf{A} = B((\alpha - 1)y, \alpha x, 0)$ will work where α is a real number.
(1.7) Hint: evaluate $[\sigma \cdot (\mathbf{p} + e\mathbf{A})]^2/2m_e$.
(1.8) Without loss of generality, the new axis can be at an angle θ to the z axis in the xz plane. Thus the operator is

$$\hat{S}_\theta = \cos\theta \hat{S}_z + \sin\theta \hat{S}_x \tag{F.9}$$

The eigenvector of this operator with eigenvalue 1 is found to be

$$(2(1 - \cos\theta))^{-1/2} \begin{pmatrix} \sin\theta \\ 1 - \cos\theta \end{pmatrix} \tag{F.10}$$

so that the required probability is

$$\frac{\sin^2\theta}{2(1 - \cos\theta)} = \cos^2\frac{\theta}{2} \tag{F.11}$$

(1.10) $H_\mathrm{d}/H_\mathrm{a}$: (a) 3×10^{-5} (b) $-8.8 \times 10^{-4} B_\mathrm{d}/B_\mathrm{a}$: (a) -6×10^{-5} (b) 1.76×10^{-3}
(1.12) Hint: $\sigma_m^2 = I$ so that

$$\sigma^n = \begin{cases} I & n \text{ even} \\ \sigma_m & n \text{ odd} \end{cases} \tag{F.12}$$

so that

$$\begin{aligned}\mathrm{e}^{\mathrm{i}\alpha\sigma_m} =& I\left(1 - \frac{\alpha^2}{2!} + \frac{\alpha^4}{4!} - \cdots\right) \\ &+ \mathrm{i}\sigma_m\left(\alpha - \frac{\alpha^3}{3!} + \frac{\alpha^5}{5!} - \cdots\right) \\ =& I\cos\alpha + \mathrm{i}\sigma_m \sin\alpha\end{aligned} \tag{F.13}$$

(1.14) (a) Torque $= I\ddot{\theta} = -p\varepsilon\sin\theta$ so that $\ddot{\theta} + p\varepsilon\sin\theta/I = 0$ which leads to simple harmonic motion when $\theta \ll 1$ with $\omega^2 = p\mathcal{E}/I$. (b) The Hamiltonian is $\mathcal{H} = \hat{L}^2/2I - p\mathcal{E}\cos\theta$. In this calculation, several commutators need to be evaluated. You should find that

$$[\hat{\theta}, \hat{L}] = \mathrm{i}\hbar \tag{F.14}$$

$$\left[\hat{\theta}, \hat{L}^2\right] = 2\mathrm{i}\hbar\hat{L} \tag{F.15}$$

$$[\hat{L}, \cos\theta] = \mathrm{i}\hbar\sin\theta \tag{F.16}$$

and the results follow. The electric dipole is different because, unlike the magnetic moment, it is not associated with angular momentum. (An electric dipole is just two, opposite charges, which are spatially separated. A magnetic dipole is two, opposite charges, which are orbiting

each other and so is associated with angular momentum.) Note that the Hamiltonian contains no dissipation, so the electric dipole oscillates forever, and the magnetic dipole precesses forever. Both dipoles will only line up with the external field if we allow them to change their energy, by exchanging energy with a heat bath. This is a dissipative process.
(2.1) $-1.65 \times 10^{-15}; 7.8 \times 10^{-12}$
(2.2) A duck is mainly water and has a mass of $2-3$ kg. An iron filing might be 0.1 $\mathrm{mm}^3 = 10^{-10}$ m^3 (as a guess) and each iron atom carries $2.2\mu_{\mathrm{B}}$. The density of iron is 7873 kg m^{-3} and the relative atomic mass is 55.847 g . This can be used to show that the required field is 1 mT (the Earth's field is about 0.05 mT). A cow might be 400 kg , and assuming that it too is made of water means that with about 200 times more mass, and hence more volume, a much smaller field ($\sim 2\mu$ T) is needed to produce the same level of magnetization. Of course we have ignored the effect of haemoglobin.
$2.74 \times 10^{-5}\mathrm{JT}^{-1}$
(2.4) Hint: use $Z = 2\cosh(\mu_{\mathrm{B}}B/k_{\mathrm{B}}T)$ and $F = -nk_{\mathrm{B}}T\log Z$. Then $E = -n(\,\mathrm{d}\log Z/\mathrm{d}\beta)$ and $C = \partial E/\partial T$. The entropy can be obtained either using $S = -(\partial F/\partial T)_B$ or $S = (E-F)/T$.
(2.5) Hint: a useful intermediate result to prove is that

$$\frac{\partial \log Z_y}{\partial y} = \mathrm{B}_J(y). \tag{F.17}$$

(2.7) (a) ${}^5\mathrm{I}_8$ (b) ${}^4\mathrm{I}_{15/2}$ (c) ${}^3\mathrm{H}_6$ (d) ${}^1\,\mathrm{S}_0$
(2.8) Hints: (1) You will need to evaluate a term containing $\mathbf{p}\cdot\mathbf{A} + \mathbf{A}\cdot\mathbf{p}$. This is proportional to

$$\mathbf{p}\cdot\boldsymbol{\mu}\times\mathbf{r} + \boldsymbol{\mu}\times\mathbf{r}\cdot\mathbf{p} = 2\mathbf{r}\times\mathbf{p}\cdot\boldsymbol{\mu} \tag{F.18}$$

where the final equality works because μ does not depend on position, so commutes with $\mathbf{r}$ and $\mathbf{p}$, and you can commute $\mathbf{r}$ and $\mathbf{p}$ because they appear in a scalar triple product so you are never multiplying the same component of $\mathbf{r}$ and $\mathbf{p}$.
(2) $\nabla^2\left(\frac{1}{r}\right) = \nabla\cdot\nabla\left(\frac{1}{r}\right) = 4\pi\delta(\mathbf{r})$. An analogous result is $(\mathrm{S}\,.\,\nabla\,)\,\nabla\left(\frac{1}{r}\right) = \mathbf{S}\frac{4\pi}{3}\delta(\mathbf{r})$, where the factor of 3 is due to the fact that $\mathbf{S}$ picks out one of out of the three components of the singularity.
$1.89 \times 10^{-4}\mathrm{emumol} - 1; 9.68 \times 10^{-7}\mathrm{emug}^{-1}$
(2.10) At fixed T,

$$T\delta S = -\mathbf{B}\cdot\delta\mathbf{M} - \mathbf{M}\cdot\delta\mathbf{B} + \frac{\partial\left(nk_{\mathrm{B}}T\log Z\right)}{\partial\mathbf{B}}\cdot\delta\mathbf{B}, \tag{F.19}$$

but using $M = nk_{\mathrm{B}}T(\partial\log Z/\partial B)$, the result follows.
(3.1) Sc^{2+} has one 3d electron $\left(l=2, s=\frac{1}{2}\right)$ and has 5 orbital states characterized by $l_z = -2,-1,0,1,2$. The energy levels are thus given by $0, A, 4A$ (degeneracies 2,4 and 4 respectively, including spin) with the ground state being twofold degenerate ($E=0$) if $A>0$, or fourfold degenerate ($E=4A$) if $A<0$.
Let us write states using the notation $|l_z, s_z\rangle$. If $A>0$, the ground state levels are $\left|0,\frac{1}{2}\right\rangle$ and $\left|0,-\frac{1}{2}\right\rangle$ and are not split by the spin-orbit interaction (because $\lambda l_z s_z = 0$). The levels are split by a magnetic field in any direction, and lead to a Curielike susceptibility providing $\mu_{\mathrm{B}}B \ll k_{\mathrm{B}}T \ll A$. This condition ensures that you do not populate excited states ($k_{\mathrm{B}}T \ll A$) and you keep on the linear part of the Brillouin function ($\mu_{\mathrm{B}}B \ll k_{\mathrm{B}}T$).
If $A<0$, the ground state levels are $\left|2,\frac{1}{2}\right\rangle, \left|2,-\frac{1}{2}\right\rangle \left|-2,\frac{1}{2}\right\rangle$ and $\left|-2,-\frac{1}{2}\right\rangle$ and the spin-orbit interaction acts so that $\left|2,-\frac{1}{2}\right\rangle$ and $\left|-2,\frac{1}{2}\right\rangle$ are lowered in energy by λ while $\left|2,\frac{1}{2}\right\rangle$ and $\left|-2,-\frac{1}{2}\right\rangle$ are raised in energy by λ. The lower levels will only now be split by a magnetic field parallel to z and so the susceptibility will be Curie-like if $\mu_{\mathrm{B}}B \ll k_{\mathrm{B}}T \ll \lambda$. The susceptibility will be

temperature independent if the field is applied perpendicular to z.
(3.2) Use

$$\begin{aligned}(1+x)^{-1/2} =& 1-\frac{x}{2}+\frac{\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)}{2!}x^2+\frac{\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\left(-\frac{5}{2}\right)}{3!}x^3 \\ &+\frac{\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\left(-\frac{5}{2}\right)\left(-\frac{7}{2}\right)}{4!}x^4+\cdots \\ =& 1-\frac{x}{2}+\frac{3x^2}{8}-\frac{5x^3}{16}+\frac{35x^4}{128}+\cdots \qquad \text{(F.20)}\end{aligned} \tag{F.20}$$

(3.3) Converting V into spherical polars produces

$$V=\frac{Dr^4}{8}\left[\sin^4\theta\left(\mathrm{e}^{4\mathrm{i}\phi}+\mathrm{e}^{-4\mathrm{i}\phi}+6\right)+8\cos^2\theta-\frac{24}{5}\right]. \tag{F.21}$$

For the matrix elements, all off-diagonal terms vanish except $V_{2,-2}$ and $V_{-2,2}$. The evaluation of the non-zero matrix elements are straightforward but tedious. Here is one example:

$$\begin{aligned}V_{1,1} =& \int_0^\infty r^6R^2(r)\mathrm{d}r\frac{D}{8}\int_0^\pi\left(6\sin^2\theta+8\cos^4\theta-\frac{24}{5}\right) \\ &\times 4\sin^3\theta\cos^2\theta\ \mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi \\ =& -4A\end{aligned} \tag{F.22}$$

The eigenvalues and eigenvectors are

Eigenvalue	Eigenvector
$A+5A\left(1+\frac{4\mu_{\mathrm{B}}^2B^2}{25A^2}\right)^{1/2}$	$\frac{(\lvert 2\rangle+x_+\lvert -2\rangle)}{\left(1+x_+^2\right)^{1/2}}$
$-4A+\mu_{\mathrm{B}}B$	$\lvert 1\rangle$
$6A$	$\lvert 0\rangle$
$-4A-\mu_{\mathrm{B}}B$	$\lvert -1\rangle$
$A-5A\left(1+\frac{4\mu_{\mathrm{B}}^2B^2}{25A^2}\right)^{1/2}$	$\frac{(\lvert 2\rangle+x_-\lvert -2\rangle)}{\left(1+x_-^2\right)^{1/2}}$

where

$$x_\pm=\pm\left(1+\left(\frac{2\mu_BB}{5A}\right)^2\right)^{1/2}-\frac{2\mu_{\mathrm{B}}B}{5A} \tag{F.23}$$

Hence in the limit of low field, the eigenvalues and eigenvectors are

Eigenvalue	Eigenvector
$6A+\frac{2\mu_{\mathrm{B}}^2B^2}{5A}$	$\frac{(\lvert 2\rangle+\lvert -2\rangle)}{\sqrt{2}}$
$-4A+\mu_{\mathrm{B}}B$	$\lvert 1\rangle$
$6A$	$\lvert 0\rangle$
$-4A-\mu_{\mathrm{B}}B$	$\lvert -1\rangle$
$-4A-\frac{2\mu_{\mathrm{B}}^2B^2}{5A}$	$\frac{(\lvert 2\rangle-\lvert -2\rangle)}{\sqrt{2}}$

and so if $k_{\rm B}T \ll A$, then you only populate the lowest levels which are a triplet if $A > 0$ or a doublet if $A < 0$. Considering these lowest levels only, the partition function can be written

$$Z = \begin{cases} 1 + {\rm e}^{-\mu_{\rm B}B/k_{\rm B}T} + {\rm e}^{\mu_{\rm B}B/k_{\rm B}T} & A > 0 \\ 1 + {\rm e}^{-2\mu_{\rm B}^2 B^2/5Ak_{\rm B}T} & A < 0 \end{cases}, \tag{F.24}$$

and the results follow using $F = -Nk_{\rm B}T\log Z$ and $M = -\partial F/\partial B$.
$I = \frac{3}{2}, A = 2 \times 10^{-6}$eV. Not much can be said about the nuclear moment. Line broadening due to phonons is large at high temperature. Low temperatures are often needed to resolve fine structure.
(3.5) Measure $g_J J = 7$ (for Gd^{3+}), 5 (for Fe^{3+}) and 3 (for Cr^{3+}). This agrees with the predictions of Hund's rules for all of them except Cr^{3+} for which one would expect $J = \frac{3}{2}$ and $g_J = \frac{2}{5}$. This is because orbital quenching occurs and means that $g_J = 2$. (Orbital quenching also occurs for Fe^{3+} but you do not notice because $L = 0$ is predicted from Hund's rules anyway.)
(3.6) The fields for 60 MHz are

^{1}H	1.41 T
^{2}H	9.18 T
^{13}C	5.60 T
^{19} F	1.50 T

(3.7) 0.322 T
(3.8) First derive matrix representations for $\hat{S}_z$ and $\hat{S}_x$ for an $S = 1$ particle:

$$\hat{S}_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{F.25}$$

$$\hat{S}_x = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \tag{F.26}$$

Hence, the Hamiltonian is

$$\hat{\mathcal{H}} = \begin{pmatrix} D + g_\parallel \mu_{\rm B} B\cos\theta & \frac{g_\perp \mu_{\rm B} B \sin\theta}{\sqrt{2}} & 0 \\ \frac{g_\perp \mu_{\rm B} B \sin\theta}{\sqrt{2}} & 0 & \frac{g_\perp \mu_{\rm B} B \sin\theta}{\sqrt{2}} \\ 0 & \frac{g_\perp \mu_{\rm B} B \sin\theta}{\sqrt{2}} & D - g_\parallel \mu_{\rm B} B\cos\theta \end{pmatrix} \tag{F.27}$$

and the eigenvalues can be found to give the required result. When $\theta = 0$ the eigenvalues simplify to

$$E = 0, D \pm g_\parallel \mu_{\rm B} B \tag{F.28}$$

When $\theta = \pi/2$ the eigenvalues simplify to

$$E = D, \frac{D}{2} \pm \sqrt{\left(\frac{D}{2}\right)^2 + (g_\perp \mu_{\rm B} B)^2}. \tag{F.29}$$

(3.9) 82.4 ms^{-1}; 7.7×10^{-19} ms^{-1}; 9.4×10^{14} Hz; 9.1×10^{-9} Hz; 2.3×10^7 Hz; 0.1μeV; 7.7×10^{-4} cm^{-1}.
(3.10) If the lowest frequency you can easily measure has a period of 20μ s, it has a frequency of 50 kHz , corresponding to a field of about 4×10^{-4} T. The fraction of muons living for 20μ s

or longer is $\mathrm{e}^{-20/2.2} = 1.1 \times 10^{-4}$ corresponding to just over 10^3 of the 10^7 muons implanted. This result follows from the exponential decay of muons with mean lifetime τ_μ. The fraction living longer than T is given by

$$\frac{\int_T^\infty \mathrm{e}^{-t/\tau_\mu}\mathrm{d}t}{\int_0^\infty \mathrm{e}^{-t/\tau_\mu}\mathrm{d}t} = \mathrm{e}^{-T/\tau_\mu} \tag{F.30}$$

If the pulse-width is 50 ns , destructive interference can occur if the spin-precession period is 100 ns , corresponding to a frequency of 10 MHz , giving an upper limit on the field of ~ 0.07 T.
If $B = 0.4$ T, $f = 54.2$MHz.
(3.11) For Fe^{3+}, $J = \frac{5}{2}$ and $g_J = 2$. The saturation moment is $g_J J\mu_\mathrm{B} = 5\mu_\mathrm{B}$. Susceptibility measures $\mu_\mathrm{eff}^2 = g_J^2 J(J+1)\mu_\mathrm{B}^2$ and hence

$$g_J\sqrt{J(J+1)}\mu_\mathrm{B} = 2 \times \sqrt{\frac{5}{2}\times\frac{7}{2}}\mu_\mathrm{B} = \sqrt{35}\mu_\mathrm{B} = 5.92\mu_\mathrm{B} \tag{F.31}$$

The saturation moment involves $\hat{J}_z$ because it is a measurement of the moment saturated along a particular direction. In contrast, susceptibility involves $\hat{J}^2$.
(4.1) There are several ways to derive this result. One way is to start with the vector potential from dipole 1 which is

$$\mathbf{A}_1 = \frac{\mu_0}{4\pi}\nabla \times \frac{\mu_1}{r} \tag{F.32}$$

and this produces a magnetic field

$$\mathbf{B}_1 = \nabla \times \mathbf{A}_1 = \frac{\mu_0}{4\pi}\left[\nabla\left(\nabla\cdot\frac{\mu_1}{r}\right) - \nabla^2\left(\frac{\mu_1}{r}\right)\right] \tag{F.33}$$

and gives rise to an energy

$$\begin{aligned} E &= -\boldsymbol{\mu}_2\cdot\mathbf{B}_1 \\ &= -\frac{\mu_0}{4\pi}\left[\left(\boldsymbol{\mu}_1\cdot\nabla\right)\left(\boldsymbol{\mu}_1\cdot\nabla\right)\frac{1}{r} - \left(\boldsymbol{\mu}_1\cdot\boldsymbol{\mu}_2\right)\nabla^2\frac{1}{r}\right] \\ &= \frac{\mu_0}{4\pi r^3}\left[\boldsymbol{\mu}_1\cdot\boldsymbol{\mu}_2 - \frac{3\left(\boldsymbol{\mu}_1\cdot\mathbf{r}\right)\left(\boldsymbol{\mu}_1\cdot\mathbf{r}\right)}{r^2}\right] \end{aligned} \tag{F.34}$$

It is also possible to derive the result by using the magnetic scalar potential $\phi_1 = \mu_1\cdot\mathbf{r}/4\pi r^3$ and then $B_1 = -\mu_0\nabla\phi_1$.
(4.2) 1Å : (a) 1.855 T (b) 1.855 mT ; 10Å : (a) 0.927 T (b) 0.927 mT .
(4.3) The lattice constant in Fe (which is bcc) is $a = 2.87$Å. The nearest neighbour distance is therefore $a/\sqrt{2}$ and so the dipoledipole energy is $\sim \left(\mu_0 \times (2.2\mu_\mathrm{B})^2/4\pi r^3\right) = 30\mu$eV. J $\approx k_\mathrm{B}T_\mathrm{C} \sim 0.09$eV, which is about 3000 times larger.
(4.4) The bandwidth ~ 0.05eV and the Coulomb energy ~ 1eV.
J $\sim -\frac{t^2}{U} = -2.5$meV.
Hence if $|\mathrm{J}| = k_\mathrm{B}T_\mathrm{N}$, then $T_\mathrm{N} = 29$ K.
(4.5) The partition function is

$$Z = 1 + \mathrm{e}^{-\beta\Delta}\left(\mathrm{e}^{\beta g\mu_\mathrm{B}B} + 1 + \mathrm{e}^{-\beta g\mu_\mathrm{B}B}\right) \tag{F.35}$$

This can be used to derive F and hence M given by

$$M = \frac{2n\beta\left(g\mu_\mathrm{B}\right)^2 B}{3 + \mathrm{e}^{\beta\Delta}}, \tag{F.36}$$

and the result for $\chi \ll 1$ then follows. Note that at high temperatures $T \gg \Delta/k_{\rm B}, \chi \to n\mu_0 g^2\mu_{\rm B}^2/2k_{\rm B}T$ which is a Curie law. It corresponds to all levels occupied, so that the susceptibility is $\frac{3}{4}$ of an $S=1$ system (for which $S(S+1)=2$) and $\frac{1}{4}$ of an $S=1$ system (for which $S(S+1)=0$) so is $\frac{3}{2}$ of a classical Curie law.
At low temperatures, $T \ll \Delta/k_{\rm B}$, then if $\Delta > 0$ one has $\chi \propto {\rm e}^{-\Delta/k_{\rm B}T} \to 0$ (non-magnetic because only the singlet is occupied), but if $\Delta < 0$ one has $\chi \to 2n\mu_0 g^2\mu_{\rm B}^2/3k_{\rm B}T$ (corresponding to the occupied triplet, $S(S+1)=2$, and so twice a classical Curie law).
(4.6) Hint: First show that

$$\mu_{I1} \cdot \mu_{I2} \propto \hat{I}_{1z}\hat{I}_{2z} + \frac{1}{2}\left(\hat{I}_{1+}\hat{I}_{2-} + \hat{I}_{1-}\hat{I}_{2+}\right) \tag{F.37}$$

and using $\mathbf{r} = r(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, show that

$$\boldsymbol{\mu}_I \cdot \hat{\mathbf{r}} \propto \cos\theta\hat{I}_z + \frac{1}{2}\sin\theta\left[{\rm e}^{-{\rm i}\phi}\hat{I}_+ + {\rm e}^{{\rm i}\phi}\hat{I}_-\right]. \tag{F.38}$$

The results then follow fairly quickly.
(4.8) Note that if $\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2 + \mathbf{S}_3$ then

$$\mathbf{S}^2 = \mathbf{S}_1^2 + \mathbf{S}_2^2 + \mathbf{S}_3^2 + 2\left(\mathbf{S}_1\cdot\mathbf{S}_2 + \mathbf{S}_2\cdot\mathbf{S}_3 + \mathbf{S}_3\cdot\mathbf{S}_1\right) \tag{F.39}$$

Hence the Hamiltonian can be rewritten in terms of $\mathbf{S}^2, \mathbf{S}_1^2, \mathbf{S}_2^2$ and S_3^2 giving $E = \mathrm{J}(6 - S(S+1))$. Hence for the possible values of S the required energies are obtained.
(5.1) $B_{\rm mf} = 2.1\times10^3$ T. The number of atoms per unit volume in Fe are $n = 8.49\times10^{28}$ m^{-3}, and using $M = n\cdot 2.2\mu_{\rm B}$ then $\mu_0 M = 2.2$ T. Hence $B_{\rm mf}$ is 10^3 times bigger than $\mu_0 M$.
(5.2) The magnetization is given by

$$\frac{M}{M_{\rm s}} = \mathrm{B}_S(y) = \alpha y - \beta y^3 + \cdots, \tag{F.40}$$

where α and β are constants and

$$y = \frac{\gamma M}{T M_{\rm S}}, \tag{F.41}$$

where γ is a positive constant given by $T_{\rm C}/\alpha$. Near $T_{\rm C}$, these equations lead to

$$\left(\frac{M}{M_{\rm S}}\right)^2 \approx \frac{\alpha^3}{\beta}\frac{(T_{\rm C}-T)}{T_{\rm C}} \tag{F.42}$$

and this leads to the required result using the values of α and β that can be obtained from eqn 5.38.
The energy is given by $E = -\frac{1}{2}\lambda M^2$ and using eqn F .42 one can show that at $T_{\rm C}$

$$\frac{{\rm d}\left(M^2\right)}{{\rm d}T} = -\frac{M_{\rm s}^2\alpha^3}{\beta T_{\rm C}}, \tag{F.43}$$

and hence the result for the specific heat can be found.
In general one can now easily show that

$$C = \begin{cases} \frac{5}{2}nk_{\rm B}\left[\frac{(2S+1)^2-1}{(2S+1)^2+1}\right]\left(3\left(\frac{T}{T_{\rm C}}\right)^2 - 2\left(\frac{T}{T_{\rm C}}\right)\right) & T \le T_{\rm C} \\ 0 & T > T_{\rm C}. \end{cases} \tag{F.44}$$

(5.4) The energy is given by

$$E = -2NS^2\left(\,\mathrm{J}_1\cos\theta + \mathrm{J}_2\cos 2\theta\right), \tag{F.45}$$

and energy minimization $\partial E/\partial\theta = 0$ implies that either $\sin\theta = 0$ (and hence $\theta = 0$ or π) or $\cos\theta = -\mathrm{J}_1/4\ \mathrm{J}_2$. These solutions are ferromagnetic ($\theta = 0$), antiferromagnetic ($\theta = \pi$) and helimagnetic ($\cos\theta = -\mathrm{J}_1/4\ \mathrm{J}_2$) and substitution back into the equation for the energy reproduces eqn 5.45 . The helimagnetic solution is only possible if $|J_1| < 4\,|J_2|$.
Which solution gives the lowest energy can be evaluated by comparing the energies of each solution. Alternatively one can look at the stability of each solution by evaluating $\partial^2 E/\partial\theta^2$.
(5.5) When $y \gg 1$,

$$\tanh y = \frac{\mathrm{e}^{y} - \mathrm{e}^{-y}}{\mathrm{e}^{y} + \mathrm{e}^{-y}} = \frac{1 - \mathrm{e}^{-2y}}{1 + \mathrm{e}^{2y}} \approx 1 - 2\mathrm{e}^{-2y}. \tag{F.46}$$

This can be used together with

$$\frac{M}{M_{\mathrm{S}}} = \tanh y, \tag{F.47}$$

and also with the fact that near T_{C} one has

$$\frac{M}{M_{\mathrm{S}}} = \frac{T}{T_{\mathrm{C}}} y, \tag{F.48}$$

to derive the final result.
Recall the mean-field theory for a ferromagnet. In that case the magnetization is given by

$$M = \frac{C}{\mu_0 T}\left(\lambda M + \mu_0 H\right) \tag{F.49}$$

where C is the Curie constant. This rearranges to

$$M\left(1 - \frac{\lambda C}{\mu_0 T}\right) = \frac{HC}{T} \tag{F.50}$$

and so the susceptibility is

$$\chi = \frac{M}{H} = \frac{C}{T - \theta} \tag{F.51}$$

where the Weiss temperature is $\theta = \lambda C/\mu_0$.
For this ferrimagnet, the magnetization in each sublatice is

$$M_1 = \frac{C_1}{\mu_0 T}\left(\mu_0 H - \lambda M_2\right) \tag{F.52}$$

$$M_2 = \frac{C_2}{\mu_0 T}\left(\mu_0 H - \lambda M_1\right) \tag{F.53}$$

which becomes

$$\begin{pmatrix} T & \lambda C_1/\mu_0 \\ \lambda C_2/\mu_0 & T \end{pmatrix}\begin{pmatrix} M_1 \\ M_2 \end{pmatrix} = H\begin{pmatrix} C_1 \\ C_2 \end{pmatrix}. \tag{F.54}$$

The determinant of the matrix in this equation is therefore zero if $T^2 - \lambda^2 C_1 C_2/\mu_0 = 0$ and hence one can deduce that the paramagnetic regime is when $T > \theta$ where $\theta = \lambda\left(C_1 C_2\right)^{1/2}/\mu_0$. The susceptibility is given by $\chi = \left(M_1 + M_2\right)/H$ which can be found by inverting eqn F. 54 to give

$$\begin{pmatrix} M_1 \\ M_2 \end{pmatrix} = \frac{H}{T^2 - \theta^2}\begin{pmatrix} T & -\lambda C_1/\mu_0 \\ -\lambda C_2/\mu_0 & T \end{pmatrix}\begin{pmatrix} C_1 \\ C_2 \end{pmatrix} \tag{F.55}$$

and hence the final answer.
The first part follows straightforwardly from an application of $\hat{S}^{\pm} = \hat{S}^x \pm \mathrm{i}\hat{S}^y$.
For the Heisenberg ferromagnet, the term $\hat{S}_i^z \hat{S}_j^z$ operating on the ground state produces S^2 and the ground state is an eigenstate of $\hat{S}_i^z \hat{S}_j^z$. There are N such terms and hence $E = -NS^2$. The terms $\hat{S}_i^+ \hat{S}_j^-$ and $\hat{S}_i^- \hat{S}_j^+$ produce nothing because in each case the raising operator annihilates the ground state. Thus the ground state is an eigenstate of both $\hat{S}_i^+ \hat{S}_j^-$ and $\hat{S}_i^- \hat{S}_j^+$ with eigenvalue zero.
For the Heisenberg antiferromagnet the two terms $\hat{S}_i^- \hat{S}_j^-$ and $\hat{S}_i^- \hat{S}_j^+$ produce states other than the 'obvious' ground state. Therefore this 'obvious' ground state is not an eigenstate of the system.
(5.8) The range of possible momentum transfer is 0 nm^{-1} to $\sqrt{k_{\mathrm{in}}} = 2\sqrt{2}\pi/\lambda = 29.6$ nm^{-1}.
(a) At 100 K , just get the chemical Bragg reflections at

$$Q = \left[\left(\frac{2\pi}{a} \right)^2 \left(h^2 + k^2 \right) + \left(\frac{2\pi}{c} \right)^2 l^2 \right]^{1/2}. \tag{F.56}$$

The symmetry is bcc, so reflections are seen when $h + k + l$ is even.
(b) At 10 K , also see the magnetic Bragg reflections.

The observed reflections for both cases are listed below. The magnitude of the scattering vector $\mathbf{Q}$ are listed, together with 2θ, the angle through which the incident beam is scattered. The magnetic reflections are indicated by an asterisk and the multiplicity, M, of each reflection is given in the final column.

h	k	l	Q (nm^{-1})	$2\theta\,(^\circ)$		M
0	1	0	12.566	34.915	*	4
1	0	0	12.566	34.915	*	
1	1	0	17.772	50.208		4
0	0	1	20.944	60.000	*	2
0	1	1	24.425	71.337		8
1	0	1	24.425	71.337		
0	2	0	25.133	73.740		4
2	0	0	25.133	73.740		
1	1	1	27.468	81.952	*	8
1	2	0	28.099	84.261	*	8
2	1	0	28.099	84.261	*	

(5.9) (a) The first expression for $\tilde{m}$ follows from the definition of $\tanh y$. This can be rearranged to give

$$x^2 = \frac{1 + \tilde{m}}{1 - \tilde{m}}, \tag{F.57}$$

and taking the log of this equation gives eqn 5.57. Using eqn F. 48 leads to eqn 5.58.
(b) Rearranging eqn 5.60 to make x the subject leads to

$$(1 - \tilde{m})x^2 - \tilde{m}x - (1 + \tilde{m}) = 0, \tag{F.58}$$

which is a quadratic. This has real solutions if $4 - 3\tilde{m}^2 \geq 0$ and are simply found to be eqn 5.62 . Now $1 \leq 4 - 3\tilde{m}^2 \leq 2$ if $0 \leq \tilde{m} \leq 1$, so for $x > 0$ we need the positive solution in eqn 5.62. This then yields

$$y = \log\left(\tilde{m} + \sqrt{4 - 3\tilde{m}^2} \right) - \log(2(1 - \tilde{m})), \tag{F.59}$$

and in this case one finds $y = 3\tilde{m}T_{\rm C}/2T$ and the final result follows.
6.1) The average energy is found to be

$$\langle E\rangle = -\frac{N\ \mathrm{J}}{4}\tanh\left(\frac{\mathrm{J}}{2k_{\rm B}T}\right), \tag{F.60}$$

and the heat capacity per spin is then

$$C = \frac{\mathrm{J}^2}{4k_{\rm B}T^2\cosh^2(\ \mathrm{J}/2k_{\rm B}T)}. \tag{F.61}$$

This behaves like $\mathrm{J}^2/4k_{\rm B}T^2$ at high temperature, and as $\mathrm{J}^2\mathrm{e}^{-\mathrm{J}/k_{\rm B}T}/4k_{\rm B}T^2$ at low temperature. It is thus zero at both $T = 0$ and $T = \infty$ and exhibits a broad maximum in between. There is thus no cusp and this reinforces the point that there are no phase transitions in purely one-dimensional models.
(6.2) (b) $\hbar\omega(\mathbf{q}) = S[\ \mathrm{J}(\mathbf{0}) - \mathrm{J}(\mathbf{q}) + K(\mathbf{0})]$.
(c) (i) $\hbar\omega(\mathbf{q}) = 2SJ_0(1-\cos qa) + 2SK_0$.
(ii) $\hbar\omega(\mathbf{q}) = 2SJ_0\,(2-\cos q_x a - \cos q_y a) + 4SK_0$.
(iii)

$$\hbar\omega(\mathbf{q}) = 8SJ_0\left(1-\cos\frac{q_x a}{2}\cos\frac{q_y a}{2}\cos\frac{q_z a}{2}\right) + 8SK_0.$$

6.3) At small q one can use $\cos qa \approx 1-(qa)^2/2$ and obtain
(i) $\hbar\omega = S\,(J_0a^2q^2 + 2K_0)$
(ii) $\hbar\omega = S\,(\ \mathrm{J}_0a^2q^2 + 4K_0)$
(iii) $\hbar\omega = S\,(J_0a^2q^2 + 8K_0)$

In the Ising case ($J_0 = 0$), one has $\hbar\omega = 2dSK_0 \equiv \Delta$ where $d = 1, 2, 4$ for cases (i), (ii) and (iii) respectively. In this model, spin waves 'cost' the same energy Δ regardless of wave vector. The number of spin waves is then proportional to $\mathrm{e}^{-\Delta/k_{\rm B}T}$, and so their energy is proportional to $\Delta\mathrm{e}^{-\Delta/k_{\rm B}T}$. The specific heat is then easily found to be proportional to

$$\frac{\Delta^2}{k_{\rm B}T^2}\mathrm{e}^{-\Delta/k_{\rm B}T},$$

in agreement with Exercise 6.1 with Δ in this problem being equal to J in that problem.
In the Heisenberg case ($K_0 = 0$), $\hbar\omega = Dq^2$ where the spinstiffness $D = SJ_0a^2$. The number of spin waves N_s is given by

$$N_s \propto \int \frac{\mathbf{d}^n q}{\mathbf{e}^{Dq^2/k_{\mathbf{B}}T} - 1}, \tag{F.62}$$

and at small q

$$\frac{1}{\mathrm{e}^{Dq^2/k_{\rm B}T}-1} \approx \frac{k_{\rm B}T}{Dq^2}. \tag{F.63}$$

The integral therefore diverges in one or two dimensions (where $\mathrm{d}^n q$ is equal to $\mathrm{d}q$ or $2\pi q\ \mathrm{d}q$ respectively), but converges for three dimensions (where $\mathrm{d}^n q$ is equal to $4\pi q^2\ \mathrm{d}q$). No long range order is therefore possible for the pure (isotropic) Heisenberg model in one or two dimensions.
(6.4) In the antiferromagnetic case, one can follow the derivation for the ferromagnetic case, but you have to remember which sublattice you are on. One finds that

$$\begin{array}{ll} \hbar\dot{S}_j^+ = 2\mathrm{i}S\left[2S_j^+ + S_{j-1}^+ + S_{j+1}^+\right] & j \text{ odd} \\ \hbar\dot{S}_j^+ = -2\mathrm{i}S\left[2S_j^+ + S_{j-1}^+ + S_{j+1}^+\right] & j \text{ even.} \end{array} \tag{F.64}$$

Hence substituting in

$$S_j^+ = \begin{cases} u\mathrm{e}^{\mathrm{i}(qja-\omega t)} & j \text{ odd} \\ v\mathrm{e}^{\mathrm{i}(qja-\omega t)} & j \text{ even} , \end{cases} \tag{F.65}$$

one finds that

$$\mathrm{i}\hbar\omega \begin{pmatrix} u \\ v \end{pmatrix} = 4\mathrm{j}JS \begin{pmatrix} 1 & \cos qa \\ -\cos qa & -1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}. \tag{F.66}$$

and hence

$$\begin{vmatrix} 1 - \frac{\hbar\omega}{4JS} & \cos qa \\ \cos qa & 1 + \frac{\hbar\omega}{4JS} \end{vmatrix} = 0 \tag{F.67}$$

so that $\hbar\omega = 4JS \sin qa$.

(6.5) Using $\partial F/\partial M = 0$, some rearrangement quickly yields

$$M^2 = -\frac{2a_0\,(T - T_\mathrm{C})}{4b} + \frac{\mu_0 H}{4bM}, \tag{F.68}$$

which is of the required form. Plotting M^2 against M/H gives a straight-line, with the intercept on the M^2 axis changing sign at $T = T_\mathrm{C}$.

The model in appropriate units is

$$\epsilon = \frac{1}{2}\sin^2(\theta - \phi) - h\cos\phi. \tag{F.69}$$

For $\theta = 0$, energy minimization gives

$$-\ \sin\phi(\cos\phi + h) = 0 \tag{F.70}$$

for which the solutions are $\phi = 0, \pi, \cos^{-1}(-h)$. Evaluation of $\partial^2 E/\partial\theta^2$ can be used to examine the stability of these solutions. For $\theta = \pi/2$, energy minimization gives

$$\sin\phi(h - \cos\phi) = 0 \tag{F.71}$$

for which the solutions are $\phi = 0, \pi, \cos^{-1}(h)$.

(6.7) The components of magnetization are

$$M_x = |\mathbf{M}| \sin\frac{\pi r}{R} \tag{F.72}$$

$$M_y = |\mathbf{M}| \cos\frac{\pi r}{R}, \tag{F.73}$$

and using results such as

$$\frac{\mathrm{d}r}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}\sqrt{x^2 + y^2 + z^2} = \frac{x}{r} \tag{F.74}$$

one can evaluate all the partial differential of M_x and M_y. Hence

$$(\nabla M_x)^2 + (\nabla M_y)^2 = |\mathbf{M}|^2 \frac{\pi^2}{R^2}, \tag{F.75}$$

from which the final result follows.

25Gbitsin^{-2} is 3.9×10^{13} bits m^{-2} in sensible units.

(a) Taking πa_0^2 to be the area of a hydrogen atom gives 3.4×10^{-7} bits (hydrogen atom area)$^{-1}$.

(b) Most Shakespearian plays are $20000 - 40000$ words long; with maybe an average 5 letters per word, and about 40 plays, one gets about 4 Mbytes or 32 Mbits (8 bits per byte). Taking a UK stamp, I calculated about 600 copies of the complete works per stamp, although that could be easily doubled with data-compression. "I am ill at all these numbers" (Hamlet Act 2 Scene 2)
DNA stores 1 bit per base pair, which is about 1 bit per cubic nanometre. If you made a sheet of this, this would be equivalent to about $6 \times 10^5 \mathrm{Gbits in}^{-2}$, but the great advantage of DNA is that it can coil and is a volume storage technique.
(6.9) The best way to tackle this is to put the magnetic field parallel to $-z$ which you can do without loss of generality. (Why $-z$? So that the ground state has all it spins aligned parallel to z. Electrons have negative charge and so their spins are antiparallel to their magnetic moments and will therefore align antiparallel to the applied field.) You should find that each magnon adds an energy $+g\mu_{\mathrm{B}}B$.
(6.11) The equation

$$\langle q \mid i \rangle = \frac{1}{\sqrt{N}} \sum_j \mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\mathbf{R}_j} \langle j \mid i \rangle \tag{F,76}$$

can be simplified using $\langle j \mid i \rangle = \delta_{ij}$. The result then follows.
Hint for the next part: S^+kills the state it is operating on unless the spin it is operating on is $< S$.
The exercise shows that each spin has a small transverse component which is perpendicular to the direction of magnetization.
The transverse components vary in space in just the manner depicted in Fig. 6.13.
(6.12) It is necessary to solve

$$\Delta E = 0 = \int_{-\infty}^{\infty} \left[2A \frac{\partial\theta}{\partial z} \frac{\partial}{\partial z}(\delta\theta) + \frac{\partial f(\theta)}{\partial\theta} \delta\theta \right] \mathrm{d}z \tag{F.77}$$

The first part of this integral can be evaluated by parts, yielding

$$\frac{\partial f}{\partial\theta} - 2A \frac{\partial^2 f}{\partial z^2} = 0. \tag{F.78}$$

The question carefully guides the reader through the rest.
(6.13) The wall energy per unit volume due to the length of walls in the structure and is therefore approximately $\sigma_{\mathrm{w}} L/D$. The spins inside the small triangles cost an anisotropy energy. The area of one of these triangles is $D^2/4$ and there are two of them in every horizontal length D, so their total cost is $2 \times K \times (D^2/4) \times (1/D) = KD/2$. Summing these terms and differentiating yields the required result. $D = 20\mu$ m.
(7.1) Using $g\left(E_{\mathrm{F}}\right) = 3n/2E_{\mathrm{F}}$, the first result follows quite quickly. To consider the temperature dependence, it is worthwhile to review some mathematical details concerning the Fermi function. Consider $h(E) = \int_0^E g\left(E'\right) \mathrm{d}E'$, so that $g(E) = \mathrm{d}h/\mathrm{d}E$ and consider the integral

$$\begin{aligned} I &= \int_0^\infty \frac{\mathrm{d}h}{\mathrm{d}E} f(E) \mathrm{d}E \\ &= [f(E)g(E)]_0^\infty - \int_0^\infty h(E) \frac{\mathrm{d}f}{\mathrm{d}E}\, \mathrm{d}E \\ &= -\int_0^\infty h(E) \frac{\mathrm{d}f}{\mathrm{d}E}\, \mathrm{d}E \end{aligned} \tag{F.79}$$

Now put $x = (E - \mu)/k_{\mathrm{B}}T$ and hence

$$\frac{\mathrm{d}f}{\mathrm{d}E} = -\frac{1}{k_\mathrm{B}T}\frac{\mathrm{e}^x}{\left(\mathrm{e}^x+1\right)^2}. \tag{F.80}$$

Writing

$$h(E) = \sum_{s=0}^{\infty}\frac{x^s}{s!}\left(\frac{\mathrm{d}^s h}{\mathrm{d}x^s}\right)_{x=0}, \tag{F.81}$$

we have that

$$I = \sum_{s=0}^{\infty}\frac{1}{s!}\left(\frac{\mathrm{d}^s h}{\mathrm{d}x^s}\right)_{x=0}\int_{-E_\mathrm{F}/k_\mathrm{B}T}^{\infty}\frac{x^s\mathrm{e}^x\,\mathrm{d}x}{\left(\mathrm{e}^x+1\right)^2}. \tag{F.82}$$

The integral part of this can be simplified by replacing the lower limit by $-\infty$. It vanishes for odd s, but for even s

$$\begin{aligned}
\int_{-\infty}^{\infty}\frac{x^s\mathrm{e}^x\,\mathrm{d}x}{\left(\mathrm{e}^x+1\right)^2} =& 2\int_0^{\infty}\frac{x^s\mathrm{e}^x\,\mathrm{d}x}{\left(\mathrm{e}^x+1\right)^2}\\
=& 2\int_0^{\infty}\mathrm{d}x\sum_{n=0}^{\infty}\mathrm{e}^x x^s\\
&\times\left[(n+1)(-1)^{n+1}\mathrm{e}^{-nx}\right]\\
=& 2\sum_{n=1}^{\infty}(-1)^{n+1}n\int_0^{\infty}x^s\mathrm{e}^{-nx}\,\mathrm{d}x\\
=& 2(s!)\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n^s}\\
=& 2(s!)\left(1-2^{1-s}\right)\zeta(s)
\end{aligned} \tag{F.83}$$

where $\zeta(s)$ is the Riemann zeta function. $\left(\zeta(0)=-\frac{1}{2}, \zeta(2)=\frac{\pi^2}{6}, \zeta(4)=\frac{\pi^4}{90}\right)$.
Thus the integral is

$$\begin{aligned}
I =& \sum_{s=0, s\text{ even}}^{\infty} 2\left(\frac{\mathrm{d}^s h}{\mathrm{d}x^s}\right)_{x=0}\left(1-2^{1-s}\right)\zeta(s)\\
=& h + \frac{\pi^2}{6}\left(\frac{\mathrm{d}^2 h}{\mathrm{d}x^2}\right)_{x=0} + \frac{7\pi^4}{360}\left(\frac{\mathrm{d}^4 h}{\mathrm{d}x^4}\right)_{x=0} + \cdots\\
=& \int_{-\infty}^{\mu} g(E)\mathrm{d}E + \frac{\pi^2}{6}\left(k_\mathrm{B}T\right)^2\left(\frac{\mathrm{d}g}{\mathrm{d}E}\right)_{E=\mu}\\
&+ \frac{7\pi^4}{360}\left(k_\mathrm{B}T\right)^4\left(\frac{\mathrm{d}^3 g}{\mathrm{d}E^3}\right)_{E=\mu} + \ldots
\end{aligned} \tag{F.84}$$

This implies that the number density is

$$n = \int_0^{\mu} g(E)\mathrm{d}E + \frac{\pi^2}{6}\left(k_\mathrm{B}T\right)^2\left(\frac{\mathrm{d}g}{\mathrm{d}E}\right)_{E=\mu} + \ldots \tag{F.85}$$

Now to first order

$$\int_0^{\mu} g(E)\mathrm{d}E = \int_0^{E_\mathrm{F}} g(E)\mathrm{d}E + \left(\mu - E_\mathrm{F}\right)g\left(E_\mathrm{F}\right) \tag{F.86}$$

so for a constant number of electrons, n takes the same value at all temperatures and so

$$\mu = E_{\mathrm{F}} - \frac{\pi^2}{6}\left(k_{\mathrm{B}}T\right)^2 \frac{g'\left(E_{\mathrm{F}}\right)}{g\left(E_{\mathrm{F}}\right)} \tag{F.87}$$

The rôle of a magnetic field is then to shift the energy levels by $\pm\mu_{\mathrm{B}}B$. The number density for each spin state is given by

$$\begin{aligned} n_{\pm} =& \frac{1}{2}\int_0^{E_{\mathrm{F}}} g(E)\mathrm{d}E + \left(\mu - E_{\mathrm{F}} \mp \mu_{\mathrm{B}}B\right) g\left(E_{\mathrm{F}} \mp \mu_{\mathrm{B}}B\right) \\ &+ \frac{\pi^2}{6}\left(k_{\mathrm{B}}T\right)^2 \left(\frac{\mathrm{d}g}{\mathrm{d}E}\right)_{E=\mu-mp\mu_{\mathrm{B}}B} + \ldots \end{aligned} \tag{F.88}$$

and using the approximate relations

$$g\left(E_{\mathrm{F}} \mp \mu_{\mathrm{B}}B\right) = g\left(E_{\mathrm{F}}\right) \mp \mu_{\mathrm{B}}B\left(\frac{\mathrm{d}g}{\mathrm{d}E}\right)_{E=\mu} \tag{F.89}$$

and

$$\left(\frac{\mathrm{d}g}{\mathrm{d}E}\right)_{E=\mu\mp\mu_{\mathrm{B}}B} = \left(\frac{\mathrm{d}g}{\mathrm{d}E}\right)_{E=\mu} \mp \mu_{\mathrm{B}}B\left(\frac{\mathrm{d}^2g}{\mathrm{d}E^2}\right)_{E=\mu} \tag{F.90}$$

one then has
$n_+ - n_- = -\frac{\pi^2}{6}\left(k_{\mathrm{B}}T\right)^2 \mu_{\mathrm{B}}B\left[\left(\frac{g'}{g}\right)^2 - \frac{g''}{g}\right] + \mu_{\mathrm{B}}Bg\left(E_{\mathrm{F}}\right).$
Hence the susceptibility is

$$\chi = \mu_0\mu_{\mathrm{B}}g\left(E_{\mathrm{F}}\right)\left[1 - \frac{\pi^2}{6}\left(k_{\mathrm{B}}T\right)^2\left[\left(\frac{g'}{g}\right)^2 - \frac{g''}{g}\right]\right]$$

For $g(E) \propto E^{1/2}$ it is simple to show that

$$\begin{aligned} \frac{g'}{g} &= \frac{1}{2E_{\mathrm{F}}} \\ \left(\frac{g'}{g}\right)^2 - \frac{g''}{g} &= \frac{1}{2E_{\mathrm{F}}^2} \end{aligned} \tag{F.93}$$

evaluated at the Fermi energy. The final result then follows.
For a metal with $E_{\mathrm{F}} \sim 10$eV (e.g. Sn), $T_{\mathrm{F}} \sim 10^5$ K, the correction is about 10^{-5}.
(7.2) (a) This is shown in the derivation of eqn 7.23.
(b) The crossover temperature is $2T_{\mathrm{F}}/3 = 27$ K. The Pauli paramagnetic susceptibility is then 1.3×10^{-8}. The Landau diamagnetic susceptibility is then

$$\chi_{\mathrm{L}} = -\left(\frac{m_e}{m^*}\right)^2 \frac{\chi_{\mathrm{P}}}{3} = -4.3 \times 10^{-7} \tag{F.94}$$

(7.3) Using $A_\phi = \frac{1}{2}Br$, the Schrödinger equation becomes

$$\begin{aligned} &\left[\frac{-\hbar^2}{2m_e}\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\phi^2} + \frac{\partial^2}{\partial z^2}\right) - \frac{\mathrm{i}\hbar eB}{2m_e}\frac{\partial}{\partial\phi}\right. \\ &\left. + \frac{e^2B^2r^2}{8m_e}\right]\psi(r,\phi,z) = E\psi(r,\phi,z) \end{aligned} \tag{F.95}$$

One can then look for solutions of the form

$$\psi(r,\phi,z) = R(r)\mathrm{e}^{il\phi}\mathrm{e}^{\mathrm{i}k_z z} \tag{F.96}$$

and this leads to

$$\left(-\frac{\hbar^2}{2m_e}\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}\right)-\frac{\hbar^2 l^2}{2m_e}+\frac{1}{2}m_e\omega_{\mathrm{c}}^2 r^2 + \frac{\hbar k_z^2}{2m_e}\right)R(r) = \left(E-\frac{l}{2}\hbar\omega_{\mathrm{c}}\right)R(r). \tag{F.97}$$

The solutions are

$$\psi(r,\theta,z) = \exp\left[\mathrm{i}\left(l\phi+k_z z\right)-\frac{eBr^2}{4\hbar}\right]r^{|l|}L_{n-1}^{|l|}\left(\frac{eBr^2}{2\hbar}\right) \tag{F.98}$$

where $L_\alpha^\beta(x)$ is an associated Laguerre polynomial (Appendix C.5), and the energy is

$$E = \left(n+\frac{l}{2}+\left|\frac{l}{2}\right|-\frac{1}{2}\right)\hbar\omega_{\mathrm{c}}+\frac{\hbar k_z^2}{2m_e}. \tag{F.99}$$

In this so-called symmetric gauge, all states circulate about the same point in space, so that rotational symmetry is captured at the expense of translational symmetry.
(7.5) Hints:

$$\frac{z+1}{z-1} = 1+\frac{2}{z-1} \tag{F.100}$$

and hence one can show

$$1+\frac{1-z^2}{2z}\log\left(\frac{z+1}{z-1}\right) = \frac{2}{3z^2}+O\left(\frac{1}{z^3}\right) \tag{F.101}$$

Also, $\log(-1) = \mathrm{i}\pi$, so that

$$\log\left|\frac{x+1}{x-1}\right| - \log\left(\frac{x+1}{x-1}\right) = \mathrm{i}\pi \tag{F.102}$$

(7.6) (b) It is useful to start with

$$\chi_q = \mu_0\mu_{\mathrm{B}}^2\frac{2m_e}{\hbar^2}I$$

where I is an integral given by

$$I = \int_{|\mathbf{k}|<k\mathbf{F}} g(\mathbf{k})\mathrm{d}^n k\left[\frac{1}{(k+q)^2-k^2}+\frac{1}{(k-q)^2-k^2}\right] \tag{F.103}$$

In one-dimension

$$\begin{aligned}\int_{|\mathbf{k}|<k_\mathrm{F}} \frac{\mathrm{d}k}{(k+q)^2-k^2} &= \int_{-k_\mathrm{F}}^{k_\mathrm{F}} \frac{\mathrm{d}k}{q(q+2k)} \\ &= \frac{1}{2q}\int_{q-2k_\mathrm{F}}^{q+2k_\mathrm{F}} \frac{\mathrm{d}u}{u} \\ &= \frac{1}{2q}\int_{|q-2k_\mathrm{F}|}^{q+2k_\mathrm{F}} \frac{\mathrm{d}u}{u} \\ &= \frac{1}{2q}\log\left|\frac{q+2k_\mathrm{F}}{q-2k_\mathrm{F}}\right| \end{aligned} \tag{F.104}$$

The result then follows.
In two-dimensions, the relevant integral to consider is

$$\begin{aligned} I &= \int_0^{k_\mathrm{F}} k\ \mathrm{d}k \int_0^{2\pi} \frac{\mathrm{d}\theta}{q(q-2k\cos\theta)} + \frac{\mathrm{d}\theta}{q(q+2k\cos\theta)} \\ &= 2\int_0^{k_\mathrm{F}} k\ \mathrm{d}k \int_0^{2\pi} \frac{\mathrm{d}\theta}{q^2-4k^2\cos\theta} \end{aligned} \tag{F.105}$$

Now

$$\int_0^{2\pi} \frac{\mathrm{d}\theta}{q^2-4k^2\cos\theta} = \begin{cases} \frac{2\pi}{q(q^2-4k^2)^{1/2}} & q > 2k_\mathrm{F} \\ 0 & q < 2k_\mathrm{F} \end{cases} \tag{F.106}$$

so that the integral becomes

$$I = \begin{cases} \pi\left[1-\sqrt{1-\left(\frac{2k_\mathrm{F}}{q}\right)}\right] & q > 2k_\mathrm{F} \\ \pi & q < 2k_\mathrm{F} \end{cases} \tag{F.107}$$

and the final result follows.

9.15 Symbols, constants and useful equations

a unit cell dimension
a_0 Bohr radius
A area
A hyperfine coupling constant
A continuum exchange constant
A mass number
A magnetic vector potential
b scattering length
B magnetic flux density (magnetic field)
$\mathbf{B}_\mathrm{a}$ applied field
$\mathbf{B}_\mathrm{d}$ demagnetizing field
$\mathbf{B}_1$ internal field
$\mathbf{B}_\mathrm{mf}$ molecular field
B_J Brillouin function
c velocity of light in free space (2.9979×10^8 m s^{-1})
D electric displacement

D axial anisotropy constant
dS surface element
dS entropy change
e $= \exp(1) = 2.718281828$
e magnitude of the electron charge (1.6022×10^{-19}C)
E energy
E_{F} Fermi energy
E_{S} singlet energy
E_{T} triplet energy
ε electric field
F Helmholtz free energy (Helmholtz function)
$g(E)$ density of states in energy space
$g(k)$ density of states in wave vector space
g g-factor
g_c electron g-factor
g_I nuclear g-factor
g_1 g-factor for orbital angular momentum
g_s g-factor for spin angular momentum
g effective g-tensor
G torque
G reciprocal lattice vector
h Planck's constant (6.626×10^{-34} J s)
$\hbar$ Planck's constant $/2\pi$ (1.0546×10^{-34} J s)
H magnetic field strength (magnetic field)
$\mathbf{H}_{\mathrm{a}}$ applied field
$\mathbf{H}_{\mathrm{d}}$ demagnetizing field
$\mathbf{H}_i$ internal field
$\mathcal{H}$ Hamiltonian
i $= \sqrt{-1}$
I current
I nuclear spin quantum number
J total angular momentum quantum number
J exchange constant/integral
k wave vector
k_{B} Boltzmann's constant $\left(1.3807 \times 10^{-23}\ \mathrm{J\ K^{-1}}\right)$
k_{F} Fermi wave vector
K Coulomb integral
K anisotropy constant
K Knight shift
K spring constant
K_{S} surface anisotropy constant
K_{V} volume anisotropy constant
l orbital angular momentum quantum number
l Landau level index
m_{e} electron mass (9.109×10^{-31} kg)
m_l orbital magnetic quantum number
m_I nuclear magnetic quantum number
m_J magnetic quantum number
m_{n} neutron mass
m_s spin magnetic quantum number

M magnetization (magnetic moment per unit volume)
$\mathbf{M}_{\pm}$magnetization on a sublattice of an antiferromagnet
$\mathbf{M}_{\mathrm{s}}$ saturation magnetization
n principal quantum number
n number of atoms/moments per unit volume
N demagnetizing tensor
N demagnetizing factor
N number of neutrons in an atomic nucleus
N_{A} Avogadro's number
p dipole moment
p canonical momentum
$\hat{\mathbf{p}}$ momentum operator
q magnon wave vector

q	helix wave vector	ϵ_0	electric permittivity of free space (8.854×
Q	scattering vector		$10^{-12}\mathrm{Fm}^{-1}$)
$\mathbf{r}$	position vector	ϵ_r	relative dielectric constant
$\hat{\mathbf{r}}$	position operator	θ	angle
R	electrical resistance	θ	Weiss temperature
R_e	extraordinary Hall coefficient	λ	wavelength
R_o	ordinary Hall coefficient	λ	spin-orbit constant
S	entropy	$\lambda_\text{el-ph}$	electron-phonon coupling constant
s, S	spin quantum number	$\lambda_\text{spin-ph}$	spin-phonon coupling constant
$\hat{\mathbf{S}}$	spin angular momentum operator	μ	magnetic moment
$\hat{S}_+$	raising operator	μ_r	relative magnetic permeability
$\hat{S}$	lowering operator	μ_B	Bohr magneton
$\tilde{\mathbf{S}}$	effective spin	μ_N	nuclear magneton
t	time	μ_eff	effective magnetic moment in μ_B per formula
t	hopping integral		unit
t_p	pulse length	μ_0	magnetic permeability of free space ($4\pi\times$
T	temperature		$10^{-7}\mathrm{Hm}^{-1}$)
T_c	critical temperature	v	frequency
T_C	Curie temperature	ξ	correlation length
T_N	Néel temperature	ρ	resistivity
T_P	Peierls transition temperature	$\rho_\mathbf{H}$	Hall resistivity
T_SP	spin-Peierls transition temperature	σ	conductivity
T_1	spin-lattice relaxation time	σ	Pauli spin matrices
T_2	spin-spin relaxation time	τ	period
T_F	Fermi temperature	τ	superparamagnetic relaxation time
U	Coulomb energy	ϕ	angle
v	velocity	Φ	magnetic flux
v	velocity vector	χ	magnetic susceptibility
V	potential or potential energy	χ_g	mass susceptibility
W	transition rate	χ_L	Landau diamagnetic susceptibility
$Y_{lm_l}(\theta, \phi)$	spherical harmonic	χ_m	molar magnetic susceptibility
Z	partition function	χ_P	Pauli spin susceptibility
Z	atomic number	$\chi_\mathbf{q}$	q -dependent susceptibility
α	fine structure constant	χ	spin wave function
β	$= 1/k_\mathrm{B}T$	ψ	wave function
γ	gyromagnetic ratio	Ψ_S	singlet wave function
γ_e	electron gyromagnetic ratio	Ψ_T	triplet wave function
γ_μ	muon gyromagnetic ratio	ω	angular frequency
γ_N	nuclear gyromagnetic ratio	ω_c	cyclotron frequency
Δ	exchange splitting	ω_L	Larmor precession frequency

9.15.1 Fundamental constants

Bohr radius	a_0	5.292×10^{-11} m
Speed of light in free space	c	2.9979×10^{8} m s^{-1}
Electronic charge	e	1.6022×10^{-19}C
Planck's constant	h	6.626×10^{-34} J s
	$h/2\pi =$	1.0546×10^{-34} J s
Boltzmann's constant	k_{B}	1.3807×10^{-23} J K^{-1}
Electron rest mass	m_{e}	9.109×10^{-31} kg
Proton rest mass	m_{p}	1.6726×10^{-27} kg
Avogadro's number	N_{A}	6.022×10^{23} mol^{-1}
Standard molar volume		22.414×10^{-3} m^3 mol^{-1}
Molar gas constant	R	8.315 J mol^{-1} K^{-1}
Fine structure constant	$\frac{e^2}{4\pi\varepsilon_0\hbar c} =$	$(137.04)^{-1}$
Permittivity of free space	ϵ_0	8.854×10^{-12}Fm^{-1}
Magnetic permeability of free space	μ_0	$4\pi \times 10^{-7}$Hm^{-1}
Bohr magneton	μ_{B}	9.274×10^{-24}Am^2 or JT^{-1}
Nuclear magneton	μ_{N}	5.051×10^{-27}Am^2 or JT^{-1}
Neutron magnetic moment	μ_{n}	$-1.9130\mu_{\mathrm{N}}$
Proton magnetic moment	μ_{p}	$2.7928\mu_{\mathrm{N}}$

9.15.2 Useful equations

(1) Identities:

$$\mathrm{e}^{\mathrm{i}\theta} = \cos\theta + \mathrm{i}\sin\theta$$
$$\sin\theta = \frac{\mathrm{e}^{\mathrm{i}\theta} - \mathrm{e}^{-\mathrm{i}\theta}}{2\mathrm{i}}$$
$$\cos\theta = \frac{\mathrm{e}^{\mathrm{i}\theta} + \mathrm{e}^{-\mathrm{i}\theta}}{2}$$
$$\sin(\theta+\phi) = \sin\theta\cos\phi + \cos\theta\sin\phi$$
$$\cos(\theta+\phi) = \cos\theta\cos\phi - \sin\theta\sin\phi$$
$$\cos^2\theta + \sin^2\theta = 1$$
$$\cos 2\theta = \cos^2\theta - \sin^2\theta$$
$$\sin 2\theta = 2\cos\theta\sin\theta$$
$$\sinh x = \frac{\mathrm{e}^{x} - \mathrm{e}^{-x}}{2}$$
$$\cosh x = \frac{\mathrm{e}^{x} + \mathrm{e}^{-x}}{2}$$

(2) Series expansions (valid for $|x| < 1$):

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + \cdots$$

$$\mathrm{e}^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots$$

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots$$

(3) Integrals:

$$\int_0^\infty x^n \mathrm{e}^{-x}\,\mathrm{d}x = n!$$

$$\int_{-\infty}^\infty \mathrm{e}^{-\alpha x^2}\,\mathrm{d}x = \sqrt{\frac{\pi}{\alpha}}$$

$$\int_{-\infty}^\infty x^{2n} \mathrm{e}^{-\alpha^2 x^2}\,\mathrm{d}x = \frac{(2n)!}{n!2^{2n}\alpha^n}\sqrt{\frac{\pi}{\alpha}}$$

Indefinite (with $a > 0$):

$$\begin{aligned}
\int \frac{\mathrm{d}x}{x^2+a^2} &= \frac{1}{a}\tan^{-1}\frac{x}{a} \\
\int \frac{\mathrm{d}x}{x^2-a^2} &= \frac{1}{2a}\ln\left|\frac{x-a}{x+a}\right| \\
\int \frac{\mathrm{d}x}{\sqrt{x^2+a^2}} &= \sinh^{-1}\frac{x}{a} \\
\int \frac{\mathrm{d}x}{\sqrt{x^2-a^2}} &= \begin{cases} \cosh^{-1}\frac{x}{a} & \text{if } x > a \\ -\cosh^{-1}\frac{x}{a} & \text{if } x < -a \end{cases} \\
\int \frac{\mathrm{d}x}{\sqrt{a^2-x^2}} &= \sin^{-1}\frac{x}{a}
\end{aligned} \tag{G.1}$$

(4) Vector operators:

- grad acts on a scalar field to produce a vector field:

$$\operatorname{grad}\psi = \nabla\psi = \left(\frac{\partial\psi}{\partial x}, \frac{\partial\psi}{\partial y}, \frac{\partial\psi}{\partial z}\right)$$

- div acts on a vector field to produce a scalar field:

$$\operatorname{div}\mathbf{A} = \nabla\cdot\mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$$

- curl acts on a vector field to produce another vector field:

$$\operatorname{curl}\mathbf{A} = \nabla\times\mathbf{A} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ A_x & A_y & A_z \end{vmatrix}$$

where $\psi(\mathbf{r})$ and $\mathbf{A}(\mathbf{r})$ are any given scalar and vector field respectively.
(5) Vector identities:

$$\begin{aligned}
\nabla\cdot(\nabla\psi) &= \nabla^2\psi \\
\nabla\times(\nabla\psi) &= 0 \\
\nabla\cdot(\nabla\times\mathbf{A}) &= 0 \\
\nabla\cdot(\psi\mathbf{A}) &= \mathbf{A}\cdot\nabla\psi + \psi\nabla\cdot\mathbf{A} \\
\nabla\times(\psi\mathbf{A}) &= \psi\nabla\times\mathbf{A} - \mathbf{A}\times\nabla\psi \\
\nabla\times(\nabla\times\mathbf{A}) &= \nabla(\nabla\cdot\mathbf{A}) - \nabla^2\mathbf{A} \\
\nabla\cdot(\mathbf{A}\times\mathbf{B}) &= \mathbf{B}\cdot\nabla\times\mathbf{A} - \mathbf{A}\cdot\nabla\times\mathbf{B} \\
\nabla(\mathbf{A}\cdot\mathbf{B}) &= (\mathbf{A}\cdot\nabla)\mathbf{B} + (\mathbf{B}\cdot\nabla)\mathbf{A} \\
&\quad + \mathbf{A}\times(\nabla\times\mathbf{B}) + \mathbf{B}\times(\nabla\times\mathbf{A}) \\
\nabla\times(\mathbf{A}\times\mathbf{B}) &= (\mathbf{B}\cdot\nabla)\mathbf{A} - (\mathbf{A}\cdot\nabla)\mathbf{B} \\
&\quad + \mathbf{A}(\nabla\cdot\mathbf{B}) - \mathbf{B}(\nabla\cdot\mathbf{A})
\end{aligned}$$

These identities can be easily proved by application of the alternating tensor and use of the summation convention. The alternating tensor ϵ_{ijk} is defined according to:

$$\epsilon_{ijk} = \begin{cases} 1 & \text{if } ijk \text{ is an even permutation of 123} \\ -1 & \text{if } ijk \text{ is an odd permutation of 123} \\ 0 & \text{if any two of } i,j \text{ or } k \text{ are equal} \end{cases}$$

so that the vector product can be written

$$(\mathbf{A}\times\mathbf{B})_i = \epsilon_{ijk}A_jB_k$$

The summation convention is used here, so that twice repeated indices are assumed summed. The scalar product is then

$$\mathbf{A}\cdot\mathbf{B} = A_iB_i$$

Use can be made of the identity

$$\epsilon_{ijk}\epsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}$$

where δ_{ij} is the Kronecker delta given by

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

The vector triple product is given by

$$\mathbf{A}\times(\mathbf{B}\times\mathbf{C}) = (\mathbf{A}\cdot\mathbf{C})\mathbf{B} - (\mathbf{A}\cdot\mathbf{B})\mathbf{C}$$

(6) Cylindrical coordinates:

$$\begin{aligned}
\nabla^2\psi &= \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\psi}{\partial\phi^2} + \frac{\partial^2\psi}{\partial z^2} \\
\nabla\psi &= \left(\frac{\partial\psi}{\partial r}, \frac{1}{r}\frac{\partial\psi}{\partial\phi}, \frac{\partial\psi}{\partial z}\right)
\end{aligned}$$

(7) Spherical polar coordinates:

$$\nabla^2\psi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2}$$

$$\nabla\psi = \left(\frac{\partial\psi}{\partial r}, \frac{1}{r}\frac{\partial\psi}{\partial\theta}, \frac{1}{r\sin\theta}\frac{\partial\psi}{\partial\phi}\right)$$

10 Quantum Theory of Magnetism by White

10.1 Preface

Perhaps as many as ten years ago a colleague of mine pointed out to me that the Second Edition of this book was "out of print". I mentioned this to my editor, who at that time was Helmut Lotsch, and he enthusiastically encouraged me to write a Third Edition. Thus began the long process that has finally resulted in this new edition. My primary goal was to get the book back into print, not necessarily to rewrite a new book. But I have taken the opportunity to change emphasis on some material as a result of recent developments and to add new material. For example, an entire chapter on thin film magnetic multilayers has been added.

The rationale for a book on magnetism is as valid today as it was when the first two editions were published. Magnetic phenomena continue to be discovered with deep scientific implications and novel applications. Since the Second Edition, for example, Giant Magneto Resistance (GMR) was discovered and the new field of "spintronics" is expanding rapidly. In addition, magnetic properties are often an important clue to our understanding of new materials. High temperature superconductors are a good example. The "parent" (undoped) compounds are antiferromagnetic. Their magnetic properties, studied by susceptibility measurements, nuclear magnetic resonance, neutron scattering, etc., have provided insight to the superconducting state. The purpose of this book is to provide a framework for understanding magnetic phenomena.

This framework is built upon linear response theory. In particular, mean field theory, or the random phase approximation, is used to determine the response of materials to magnetic fields. This approach provides a physical description of most magnetic phenomena. But it is not as powerful and elegant as other approaches applied to many interesting problems represented by magnetic systems. For example, I do not cover the renormalization group or the techniques used to obtain exact solutions to lower dimensional systems. Thus, this book may be thought of as a poor man's theory of magnetic phenomena.

One of the challenges of producing this edition was that the previous editions had been type-set, so the material did not exist in a digital format. This

VIII Preface

meant retyping everything using LaTeX. I am extremely grateful to Ferna Hartman of Carnegie Mellon for recreating not only the text but the many equations as well. The figures also had to be scanned and new ones created. For this I am grateful to Dr. Chando Park. Sergio Rezende and Vladimir Safonov kindly read the entire manuscript and offered helpful suggestions. I want to thank Jeff Lynn for his comments on the chapter on neutron scattering; Luc Berger for reviewing the section on spin transfer; and Christian Ruegg for reading the section on quantum phase transitions. I would also like to thank the Materials Science and Engineering Department at Stanford, Shan Wang in particular, and the Geballe Laboratory for Advanced Materials for their hospitality during the final stage of writing.

Palo Alto, September 2006
R.M. White

10.2 The Magnetic Susceptibility

Any system may be characterized by its response to external stimuli. For example, in electronics the proverbial "black box" is characterized by its measured output voltage when an input current is applied. This transfer impedance, as it is called, provides all the information necessary to understand the operation of the black box. If we know what is in the black box - for example, the detailed arrangement of resistors, diodes, etc. - then we can predict, through analysis, what the transfer impedance will be.

Similarly, a system of charges and currents, such as a crystal, may be characterized by a response function. In this text we shall be concerned mainly with the response of such a system to a magnetic field. In this case the "output" is the magnetization and the response function is the magnetic susceptibility. A complete analysis of the magnetic susceptibility is virtually impossible since the system consists of about 10^{21} particles. Therefore we usually look to a measured susceptibility for clues to the important mechanisms active in the system and then use these to analyze the system. In order to carry out such a program, we must know what possible mechanisms exist and what effect they have on the susceptibility.

Determination of the susceptibility entails evaluation of the magnetization produced by an applied magnetic field. In general, this applied field may depend on space and time. The resulting magnetization will also vary in space and time. If the spatial dependence of the applied field is characterized by a wave vector $\boldsymbol{q}$ and its time dependence is characterized by a frequency ω, and if we restrict ourselves for the time being to the magnetization with this wave vector and frequency, we obtain the susceptibility $\chi(\boldsymbol{q},\omega)$. As we shall see shortly, the magnetization is the average magnetic moment. The magnetic moment itself is a well-defined quantity. The problem, however, is the computation of its average value. In order to compute this average it is necessary to know the probabilities of the system being in its various configurations. This information is contained in the distribution function associated with the system.

We shall see in this chapter that the distribution function depends on the total energy, or Hamiltonian, of the system. Therefore the first step in understanding magnetic properties is the identification of those interactions relevant to magnetism. In Chap. 2 the origin of these interactions is discussed, and they are expressed in a form which facilitates their application in later chapters. The reader is asked to keep in mind that Chaps. 1 and 2 both constitute background material for the theoretical development which begins in Chap. 3. The motivation for the material in these first two chapters should become clear as this theory unfolds.

In the absence of time-dependent fields we may assume that the system is in thermal equilibrium. In this case the distribution function is easily obtained. In Chap. 3 this is used to compute the response of noninteracting moments to a static field. This computation leads to the susceptibility $\chi(\boldsymbol{q},0)$. In Chaps. 4 and 5 the response $\chi(\boldsymbol{q},0)$ of an interacting system of moments to a static field is investigated in the random-phase approximation.

In the presence of time-dependent fields the distribution function must be obtained from its equation of motion. In the case of localized moments this consists of solving the Bloch equations. For itinerant moments the distribution function is obtained from a Boltzmann equation. In Chaps. 6 and 7 these equations are solved for weakly interacting systems to obtain the generalized susceptibility $\chi(\boldsymbol{q},\omega)$. Finally, in Chap. 8 the generalized susceptibility associated with strongly interacting systems is investigated. This function is of particular interest because its singularities determine the magnetic-excitation spectrum of the system.

With the development of thin film deposition techniques it became possible to fabricate inhomogeneous magnetic materials, particularly thin films. Chapter 9 describes some of the

phenomena associated with such structures.

One of the most powerful techniques for studying the spatial and temporal behavior of magnetic materials is neutron scattering. While pulsed and "cold" sources have expanded the range of neutron studies since the first edition of this book, the scattering description provided in Chap. 10 remains valid.

The next few sections introduce the basic quantities with which we shall be concerned throughout this text. Since these quantities may be defined in various ways, the reader may find it informative to compare other approaches (especially the classic work [1]).

10.2.1 1.1 The Magnetic Moment

Let us begin by discussing the magnetic moment. To see why this particular object is of interest let us consider the classical description of a system of charges and currents. Such a system is governed by Maxwell's equations. The appropriate forms of these equations in a medium are the so-called macroscopic Maxwell equations, which are obtained from the microscopic equations by averaging over a large number of particles, see [2]. The microscopic equation
in which we shall be particularly interested is the one representing Ampere's law, which has the differential form

$$\nabla \times \boldsymbol{h} = \frac{4\pi}{c}\boldsymbol{j} + \frac{\partial \boldsymbol{e}}{\partial t} \tag{1.1}$$

We define the average fields

$$\begin{aligned} \langle \boldsymbol{h} \rangle &\equiv \boldsymbol{B}, \\ \langle \boldsymbol{e} \rangle &\equiv \boldsymbol{E}. \end{aligned} \tag{1.2}$$

Here $\langle \ldots \rangle$ is a spatial average over a region which is small compared with the size of the sample, yet large enough to contain many atomic systems (the lower limit to the macroscopic domain would typically be 10 nm).

When we write $\boldsymbol{B}(\boldsymbol{r})$ or $\boldsymbol{E}(\boldsymbol{r})$, the coordinate $\boldsymbol{r}$ refers to the center of the region over which the average is taken. Thus the first equation of (1.2) might have been written as $\boldsymbol{B}(\boldsymbol{r}) = \langle \boldsymbol{h} \rangle_r$. In this description it is assumed that any spatial variations are large in comparison with interatomic spacings. The actual details of the averaging will be discussed in Sect.1.2. With this notation the macroscopic version of (1.1) becomes

$$\nabla \times \boldsymbol{B} = \frac{4\pi}{c}\langle \boldsymbol{j} \rangle + \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t} \tag{1.3}$$

The objective now is to calculate the average current density. To do this we separate the total current density into two parts, that associated with conduction electrons and that localized at an ionic site. The average value of the conduction electron current density is the free current density $\boldsymbol{j}_{\text{free}}$.

The ionic current density may be further separated into two contributions. First of all, the ion may possess an electric-dipole moment which is characterized by a dipole charge density ρ_{dip} . If this charge density is time dependent, there is a polarization current density $\boldsymbol{j}_{\text{pol}}$ which satisfies the continuity equation

$$\nabla \cdot \boldsymbol{j}_{\text{pol}} = -\frac{\partial \rho_{\text{dip}}}{\partial t} \tag{1.4}$$

Taking the average of this equation and assuming that the average commutes with the time and space derivatives, we obtain

$$\left\langle \sum_{\text{ions}} \boldsymbol{j}_{\text{pol}} \right\rangle = \frac{\partial \boldsymbol{P}}{\partial t}, \tag{1.5}$$

where the sum is over those ions within the averaging volume and $\boldsymbol{P}$ is the electric polarization defined by

$$\left\langle \sum_{\text{ions}} \rho_{\text{dip}} \right\rangle = -\nabla \cdot \boldsymbol{P}.$$

The second contribution to the ionic current density arises from the internal motion of the ionic electrons. Since this current density $\boldsymbol{j}_{\text{mag}}$ is stationary, $\nabla \cdot \boldsymbol{j}_{\text{mag}} = 0$. This is the current density responsible for the magnetic moment $\boldsymbol{m}$ of the ion. If the center of mass of the ion is at $\boldsymbol{R}$, the magnetic moment is defined as

$$\boldsymbol{m} = \frac{1}{2c} \int d\boldsymbol{r} (\boldsymbol{r} - \boldsymbol{R}) \times \boldsymbol{j}_{\text{mag}} \tag{1.6}$$

A convenient representation for $\boldsymbol{j}_{\text{mag}}$ which has zero divergence and satisfies (1.6) is

$$\boldsymbol{j}_{\text{mag}} = -c\boldsymbol{m} \times \nabla f(|\boldsymbol{r} - \boldsymbol{R}|) \tag{1.7}$$

where $f(|\boldsymbol{r} - \boldsymbol{R}|)$ a smoothly varying function centered at $\boldsymbol{R}$ which goes to 0 at the ionic radius and is normalized to 1 . In Chap. 2 we shall see that this function has a quantum mechanical interpretation. Then

$$\left\langle \sum_{\text{ions}} \boldsymbol{j}_{\text{mag}} \right\rangle = c \left\langle \sum_{\text{ions}} \nabla f(|\boldsymbol{r} - \boldsymbol{R}|) \times \boldsymbol{m} \right\rangle = c\nabla \times \left\langle \sum_{\text{ions}} f(|\boldsymbol{r} - \boldsymbol{R}|) \boldsymbol{m} \right\rangle. \tag{1.8}$$

The last average in (1.8) is the magnetization $\boldsymbol{M}$, defined by

$$\boldsymbol{M} \equiv \left\langle \sum_{\text{ions}} f(|\boldsymbol{r} - \boldsymbol{R}|) \boldsymbol{m} \right\rangle. \tag{1.9}$$

Combining these results, we may now write (1.3) as

$$\nabla \times \boldsymbol{B} = \frac{4\pi}{c} \boldsymbol{j}_{\text{free}} + \frac{4\pi}{c} \frac{\partial \boldsymbol{P}}{\partial t} + 4\pi \nabla \times \boldsymbol{M} + \frac{1}{c} \frac{\partial \boldsymbol{E}}{\partial t}. \tag{1.10}$$

Defining

$$\boldsymbol{H} = \boldsymbol{B} - 4\pi \boldsymbol{M} \tag{1.11}$$

and

$$\boldsymbol{D} = \boldsymbol{E} + 4\pi \boldsymbol{P} \tag{1.12}$$

we have the familiar result

$$\nabla \times \boldsymbol{H} = \frac{4\pi}{c} \boldsymbol{j}_{\text{free}} + \frac{1}{c} \frac{\partial \boldsymbol{D}}{\partial t} \tag{1.13}$$

Thus we see that the magnetization which appears in the macroscopic Maxwell's equations is the average of the ionic magnetic moment density. Since $f(|\boldsymbol{r} - \boldsymbol{R}|)$ is normalized to the volume the magnetization is the magnetic moment per unit volume.

As an example of the use of definition (1.6), let us neglect the possibility of nuclear currents and consider only the electron currents within the ion. Then

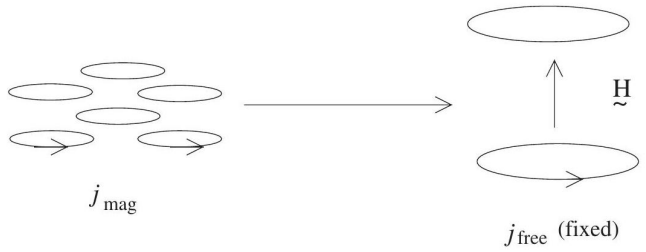


Fig. 1.1. Geometry envisioned in deriving the magnetic energy

$$\boldsymbol{j}_{\text{mag}}(\boldsymbol{r}) = \sum_{\alpha} e\boldsymbol{v}_{\alpha}\delta\left(\boldsymbol{r} - \boldsymbol{r}_{\alpha}\right) \tag{1.14}$$

where e is the charge on the electron, which is $-|e|$, and $\boldsymbol{v}_a$ is the velocity of the α th electron. From (1.6) we find for the total magnetic moment of the ion

$$\boldsymbol{m} = \frac{e}{2c}\sum_{\alpha} \boldsymbol{r}_{\alpha} \times \boldsymbol{v}_{\alpha}. \tag{1.15}$$

Recalling that the orbital angular momentum of an electron is

$$\boldsymbol{l}_{\alpha} = \boldsymbol{r}_{\alpha} \times m\boldsymbol{v}_{\alpha}, \tag{1.16}$$

we have

$$\boldsymbol{m} = \sum_{\alpha} \frac{e}{2mc}\boldsymbol{l}_{\alpha}. \tag{1.17}$$

Since $e = -|e|$, we see that the orbital magnetic moment of an electron is in the opposite direction to its orbital angular momentum.

We shall find it convenient to adopt a more general definition of the magnetic moment than that given by (1.6). This definition is based on the energy of the magnetic system (magnetic energy is discussed in [3] and [4]). The form of the magnetic energy depends upon the definition of the magnetic system. Let us define our magnetic system by the ionic magnetic current density $\boldsymbol{j}_{\text{mag}}$. This excludes the free currents, $\boldsymbol{j}_{\text{free}}$, which are assumed to be fixed and are the source of an external field $\boldsymbol{H}$ in which our magnetic ion is to be located. [1] We now want to know the change in energy of this magnetic system when the field $\boldsymbol{H}$ is applied or, equivalently, we may think of bringing the currents $\boldsymbol{j}_{\text{mag}}$ in from infinity to a position in the field (see Fig.1.1).

The energy difference results from the work done by the magnetic currents as they accommodate to the increasing external field. Since the magnetic field itself does no work on moving charges, we must use the induced electric field

which is present while the external magnetic field is being turned on (due to the relative motion). This is given by

$$\nabla \times \boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{H}}{\partial t}. \tag{1.18}$$

The work done by the magnetic currents in a time δt is

$$\delta W = \int \boldsymbol{j}_{\text{mag}} \cdot \boldsymbol{E} d\boldsymbol{r}\delta t \tag{1.19}$$

[01] The magnetic field $\boldsymbol{H}$ is understood to be the field in vacuum. Strictly speaking, this is the magnetic induction on flux density, $\boldsymbol{B}$. But in vacuum and in cgs units, $\boldsymbol{B} = \boldsymbol{H}$. It has become common practice to denote the field in vacuum as $\boldsymbol{H}$.

Making use of the representation (1.7) for $\boldsymbol{j}_{\text{mag}}$, integrating by parts, and then using (1.18), we obtain

$$\delta W = -\int f(|\boldsymbol{r}-\boldsymbol{R}|)\boldsymbol{m}\cdot\delta\boldsymbol{H}d\boldsymbol{r} \tag{1.20}$$

If the field $\boldsymbol{H}$ is uniform over the ionic dimension, $\delta\boldsymbol{H}$ may be taken outside the integral. Since $\boldsymbol{m}$ is just a constant vector and $f(|\boldsymbol{r}-\boldsymbol{R}|)$ is normalized to unity,

$$\delta W = -\boldsymbol{m}\cdot\delta\boldsymbol{H} \tag{1.21}$$

This work corresponds to Kittel's "scheme A" for applying the field [4]. Kittel also calculates the work needed to create the magnetized material in zero field in the first place ("scheme B"). The work associated with scheme A is important because this is the work that results in the change in the energy of the system given by its quantum mechanical eigenvalues [4].

Neither of these results give the total change in energy of the system when the magnetic material is introduced into the field since they do not include the work done by the source in keeping $\boldsymbol{j}_{\text{free}}$ fixed. Jackson shows that the total change in energy is given by

$$W = \frac{1}{2}\int \boldsymbol{M}\cdot\boldsymbol{H}_0 d\boldsymbol{r}$$

where $\boldsymbol{H}_0$ is the field ($\boldsymbol{B}_0$) in the absence of the magnetic material, and the $\frac{1}{2}$ arises from an assumed linear relation between $\boldsymbol{M}$ and $\boldsymbol{B}$.

The resulting change in the energy of the magnetic system is $\delta E = W$. Thus from (1.21)

$$\boldsymbol{m} = -\frac{\partial E}{\partial \boldsymbol{H}} \tag{1.22}$$

As an example of the application of this definition, consider the ionic system of electrons which gave rise to the current density of (1.14). In the presence of a uniform field $\boldsymbol{H}$, which may be obtained from a vector potential $\boldsymbol{A}$ by $\boldsymbol{H} = \nabla\times\boldsymbol{A}$, the energy of such a system is

$$E = \sum_\alpha \frac{1}{2}m\boldsymbol{v}_\alpha^2 + \sum_\alpha e\phi_\alpha \tag{1.23}$$

where ϕ_α is the ionic potential. We see that the magnetic field $\boldsymbol{H}$ does not appear explicitly in this energy. However, the velocity is, in fact, a function of the field. In general, the task of finding the field dependence of the total energy of the system is a difficult one. However, in this case the actual field dependence is revealed by expressing the energy in terms of the canonical coordinates of the system. The reason for this is that in a slowly varying uniform field the canonical momentum does not change. When expressed in canonical coordinates, the energy is the same as the Hamiltonian function. For this reason the Hamiltonian $\mathcal{H}$ is often used in place of the energy E in definition (1.22).

In Chap. 2 we shall find that the Hamiltonian is

$$\mathcal{H} = \sum_\alpha \frac{1}{2m}\left(\boldsymbol{p}_\alpha - \frac{e}{c}\boldsymbol{A}\right)^2 + \sum_\alpha e\phi_\alpha. \tag{1.24}$$

With the gauge $\boldsymbol{A} = \frac{1}{2}\boldsymbol{H}\times\boldsymbol{r}$ this becomes

$$\mathcal{H} = \sum_a \frac{p_a^2}{2m} - \sum_a \frac{e}{2mc}(\boldsymbol{r}_a\times\boldsymbol{p}_a)\cdot\boldsymbol{H} + \frac{e^2}{8mc^2}\sum_a(\boldsymbol{H}\times\boldsymbol{r}_a)^2 + \sum_a e\phi_a. \tag{1.25}$$

Differentiating with respect to $\boldsymbol{H}$ and using the fact that the canonical momentum is given by $\boldsymbol{p}_\alpha = m\boldsymbol{v}_\alpha + (e/c)\boldsymbol{A}$, we obtain (1.15).

10.2.2 1.2 The Magnetization

The magnetization is obtained by averaging the ionic moments over a region of space which is large enough to give such an average a meaning but smaller than spatial variations in the system. In order to perform this average we must know the ionic-current distributions. In general they are not known. In fact, herein lies the principle difficulty in the theory of magnetism. In any real system the motion of charge in one region is governed by the charge and currents throughout the system. Thus we have a many-body problem. Historically, there have been two ways of describing magnetic systems, that of localized moments and that of itinerant moments. The choice between these two descriptions depends on the nature of the material and in many cases is a difficult one to make.

In certain cases the relevant current distributions are localized within a lattice cell. In such cases the ionic magnetic moment is relatively unambiguous. The interaction with external charge and current distributions is then expressed in terms of this moment. This approach leads to the spin Hamiltonian, which has proved extremely useful. In other cases we begin by assuming that the current distributions are those associated with free electrons. Thus, although these electrons may extend throughout the lattice, the fact that they may be approximated as a "gas" provides a certain simplification. With these two types of current distributions-that corresponding to very
localized electrons and that corresponding to itinerant electrons-we can proceed to determine the average moment. The computation of averages necessarily relies on the techniques of statistical mechanics. Since this is not the place to develop such techniques, they will be introduced in a rather ad hoc manner. A more thorough derivation may be found in texts on statistical physics [5].

If the system possesses translational invariance, then the statistical average over numerous unit cells of the crystal is equivalent to the time average over one cell. This average is determined by the probability that the system will have some particular current distribution. For example, in the case of a magnetic insulator the average over a cell is an average of the magnetic ion. If this ion consists of h electrons, then, classically, the state of the ion is characterized by the $6h$ coordinates and momenta, $(q_1, \ldots, q_{3h}, p_1, \ldots, p_{3h})$. The magnetization is obtained by multiplying the magnetic moment, which is a function of all the coordinates and momenta, by the probability that the system is in the state ($q_1, \ldots, p_1, \ldots$) and then integrating over all the coordinates and momenta. This probability is determined by the ion's environment, which defines a temperature T. For the most part this will be the temperature of the lattice in which the ions are located. Classically, the equilibrium probability function is the Boltmann distribution function $\exp(-\beta\mathcal{H}_{\text{ion}})$, where $\mathcal{H}_{\text{ion}}$, is the Hamiltonian for the ion and $\beta = 1/k_B T$. Therefore the equilibrium magnetization associated with N/V ions per unit volume is

$$\boldsymbol{M} = \frac{N}{V}\langle \boldsymbol{m} \rangle = \frac{N}{V}\frac{\int \ldots \int \boldsymbol{m} \exp(-\beta\mathcal{H}_{\text{ion}})\, dq_1 \ldots dp_1 \ldots}{\int \ldots \int \exp(-\beta\mathcal{H}_{\text{ion}})\, dq_1 \ldots dp_1 \ldots} \tag{1.26}$$

It is interesting that this classical averaging procedure leads to the conclusion that there can be no magnetism in thermodynamic equilibrium. The reason for this is that the integrals over the momenta in (1.26) run from $-\infty$ to $+\infty$. Therefore adding a vector potential may shift the momentum origin, but it will not affect the limits of integration. Since the vector potential always enters the integrand as an addition to the momentum, it may be transformed away. This is readily seen by considering the partition function Z, which is just the integral of $\exp(-\beta\mathcal{H}_{\text{ion}})$ over phase space,

$$Z = \int_{-\infty}^{\infty} dxdydz \int_{-\infty}^{\infty} dp_x dp_y dp_z \exp\left\{-\beta\left[p_x - (e/c)A_x\right]^2/2m + \ldots\right\} \tag{1.27}$$

This function is of importance because the equilibrium thermodynamic properties of the system can be calculated from it. For example, the energy is $E = -\partial \ell n Z/\partial\beta$.

We introduce

$$u = p_x - \frac{e}{c}A_x, \ldots, \tag{1.28}$$

where, in general, $\boldsymbol{A}$ may be a function of $\boldsymbol{r}$. Then

$$Z = V\int_{-\infty}^{\infty} dudvdw \exp\left[-\beta\left(u^2+v^2+w^2\right)/2m\right] \tag{1.29}$$

which is independent of $\boldsymbol{A}$. Therefore the derivative of Z with respect to the field $\boldsymbol{H}$, which can be shown to be proportional to the magnetization, is 0 . This result, known as Miss van Leeuwen's theorem, forces us to consider the discreteness of the eigenvalues of the system and hence its quantum-mechanical nature. This interesting result has the following physical interpretation (see the discussion in [1] Sect. 26). In the presence of a magnetic field the electrons move in circular orbits in the plane perpendicular to the field. Those electrons that complete such orbits contribute a diamagnetic moment. However, those electrons which strike the boundary have their orbits interrupted with the result that they creep around the boundary giving rise to a paramagnetic moment. It turns out that this paramagnetic moment just cancels the diamagnetic moment. Furthermore, it is independent of the size and nature of the boundary.

Quantum mechanically, the magnetic system is described by a Hamiltonian operator $\mathcal{H}$ which has eigenfunctions ψ with eigenvalues E. The total magnetic moment of the system when it is in the state ψ is, according to (1.22),

$$MV = -\frac{\partial E}{\partial \boldsymbol{H}} \tag{1.30}$$

This may be written in a more useful form. First we differentiate the eigenvalue relation

$$(\mathcal{H}-E)\psi = 0 \tag{1.31}$$

with respect to $\boldsymbol{H}$ to obtain

$$\left(\frac{\partial \mathcal{H}}{\partial \boldsymbol{H}} - \frac{\partial E}{\partial \boldsymbol{H}}\right)\psi = -(\mathcal{H}-E)\frac{\partial \psi}{\partial \boldsymbol{H}}. \tag{1.32}$$

Forming the scalar product with ψ and using the fact that $\mathcal{H}$ is a hermitian operator, i.e., $\mathcal{H}_{ij} = \mathcal{H}^*_{ji}$, we find

$$\begin{aligned}\langle\psi|\frac{\partial \mathcal{H}}{\partial \boldsymbol{H}}|\psi\rangle - \frac{\partial E}{\partial \boldsymbol{H}} &= -\langle\psi|(\mathcal{H}-E)\left|\frac{\partial \psi}{\partial \boldsymbol{H}}\right\rangle \\ &= -\left\langle\frac{\partial \psi}{\partial \boldsymbol{H}}\right|(\mathcal{H}-E)|\psi\rangle^* = 0.\end{aligned} \tag{1.33}$$

Therefore

$$\boldsymbol{M}V = -\langle\psi|\frac{\partial \mathcal{H}}{\partial \boldsymbol{H}}|\psi\rangle \tag{1.34}$$

This leads us to define a magnetic-moment operator

$$\mathcal{M} = -\frac{\partial \mathcal{H}}{\partial \boldsymbol{H}} \tag{1.35}$$

Hereafter $\mathcal{M}$ will be understood to be an operator.

Since the derivation above was independent of the detailed form of the Hamiltonian and its eigenfunctions, the result (1.35) is quite general. For example, the magnetic-moment operator for a particle governed by the nonrelativistic Schrödinger Hamiltonian (1.25) is

$$\mathcal{M}_z = \frac{e}{2mc}\left(xp_y - yp_x\right) - \frac{He^2}{4mc^2}\left(x^2 + y^2\right) = \frac{e}{2c}(x\dot{y} - y\dot{x}). \tag{1.36}$$

For a relativistic electron governed by the Dirac equation (which will be discussed briefly in Chap. 2) the magnetic moment becomes

$$\mathcal{M}_z = -\frac{e}{2}\left(\alpha_x y - \alpha_y x\right) \tag{1.37}$$

Since the α 's are 4×4 matrices acting on negative as well as positive energy states, the physical meaning of the operator is not clear. However, if we apply a transformation to this operator which separates the positive and negative energy states, then we find that an intrinsic spin contribution to the magnetic moment emerges automatically.

To find the magnetization we must take the expectation value of the magnetic moment operator,

$$\langle \mathcal{M} \rangle = \int \psi^* \mathcal{M} \psi \prod_i d\boldsymbol{r}_i \tag{1.38}$$

If we knew the wavefunction ψ this would be straightforward. But the fact that we are describing the system at a temperature T implies that the system is in equilibrium with some temperature bath. Let us describe the system in terms of its eigenfunctions, $\varphi_k\left(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N\right)$. The effect of the temperature bath is to cause the system to move through different states k much as a classical system moves through phase space. That is, the wavefunction ψ may be written as a superposition of states,

$$\psi\left(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N, t\right) = \sum_k c_k(t)\varphi_k\left(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N\right) \tag{1.39}$$

The expectation value then becomes

$$\langle \mathcal{M} \rangle = \sum_k \sum_{k'} c_k^*(t) c_{k'}(t) \mathcal{M}_{kk'} \tag{1.40}$$

where

$$\mathcal{M}_{kk'} = \int \varphi_k^* \mathcal{M} \varphi_{k'} \prod_i d\boldsymbol{r}_i \tag{1.41}$$

When we measure the magnetization we actually do a time average,

$$\langle \overline{\mathcal{M}} \rangle = \sum_k \sum_{k'} \overline{c_k^*(t) c_{k'}(t)} \mathcal{M}_{kk'} \tag{1.42}$$

The quantity $\overline{c_k^*(t) c_{k'}(t)}$ is defined as the statistical density matrix, $\rho_{kk'}$. Thus

$$\langle \overline{\mathcal{M}} \rangle = \mathrm{Tr}(\rho \mathcal{M}). \tag{1.43}$$

If the system is isolated from the temperature bath then the c_k 's are independent of time and $\rho_{kk'} = \left|c_k\right|^2 \delta_{kk'}$. The states comprise the microcanonical ensemble. In contact with the temperature bath $\rho_{kk'} = \exp\left(-\beta E_k\right) \delta_{kk'}$ and we speak of this weighted set of states as the canonical ensemble. In the presence of time-dependent fields it is necessary to solve for $\rho_{kk'}$ from its equation of motion. From the Schrödinger equations

$$i\hbar\frac{\partial\phi(t)^*}{\partial t} = -\mathcal{H}\phi(t)^* \tag{1.44}$$

$$i\hbar\frac{\partial\phi(t)}{\partial t} = \mathcal{H}\phi(t) \tag{1.45}$$

we obtain the equations for the expansion coefficients. From our definition of the density matrix, we find that

$$i\hbar\frac{\partial\rho_{k'k}}{\partial t} = -\sum_{k''}\rho_{k'k''}\mathcal{H}_{k''k} + \sum_{k''}\mathcal{H}_{k'k''}\rho_{k''k} \tag{1.46}$$

or

$$i\hbar\frac{\partial\rho}{\partial t} = [\mathcal{H},\rho]. \tag{1.47}$$

This is often a more convenient approach to the density matrix, for, as we shall see below, when perturbation theory applies, (1.47) may be solved iteratively.

Notice that (1.42) gives the average of the magnetic moment over the entire system. If we are interested in the magnetization at point $\boldsymbol{r}, \boldsymbol{M}(\boldsymbol{r})$ this behavior can be projected out by introducing the Dirac delta function,

$$\boldsymbol{\mathcal{M}}(\boldsymbol{r}) = \frac{1}{2}\sum_{\alpha}\left[\boldsymbol{\mu}_\alpha\delta\left(\boldsymbol{r}-\boldsymbol{r}_\alpha\right) + \delta\left(\boldsymbol{r}-\boldsymbol{r}_\alpha\right)\boldsymbol{\mu}_\alpha\right]. \tag{1.48}$$

Since the delta functions have dimensions of a reciprocal volume, $\mathcal{M}(\boldsymbol{r})$ is the magnetic-moment operator per unit volume. Here $\boldsymbol{\mu}_a$ is the magneticmoment operator associated with the α th electron. Notice that since this is a function of $\boldsymbol{r}_\alpha$ and $\boldsymbol{p}_\alpha$, we must form the symmetric product denoted by $\{\cdots\}$. Therefore the magnetization becomes

$$\boldsymbol{M}(\boldsymbol{r}) = \mathrm{Tr}\{\rho\boldsymbol{\mathcal{M}}(\boldsymbol{r})\}. \tag{1.49}$$

10.2.3 1.3 The Generalized Susceptibility

When we speak of a susceptibility, we are usually referring to a medium in which the response is proportional in some sense to the excitation. If the medium is linear, the response is directly proportional to the excitation. If the medium is nonlinear, the proportionality involves higher powers of the excitation. However, if the excitation is very small, the response will be given to a good approximation by the linear susceptibility. Since time- and spacevarying magnetic fields are generally quite small, a linear response theory is usually adequate. Nonlinear effects become important in dealing with hysteresis phenomena or high-power absorption in magnetic materials. For the most part, then, we shall be concerned with a linear response theory. In this section we shall define the wave-vector-dependent frequency-dependent linear susceptibility and investigate some of its properties.

Let us consider the magnetization $\boldsymbol{M}(\boldsymbol{r},t)$ associated with a particular magnetic field $\boldsymbol{H}(\boldsymbol{r},t)$. These quantities are related to their Fourier components by

$$\boldsymbol{M}(\boldsymbol{r},t) = \frac{1}{2\pi V}\sum_{k}\int d\Omega\boldsymbol{M}(\boldsymbol{k},\Omega)e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\Omega t)} \tag{1.50}$$

$$\boldsymbol{H}(\boldsymbol{r},t) = \frac{1}{2\pi V}\sum_{\boldsymbol{q}}\int d\omega\boldsymbol{H}(\boldsymbol{q},\omega)e^{i(\boldsymbol{q}\cdot\boldsymbol{r}-\omega t)} \tag{1.51}$$

To invert these expansions we use the following relations:

$$\int d\boldsymbol{r} e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}} = V\Delta\left(\boldsymbol{k}-\boldsymbol{k}'\right) \tag{1.52}$$

$$\int dt e^{-i(\Omega-\Omega')t} = 2\pi\delta\left(\Omega-\Omega'\right) \tag{1.53}$$

$$\sum_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')} = \frac{V}{(2\pi)^3}\int d\boldsymbol{k} e^{i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')} = V\delta\left(\boldsymbol{r}-\boldsymbol{r}'\right) \tag{1.54}$$

Here $\Delta\left(\boldsymbol{k}-\boldsymbol{k}'\right)$ is the Kronecker delta function and $\delta\left(\boldsymbol{r}-\boldsymbol{r}'\right)$ is the Dirac delta function.

We now define the generalized wave-vector-dependent frequency-dependent susceptibility by

$$M_\nu(\boldsymbol{k},\Omega) = \sum_{\boldsymbol{q}}\int d\omega \sum_{\mu}\chi_{\nu\mu}(\boldsymbol{k},\boldsymbol{q};\Omega,\omega)H_\mu(\boldsymbol{q},\omega) \tag{1.55}$$

where ν and $\mu = x, y,$ or z. This may be written in the more convenient dyadic form

$$\boldsymbol{M}(\boldsymbol{k},\Omega) = \sum_{\boldsymbol{q}}\int d\omega \boldsymbol{\chi}(\boldsymbol{k},\boldsymbol{q};\Omega,\omega)\cdot\boldsymbol{H}(\boldsymbol{q},\omega)$$

In general $\boldsymbol{\chi}(\boldsymbol{k},\boldsymbol{q};\Omega,\omega)$ will depend on the particular form of $\boldsymbol{H}(\boldsymbol{r},t)$, or equivalently, $\boldsymbol{H}(\boldsymbol{q},\omega)$; that is, the susceptibility is a functional of the field. The susceptibility is also a tensor. Furthermore, since the magnetization may be out of phase with the exciting field, the susceptibility is also complex. Substituting this expression into (1.50) gives

$$\boldsymbol{M}(\boldsymbol{r},t) = \frac{1}{2\pi V}\sum_{\boldsymbol{k}}\int d\Omega \sum_{\boldsymbol{q}}\int d\omega\chi(\boldsymbol{k},\boldsymbol{q};\Omega,\omega)\cdot\boldsymbol{H}(\boldsymbol{q},\omega)e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\Omega t)} \tag{1.56}$$

or

$$\begin{aligned}\boldsymbol{M}(\boldsymbol{r},t) = &\iint d\boldsymbol{r}'dt' \\ &\times\left\{\left[\frac{1}{2\pi V}\sum_{\boldsymbol{k}}\int d\Omega\sum_{\boldsymbol{q}}\int d\omega\boldsymbol{\chi}(\boldsymbol{k},\boldsymbol{q};\Omega,)e^{i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')}e^{-i\Omega(t-t')}\right]\right. \\ &\left.\times e^{i(\boldsymbol{k}-\boldsymbol{q})\cdot\boldsymbol{r}'}e^{-i(\Omega-\omega)t'}\right\}\cdot\boldsymbol{H}\left(\boldsymbol{r}',t'\right)\end{aligned} \tag{1.57}$$

where the quantity in braces defines a general spatial-temporal susceptibility density $\boldsymbol{\chi}\left(\boldsymbol{r},\boldsymbol{r}';t,t'\right)$.

If the magnetic medium possesses translational invariance, then this susceptibility must be a function only of the relative coordinate $\boldsymbol{r}-\boldsymbol{r}'$. From the expression above we see that this implies that in the wave-vector-dependent susceptibility $\boldsymbol{q}$ is equal to $\boldsymbol{k}$. Furthermore, if the medium is stationary, it can be shown that the temporal dependence is $t-t'$, which implies a monochromatic response to a monochromatic excitation with the same frequency, that is, $\Omega=\omega$. Therefore, when these conditions are satisfied, the susceptibility takes the form

$$\chi(k,q;\Omega,\omega) = \chi(q,\omega)\Delta(k-q)\delta(\Omega-\omega)$$

Thus

$$\boldsymbol{M}(\boldsymbol{r},t)=\iint d\boldsymbol{r}'dt'\boldsymbol{\chi}\left(\boldsymbol{r}-\boldsymbol{r}',t-t'\right)\cdot\boldsymbol{H}\left(\boldsymbol{r}',t'\right) \tag{1.58}$$

where

$$\chi\left(\boldsymbol{r}-\boldsymbol{r}',t-t'\right)=\frac{1}{2\pi V}\sum_{\boldsymbol{q}}\int d\omega\chi(\boldsymbol{q},\omega)e^{i\boldsymbol{q}\cdot(\boldsymbol{r}-\boldsymbol{r}')}e^{-i\omega(t-t')} \tag{1.59}$$

and its Fourier transform is

$$\chi(\boldsymbol{q},\omega)=\int d\left(t-t'\right)\int d\left(\boldsymbol{r}-\boldsymbol{r}'\right)\chi\left(\boldsymbol{r}-\boldsymbol{r}',t-t'\right)e^{-i\boldsymbol{q}\cdot(\boldsymbol{r}-\boldsymbol{r}')}e^{i\omega(t-t')} \tag{1.60}$$

The more general susceptibility is required whenever the presence of impurities destroys the translational invariance. The response of a typical paramagnet is such a case.

Since the susceptibility has such a general nature, it should not be surprising that there are various important theorems involving this quantity. We shall consider three of these now. The first theorem, known as the Kramers-Kronig relations, relates the real and imaginary parts of the susceptibility. The second is the fluctuation-dissipation theorem, which relates the susceptibility to thermal fluctuations in the magnetization. Finally, we shall present a derivation of the so-called Onsager relation that describes the symmetry of the susceptibility tensor.

1.3.1 The Kramers-Kronig Relations

As a consequence of some rather general properties of

$$\boldsymbol{\chi}(\boldsymbol{q},\omega)=\boldsymbol{\chi}'(\boldsymbol{q},\omega)+i\boldsymbol{\chi}''(\boldsymbol{q},\omega)$$

its real part $\chi'(q,\omega)$ and its imaginary part $\chi''(q,\omega)$ are connected on the real axis ω by integral relations known as the Kramers-Kronig relations, or just as dispersion relations. To the electrical engineer, the real and imaginary parts of the response function are related by the Hilbert transform. Let us consider a medium which is linear and stationary (and translationally invariant, although this is not a necessary condition). Then $\boldsymbol{\chi}(\boldsymbol{q},\omega)$ is related to $\boldsymbol{\chi}\left(\boldsymbol{r}-\boldsymbol{r}',t-t'\right)$ by (1.60). If the system obeys the principle of causality, then $\chi\left(\boldsymbol{r}-\boldsymbol{r}',t-t'\right)=0$ for $t<t'$. Hence the time integral in (1.60) runs only from 0 to ∞; that is,

$$\boldsymbol{\chi}(\boldsymbol{q},\omega)=\int_0^\infty dt\boldsymbol{\chi}(\boldsymbol{q},t)e^{i\omega t} \tag{1.61}$$

Therefore the function $\boldsymbol{\chi}(\boldsymbol{q},\omega)$ is a complex function of ω which has no singularities at the ends of the real axis, provided that

$$\int_0^\infty \boldsymbol{\chi}(\boldsymbol{q},t)dt$$

is finite. This is equivalent to the assumption that the response to a finite excitation is finite. The finite values of $\boldsymbol{\chi}(\boldsymbol{q},\omega)$ at the ends of the real axis may be identified with the real part of the susceptibility $\boldsymbol{\chi}'(\boldsymbol{q},\infty)$. The fact that $\chi''(\boldsymbol{q},\omega)$ vanishes as $\omega\to\infty$ may be obtained from the following physical argument. As we shall see in Chap. 5, the rate of energy absorption by a magnetic system is proportional to $\omega\chi''(\boldsymbol{q},\omega)$. If this is to remain finite as $\omega\to\infty$, then $\chi''(\boldsymbol{q},\omega)$ must go to 0 as $\omega\to\infty$. This result may also be derived from the finite-response assumption.

There is no reason for the real part of the susceptibility to vanish as $\omega\to\infty$. Therefore let us define $\lim_{\omega\to\infty}\chi'(\boldsymbol{q},\omega)=\chi(\boldsymbol{q},\infty)$. The quantity $\boldsymbol{\chi}(\boldsymbol{q},\omega)-\boldsymbol{\chi}(\boldsymbol{q},\infty)$ is then a complex

function which vanishes at the ends of the real axis. The theory of complex variables tells us that the function
$\boldsymbol{\chi}(\boldsymbol{q}, z) - \boldsymbol{\chi}(\boldsymbol{q}, \infty)$, where z is a complex variable, will be analytic in the upper half plane. The residue theorem then says

$$\oint_c \frac{\boldsymbol{\chi}(\boldsymbol{q}, z) - \boldsymbol{\chi}(\boldsymbol{q}, \infty)}{z - \omega} dz = 0 \tag{1.62}$$

where the contour C runs from $-\infty$ to $+\infty$ along the real axis and closes in the upper half plane. In terms of its principal value, this integral may be written as

$$\mathcal{P} \int_{-\infty}^{\infty} \frac{\chi\left(\boldsymbol{q}, \omega'\right) - \chi(\boldsymbol{q}, \infty)}{\omega' - \omega} d\omega' - i\pi[\chi(\boldsymbol{q}, \omega) - \chi(\boldsymbol{q}, \infty)] = 0 \tag{1.63}$$

Equating the real and imaginary parts to 0 separately gives the result

$$\chi'(\boldsymbol{q}, \omega) - \chi(\boldsymbol{q}, \infty) = \frac{1}{\pi} \mathcal{P} \int_{-\infty}^{\infty} \frac{\chi''\left(\boldsymbol{q}, \omega'\right)}{\omega' - \omega} d\omega' \tag{1.64}$$

$$\chi''(\boldsymbol{q}, \omega) = \frac{1}{\pi} \mathcal{P} \int_{-\infty}^{\infty} \frac{\chi'\left(\boldsymbol{q}, \omega'\right) - \chi(\boldsymbol{q}, \infty)}{\omega' - \omega} d\omega' \tag{1.65}$$

The usefulness of this result lies in the fact that χ'' is proportional to the absorption spectrum of the medium. Therefore (1.64) tells us, for example, that the static susceptibility may be obtained by integrating over the absorption spectrum. This is, in fact, an experimental technique used to obtain the static susceptibility of certain systems.

Since the response $M(\boldsymbol{r}, t)$, is a real quantity, it follows that $\chi'(\boldsymbol{q}, \omega)$ is an even function of ω while $\chi''(\boldsymbol{q}, \omega)$ is odd. This enables us to express the Kramers-Kronig relations in terms of integrals over positive frequencies. In particular,

$$\chi''(\boldsymbol{q}, \omega) = -\frac{2\omega}{\pi} \mathcal{P} \int_0^{\infty} \frac{\chi'\left(\boldsymbol{q}, \omega'\right)}{\omega'^2 - \omega^2} d\omega' \tag{1.66}$$

The term involving $\chi(\boldsymbol{q}, \infty)$ vanishes because the principal value of the integral of $1/\left(\omega'^2 - \omega^2\right)$ is zero.

1.3.2 The Fluctuation-Dissipation Theorem

It is well known that a colloidal suspension of particles exhibits Brownian motion; that is, the particles move about irregularly because they are being bombarded by the molecules of the liquid. Now, suppose that these particles are charged, and we attempt to accelerate them with an external electric field. Because of the impacts with the molecules, the particles experience a resistive force which is proportional to their velocity. Thus the mechanism that produces the random fluctuations in the position of the particle is also
responsible for its response to an external excitation. The relationship between the response of a system and its thermal fluctuation spectrum is called the fluctuation-dissipation theorem. This relationship is a very general one, and we shall consider only its specific application to a magnetic medium.

Let us consider a linearly polarized magnetic field of amplitude $H_\mu \cos(\boldsymbol{q} \cdot \boldsymbol{r})$ oscillating at a frequency ω in the μ direction, $H_\mu \cos(\boldsymbol{q} \cdot \boldsymbol{r}) \cos \omega t$. Since we have a linear system, the principle of superposition applies. Therefore we may construct the response to an arbitrary field if we know the response to this particular field. The response in the ν direction to such an excitation is given by (1.55). Since

$$H_\mu(\boldsymbol{q}',\omega') = \frac{\pi H_1 V}{2}\left[\Delta(\boldsymbol{q}'-\boldsymbol{q})\,\delta(\omega'+\omega) + \Delta(\boldsymbol{q}'-\boldsymbol{q})\,\delta(\omega'-\omega) + \Delta(\boldsymbol{q}'+\boldsymbol{q})\,\delta(\omega'+\omega) + \Delta(\boldsymbol{q}'+\boldsymbol{q})\,\delta(\omega'-\omega)\right] \tag{1.67}$$

we obtain

$$\boldsymbol{M}_\nu(\boldsymbol{k},\Omega) = \frac{\pi H_1 V}{2}\left[\chi_{\nu\mu}(\boldsymbol{k},\boldsymbol{q};\Omega,-\omega) + \chi_{\nu\mu}(\boldsymbol{k},\boldsymbol{q};\Omega,\omega) + \chi_{\nu\mu}(\boldsymbol{k},-\boldsymbol{q};\Omega,-\omega) + \chi_{\nu\mu}(\boldsymbol{k},-\boldsymbol{q};\Omega,\omega)\right]. \tag{1.68}$$

Let us now compute $M_\nu(\boldsymbol{k},\Omega)$, using the prescription given in Sect.1.2. The magnetization is

$$\boldsymbol{M}_\nu(\boldsymbol{r},t) = \mathrm{Tr}\left\{\rho\mathcal{M}_\nu(\boldsymbol{r})\right\} \tag{1.69}$$

Although ρ is a function of time, we shall not display this dependence explicitly, for a reason that will be apparent later. Since the time-varying field now disrupts the thermodynamic equilibrium, we must solve for ρ. We write the total Hamiltonian as

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1, \tag{1.70}$$

where $\mathcal{H}_1 = -\int d\boldsymbol{r}\mathcal{M}(\boldsymbol{r})\cdot\boldsymbol{H}(\boldsymbol{r},t)$. For the particular field we are considering this becomes $\mathcal{H}_1 = -\boldsymbol{H}_1\int d\boldsymbol{r}\mathcal{M}_\mu(\boldsymbol{r})\cos(\boldsymbol{q}\cdot\boldsymbol{r})\cos\omega t = -\frac{H_1}{2}\left[\mathcal{M}_\mu(\boldsymbol{q}) + \mathcal{M}_\mu(-\boldsymbol{q})\right]\cos\omega t$.

The equation of motion for the density matrix is

$$\frac{\partial\rho}{\partial t} = \frac{i}{\hbar}\left[\rho,\mathcal{H}_0+\mathcal{H}_1\right]. \tag{1.72}$$

It is now convenient to introduce

$$\rho(t) \equiv \exp\left(\frac{i\mathcal{H}_0 t}{\hbar}\right)\rho\exp\left(\frac{-i\mathcal{H}_0 t}{\hbar}\right) \tag{1.73}$$

Differentiating (1.73) and using (1.72) gives

$$\frac{d\rho(t)}{dt} = \frac{i}{\hbar}\left[\rho(t),\exp\left(\frac{i\mathcal{H}_0 t}{\hbar}\right)\mathcal{H}_1\exp\left(\frac{-i\mathcal{H}_0 t}{h}\right)\right]. \tag{1.74}$$

This has the solution

$$\rho(t) = \rho(-\infty) + \frac{i}{\hbar}\int_{-\infty}^{t}\left[\rho(t'),\exp\left(\frac{i\mathcal{H}_0 t'}{\hbar}\right)\mathcal{H}_1\exp\left(\frac{-i\mathcal{H}_0 t'}{\hbar}\right)\right]dt' \tag{1.75}$$

If the interaction is turned on adiabatically, then $\rho(-\infty) = \rho_0$, which is the equilibrium density matrix $\rho_0 = \exp(-\beta\mathcal{H}_0)/Z$ where $Z = \mathrm{Tr}\left\{\exp(-\beta\mathcal{H}_0)\right\}$. Inverting (1.75), using (1.71), and replacing ρ within the commutator by ρ_0, we have

$$\rho \simeq \rho_0 - i\frac{H_1}{2\hbar}\int_0^\infty\left\{\rho_0,\exp\left(\frac{-i\mathcal{H}_0 t'}{\hbar}\right)\left[\mathcal{M}_\mu(\boldsymbol{q})+\mathcal{M}_\mu(-\boldsymbol{q})\right]\exp\left(\frac{i\mathcal{H}_0 t'}{\hbar}\right)\right\}. \tag{1.76}$$

The magnetization is obtained from (1.69). If the system is ordered in the absence of the applied field, then $\mathrm{Tr}\left\{\rho_0\mathcal{M}_\nu\right\} \equiv M_\nu(-\infty)$ is nonzero. The response of such a system is then defined by the difference $M_\nu(\boldsymbol{r},t) - M_\nu(-\infty)$ resulting from the applied field. In the following we shall understand $M_\nu(\boldsymbol{r},t)$ to be the response to the applied field. Then,

$$M_\nu(\boldsymbol{r},t) = -\,i\frac{H_1}{2\hbar} \times \mathrm{Tr}\left\{\int_0^\infty \left\{\rho_0, \exp\left(\frac{-i\mathcal{H}_0 t'}{\hbar}\right)\left[\mathcal{M}_\mu(\boldsymbol{q}) + \mathcal{M}_\mu(-\boldsymbol{q})\right] \times \exp\left(\frac{i\mathcal{H}_0 t'}{\hbar}\right)\right\}\mathcal{M}_\nu(\boldsymbol{r})\right\}\cos\omega\,(t-t')\,dt' \tag{1.77}$$

Taking the Fourier transform of this equation gives

$$M_\nu(\boldsymbol{k},\Omega) = -\,\frac{\pi H_1}{2\hbar}\,\mathrm{Tr}\left\{\int_0^\infty \left[\rho_0, \mathcal{M}_\mu\left(\boldsymbol{q},-t'\right)\right]\mathcal{M}_\nu(\boldsymbol{k})\right\} e^{i\omega t'} dt' \delta(\Omega+\omega) + (\text{ terms involving } -\boldsymbol{q} \text{ and } -\omega) \tag{1.78}$$

Here $\mathcal{M}_\mu(\boldsymbol{q},t)$ is defined in a manner identical to (1.73). The delta function involving the frequency results from our having linearized the expression for ρ when we replaced ρ by ρ_0 within the commutator.

If we now commute the integral with the trace in (1.78) and make use of the cyclic invariance of the trace, we have

$$\mathrm{Tr}\left\{\int_0^\infty \left[\rho_0, \mathcal{M}_\mu\left(\boldsymbol{q},-t'\right)\right]\mathcal{M}_\nu(\boldsymbol{k})\right\} e^{-i\omega t'} dt' = \int \left\langle\left[\mathcal{M}_\mu\left(\boldsymbol{q},-t'\right), \mathcal{M}_\nu(\boldsymbol{k})\right]\right\rangle e^{i\omega t'} dt' \tag{1.79}$$

By comparing the resulting expression for $M_\nu(\boldsymbol{k},\Omega)$ with (1.68) we make the following identification:

$$\chi_{\nu\mu}(\boldsymbol{k},\boldsymbol{q};\Omega,\omega) = \frac{i}{\hbar V}\int_0^\infty \left\langle\left[\mathcal{M}_\nu(\boldsymbol{k},t), \mathcal{M}_\mu(-\boldsymbol{q})\right]\right\rangle e^{i\omega t} dt \delta(\Omega-\omega) \tag{1.80}$$

Since the $\boldsymbol{q}$ component of the applied field couples to the $-\boldsymbol{q}$ component of the magnetization, let us consider $\chi_{\nu\mu}(\boldsymbol{q},\boldsymbol{q},\omega)$ which we write as $\chi_{\nu\mu}(\boldsymbol{q},\omega)$. Therefore,

$$\chi_{\nu\mu}(\boldsymbol{q},\omega) = \frac{i}{\hbar V}\int_0^\infty \left\langle\left[\mathcal{M}_\nu(\boldsymbol{q},t), \mathcal{M}_\mu(-\boldsymbol{q})\right]\right\rangle e^{i\omega t} dt \tag{1.81}$$

The quantity

$$(i/\hbar)\left\langle\left[\mathcal{M}_\nu(\boldsymbol{q},t), \mathcal{M}_\mu(-\boldsymbol{q})\right]\right\rangle$$

or equivalently

$$\frac{i}{\hbar}\,\mathrm{Tr}\left\{\left[\mathcal{M}_\mu(-\boldsymbol{q}), \rho_0\right]\mathcal{M}_\nu(\boldsymbol{q},t)\right\}$$

is referred to in the literature as the response function of the system. The susceptibility may also be written as an integral over the full range of time by multiplying the integrand by the theta function $\theta(t)$ which equals 1 for $t > 0$ and 0 for $t < 0$. The product of the response function and this theta function is called the double-time-retarded Green's function and represented by double angular brackets:

$$\left\langle\left\langle\mathcal{M}_\nu(\boldsymbol{q},t), \mathcal{M}_\mu(-\boldsymbol{q})\right\rangle\right\rangle \equiv -i\left\langle\left[\mathcal{M}_\nu(\boldsymbol{q},t), \mathcal{M}_\mu(-\boldsymbol{q})\right]\right\rangle\theta(t)$$

These functions are very useful in calculating thermodynamic properties but are beyond the scope of this monograph.

Since the response function does not have a classical analog and is not a well-defined observable, it is more convenient to relate the susceptibility to the correlation function $\langle\{\mathcal{M}_\nu(\boldsymbol{q},t)\mathcal{M}_\mu(-\boldsymbol{q})\}\rangle$, where $\{\cdots\}$ is the symmetrized product, which is defined by

$$\{\mathcal{M}_\nu(\boldsymbol{q},t)\mathcal{M}_\mu(-\boldsymbol{q})\} \equiv \frac{1}{2}\left[\mathcal{M}_\nu(\boldsymbol{q},t)\mathcal{M}_\mu(-\boldsymbol{q}) + \mathcal{M}_\mu(-\boldsymbol{q})\mathcal{M}_\nu(\boldsymbol{q},t)\right]$$

In order to relate the response function to the correlation function let us consider their Fourier transforms,

$$f_{\nu\mu}(\boldsymbol{q},\omega) = \frac{i}{\hbar}\int_{-\infty}^{\infty} \langle[\mathcal{M}_\nu(\boldsymbol{q},t),\mathcal{M}_\mu(-\boldsymbol{q})]\rangle\, e^{i\omega t}dt \tag{1.82}$$

$$g_{\nu\mu}(\boldsymbol{q},\omega) = \int_{-\infty}^{\infty} \langle\{\mathcal{M}_\nu(\boldsymbol{q},t)\mathcal{M}_\mu(-\boldsymbol{q})\}\rangle\, e^{i\omega t}dt \tag{1.83}$$

We can rewrite (1.82) by using the following relation:

$$\begin{aligned}
&\int_{-\infty}^{\infty} dt\,\langle\mathcal{M}_\mu(-\boldsymbol{q})\mathcal{M}_\nu(\boldsymbol{q},t)\rangle\, e^{i\omega t} \\
&= \int_{-\infty}^{\infty} dt\,\mathrm{Tr}\left\{\exp\left(-\beta\mathcal{H}_0\right)\mathcal{M}_\mu(\boldsymbol{q})\exp\left(\frac{i\mathcal{H}_0 t}{\hbar}\right)\mathcal{M}_\nu(\boldsymbol{q})\exp\left(\frac{-i\mathcal{H}_0 t}{\hbar}\right)\right\} e^{i\omega t} \\
&= \int_{-\infty}^{\infty} dt\,\mathrm{Tr}\left\{\exp\left(-\beta\mathcal{H}_0\right)\exp\left(\frac{i\mathcal{H}_0(t-i\hbar\beta)}{\hbar}\right)\mathcal{M}_\nu(\boldsymbol{q})\right. \\
&\quad\left.\times\exp\left[\frac{-i\mathcal{H}_0(t-i\hbar\beta)}{\hbar}\right]\mathcal{M}_\mu(-\boldsymbol{q})\right\} e^{i\omega t} \\
&= e^{-\beta\hbar\omega}\int_{-\infty}^{\infty} dt\,\langle\mathcal{M}_\nu(\boldsymbol{q},t)\mathcal{M}_\mu(-\boldsymbol{q})\rangle\, e^{i\omega t}
\end{aligned} \tag{1.84}$$

Therefore

$$f_{\nu\mu}(\boldsymbol{q},\omega) = \frac{i}{\hbar}\left(1-e^{-\beta\hbar\omega}\right)\int_{-\infty}^{\infty} dt\,\langle\mathcal{M}_\nu(\boldsymbol{q},t)\mathcal{M}_\mu(-\boldsymbol{q})\rangle\, e^{i\omega t} \tag{1.85}$$

From the definition of $g_{\nu\mu}$ we see that its relation to $f_{\nu\mu}$ is

$$g_{\nu\mu}(\boldsymbol{q},\omega) = (\hbar/2i)\coth(\beta\hbar\omega/2) f_{\nu\mu}(\boldsymbol{q},\omega) \tag{1.86}$$

We can also relate $f_{\nu\mu}$ to the susceptibility by separating the time integral as follows:

$$\begin{aligned}
f_{\nu\mu}(\boldsymbol{q},\omega) =& \frac{i}{\hbar}\int_0^{\infty} dt\,\langle[\mathcal{M}_\nu(\boldsymbol{q},t),\mathcal{M}_\mu(-\boldsymbol{q})]\rangle\, e^{i\omega t} \\
&+ \frac{i}{\hbar}\int_{-\infty}^{0} dt\,\langle[\mathcal{M}_\nu(\boldsymbol{q},t),\mathcal{M}_\mu(-\boldsymbol{q})]\rangle\, e^{i\omega t}
\end{aligned} \tag{1.87}$$

If we now make the transformation $t \to -t$ in the second integral and use the fact that $\chi_{\mu\nu}(-\boldsymbol{q},-\omega) = \chi^*_{\mu\nu}(\boldsymbol{q},\omega)$, which follows from (1.81), then

$$f_{\nu\mu}(\boldsymbol{q},\omega) = \left[\chi_{\nu\mu}(\boldsymbol{q},\omega) - \chi^*_{\mu\nu}(\boldsymbol{q},\omega)\right] V$$

and

$$g_{\nu\mu}(\boldsymbol{q},\omega) = (\hbar V/2i)\coth(\beta\hbar\omega/2)\left[\chi_{\nu\mu}(\boldsymbol{q},\omega) - \chi^*_{\mu\nu}(\boldsymbol{q},\omega)\right]$$

Therefore

$$\int_{-\infty}^{\infty} dt\,\langle\{\mathcal{M}_\nu(\boldsymbol{q},t)\mathcal{M}_\mu(-\boldsymbol{q})\}\rangle_s\, e^{i\omega t} = \hbar V\coth(\beta\hbar\omega/2)\chi''_{\nu\mu}(\boldsymbol{q},\omega)_s \tag{1.88}$$

where the subscript s indicates the symmetric part of the tensor.

This is the result we had set out to find. It tells us that the Fourier transform of the correlation function is proportional to the imaginary part of the susceptibility. The derivation has been presented here in detail because it is a very important relationship, and we shall make frequent reference to this result throughout the text.

The imaginary part of the susceptibility describes the absorptive or lossy, response of the magnetic system. The fluctuation-dissipation theorem therefore relates the fluctuations in the magnetization to energy loss. This is a general result that applies to many systems. One example that may be familiar to the reader is Johnson noise. This is the noise associated with thermal fluctuations in a resistive electrical circuit element. The voltage noise power is given by $\langle v^2\rangle/R$. Since this power is due to thermal fluctuations it may also be expressed as the unit of thermal energy, k_BT, divided by the measurement time, which is the reciprocal of the frequency bandwidth, Δf. A rigorous derivation gives

$$\left\langle v^2\right\rangle = 4k_BTR\Delta f$$

This form of the fluctuation-dissipation theorem is known as the Nyquist theorem.

1.3.3 Onsager Relation

Generally, when we probe a magnetic system it is in the presence of a dc field $\boldsymbol{H}$. Therefore $\mathcal{H}_0$ and hence the response function, is a function of this field. In 1931 Onsager pointed out that microscopic reversibility requires the simultaneous reversal of both the magnetic field and time. To see this, consider the response function for the susceptibility:

$$\langle[\mathcal{M}_\nu(\boldsymbol{q},t),\mathcal{M}_\mu(-\boldsymbol{q})]\rangle = \sum_n\langle n|\rho_0\left(\mathcal{H}_0\right)\left[\mathcal{M}_\nu(\boldsymbol{q},t),\mathcal{M}_\mu(-\boldsymbol{q})\right]|n\rangle$$

Let us introduce time reversal through the time reversal operator, T, which satisfies the relation

$$\langle\psi\mid\varphi\rangle = \langle T\psi\mid T\varphi\rangle^* \tag{1.89}$$

Therefore

$$\langle[\mathcal{M}_\nu(\boldsymbol{q},t),\mathcal{M}_\mu(-\boldsymbol{q})]\rangle = \sum_n\langle Tn|T\rho_0\left(\mathcal{H}_0\right)\left[\mathcal{M}_\nu(\boldsymbol{q},t),\mathcal{M}_\mu(-\boldsymbol{q})\right]|n\rangle^* \tag{1.90}$$

We now insert $T^{-1}T$ between all the factors. If the system is invariant under time reversal then $T\rho_0\left(\mathcal{H}_0\right)T^{-1} = \rho_0\left(\mathcal{H}_0\right)$, where any magnetic field has been reversed. Let us display this explicitly as $\rho_0(-H)$. We also have factors like $T\mathcal{M}_\nu(\boldsymbol{q},t)T^{-1}$. Since the magnetic moment changes sign under time reversal,

$$T\mathcal{M}_\nu(\boldsymbol{q},t)T^{-1} = -\mathcal{M}_\nu(\boldsymbol{q},-t) \tag{1.91}$$

If we also recognize that the state $T\mid n>$ is some eigenstate $\mid m>$, then the sum over n is the same as the sum over m, which can then be relabled as n. These steps enable us to carry out all the time reversal operations in the expression for the response function above,

$$\langle[\mathcal{M}_\nu(\boldsymbol{q},t),\mathcal{M}_\mu(-\boldsymbol{q})]\rangle = \sum_n \langle n|\rho_0(-H)\left[\mathcal{M}_\nu(\boldsymbol{q},t),\mathcal{M}_\mu(-\boldsymbol{q})\right]|n\rangle^* \tag{1.92}$$

To remove the complex conjugation we introduce the Hermitian adjoint $0^\dagger$ of an arbitrary operator 0 by the relation

$$\langle\psi|0|\varphi\rangle^* = \langle\varphi|0^\dagger|\psi\rangle. \tag{1.93}$$

The first term in the commutator becomes

$$\begin{aligned}(\rho_0(-H)\mathcal{M}_\nu(\boldsymbol{q},-t)\mathcal{M}_\mu(-\boldsymbol{q}))^\dagger &= \mathcal{M}_\mu(-\boldsymbol{q})^\dagger\mathcal{M}_\nu(\boldsymbol{q},-t)^\dagger\rho_0(-H)^\dagger \\ &= \mathcal{M}_\mu(\boldsymbol{q})\mathcal{M}_\nu(-\boldsymbol{q},-t)\rho_0(-H),\end{aligned} \tag{1.94}$$

where we have used the fact that the magnetic moment and the Hamiltonian are Hermitian operators. We now use the cyclic invariance of the trace to write

$$\begin{aligned}\langle[\mathcal{M}_\nu(\boldsymbol{q},t),\mathcal{M}_\mu(-\boldsymbol{q})]\rangle &= \sum_n \langle n|\rho_0\left(-H\left[\mathcal{M}_\mu(\boldsymbol{q}),\mathcal{M}_\nu(-\boldsymbol{q},-t)\right]\right)|n\rangle \\ &= \sum_n \langle n|\rho_0(-H)\left[\mathcal{M}_\mu(\boldsymbol{q},t),\mathcal{M}_\nu(-\boldsymbol{q})\right]|n\rangle\end{aligned} \tag{1.95}$$

Therefore we have the relation

$$\chi_{\nu\mu}(\boldsymbol{q},\omega,H) = \chi_{\mu\nu}(\boldsymbol{q},\omega,-H) \tag{1.96}$$

which is known as the Onsager relation. Note that this tells us immediately that the diagonal components of the susceptibility tensor must be even functions of the field.

10.2.4 1.4 Second Quantization

Magnetism, particularly in metals, is a many-body phenomenon. It is therefore important to incorporate the fermion statistics of the electrons that govern the magnetic behavior. In this section we shall briefly develop the technique of second quantization which facilitates this description. We shall have occasion to use these results later, particularly in setting up the so-called Anderson and Hubbard Hamiltonians.

Let us begin by considering a system of N interacting particles described by the Hamiltonian

$$\mathcal{H} = \sum_{i=1}^{N} T\left(\boldsymbol{r}_i,\dot{\boldsymbol{r}}_i\right) + \frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{N} V\left(\boldsymbol{r}_i,\boldsymbol{r}_j\right) \tag{1.97}$$

The many-body wave function $\psi\left(\boldsymbol{r}_i,\ldots,\boldsymbol{r}_N,t\right)$ satisfies the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\psi\left(\boldsymbol{r}_1,\ldots\boldsymbol{r}_N,t\right) = \mathcal{H}\psi\left(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N,t\right). \tag{1.98}$$

We now expand this wave function in terms of products of single-particle wave functions characterized by quantum numbers E_i,

$$\psi\left(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N,t\right) = \sum_{(E_1,\ldots,E_N)} c\left(E_1,\ldots,E_N,t\right)\varphi_{E_1}\left(\boldsymbol{r}_1\right)\varphi_{E_2}\left(\boldsymbol{r}_2\right)\ldots\varphi_{E_N}\left(\boldsymbol{r}_N\right) \tag{1.99}$$

where the sum is over all possible sets of quantum numbers. The statistical nature of the particles is contained in the coefficients $c(E_1,\ldots,E_N,t)$. For example, if the particles are bosons, the sign of the coefficient is invariant under particle interchange,

$$c(E_1,\ldots,E_k,\ldots,E_i,\ldots,E_N,t)=c(E_1,\ldots,E_i,\ldots,E_k,\ldots,E_N,t) \tag{1.100}$$

and any number of particles may occupy a given state. If the particles are fermions,

$$c(E_1,\ldots,E_k,\ldots,E_i,\ldots,E_N,t)=-c(E_1,\ldots,E_i,\ldots,E_k,\ldots,E_N,t) \tag{1.101}$$

This insures that there may not be more than one particle in a particular state.

Since we shall be concerned mainly with electrons, which are fermions, we shall be faced with the problem of keeping track of the minus sign introduced when two electrons are interchanged. It is to facilitate this bookkeeping that the concept of second quantization is introduced.

The coefficients in the expansion of ψ above are characterized by the set of N quantum numbers. We could just as well, however, have chosen coefficients characterized by the number of electrons in each of the possible states. That is, instead of the set of N numbers $\{E_1,\ldots,E_N\}$ we could have used the infinite set of numbers $\{n_1,\ldots,n_\infty\}$, where for fermions $n=0$ or 1 . We must be very careful in making this transcription. For example, suppose that the electron at $\boldsymbol{r}_i$ is in a state E_i and the electron $\boldsymbol{r}_k$ is in a state E_k. In this case the occupation-number description would be the same as if the electron at $\boldsymbol{r}_i$, were in state E_k and the electron at $\boldsymbol{r}_k$ were in state E_i. However,

$$c(E_1,\ldots,E_k,\ldots,E_i,\ldots,E_N,t)=-c(E_1,\ldots,E_i,\ldots,E_k,\ldots,E_N,t) \tag{1.102}$$

If we wish to use the occupation-number scheme, we must account for this minus sign. This is done by arbitrarily assigning a certain order to the particular set of quantum numbers $\{E_1,\ldots,E_N\}$. Then the relative sign of any permutation of the electrons from this order is automatically given by writing the single-particle wave functions as a Slater determinant:

$$\begin{aligned} &c(E_1,\ldots,E_N,t)\,\varphi_{E_1}(\boldsymbol{r}_1)\ldots\varphi_{E_N}(\boldsymbol{r}_N)\\ &\quad+(\text{ all permutations within the set }\{E_1,\ldots,E_N\})\\ &=f(n_1,\ldots,n_\infty,t)\begin{vmatrix}\varphi_{E_1}(\boldsymbol{r}_1)\cdots\varphi_{E_1}(\boldsymbol{r}_N)\\ \cdots\cdots\cdots\cdots\cdots\\ \varphi_{E_N}(\boldsymbol{r}_1)\cdots\varphi_{E_N}(\boldsymbol{r}_N)\end{vmatrix} \end{aligned} \tag{1.103}$$

where $f(n_1,\ldots,n_\infty,t)$ has the sign and magnitude of the first c. Summing over all sets $\{E_1,\ldots,E_N\}$ is equivalent to summing over all combinations of occuped states. Therefore

$$\psi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N,t)=\sum_{n_1,\cdots,n_\infty}f(n_1,\ldots,n_\infty,t)\,\frac{1}{\sqrt{N!}}\begin{vmatrix}\varphi_{E_1}(\boldsymbol{r}_1)\cdots\varphi_{E_1}(\boldsymbol{r}_N)\\ \cdots\cdots\cdots\cdots\cdots\\ \varphi_{E_N}(\boldsymbol{r}_1)\cdots\varphi_{E_N}(\boldsymbol{r}_N)\end{vmatrix}. \tag{1.104}$$

The states used in constructing the determinant are, of course, those occupied.

By using this occupation-number description we have succeeded in moving the statistics from the expansion coefficients into the basis functions, which, in fact, form an orthonormal anti-symmetric set.

Let us now define an abstract vector space, or Hilbert space, spanned by vectors $|n_1,n_2,\ldots,n_\infty\rangle$. We introduce operators which satisfy the anticommutation relations

$$\left\{a_i,a_j^\dagger\right\}\equiv a_ia_j^\dagger+a_j^\dagger a_i=\delta_{ij},\quad \{a_i,a_j\}=\left\{a_i^\dagger a_j^\dagger\right\}=0. \tag{1.105}$$

From these relations it can be shown that $a_i^\dagger$ creates an entry in position i (provided one does not already exist there) and a_i destroys an entry at i. Therefore we may represent a basis vector of our Hilbert space as

$$|n_1, \ldots, n_\infty\rangle = \left(a_1^\dagger\right)^{n_1} \left(a_2^\dagger\right)^{n_2} \left(a_\infty^\dagger\right)^{n_\infty} |0\rangle. \tag{1.106}$$

Now consider operating on this with a_k. If $n_k = 0$, then a_k can be commuted all the way over to the "vacuum", where it gives 0 . If $n_k = 1$, then a_k will commute until it comes to $a_k^\dagger$.

$$a_k |n_1, \ldots, n_k, \ldots, n_\infty\rangle = (-1)^{\sum_k \left(a_1^\dagger\right)^n \ldots a_k a_k^\dagger \ldots \left(a_\infty^\dagger\right)^{n_\infty} |0\rangle.} \tag{1.107}$$

Here $\sum_k = n_1 + n_2 + \cdots n_{k-1}$ accounts for all the sign changes that a_k left in its wake as it commuted over to $a_k^\dagger$. We now use $a_k a_k^\dagger = 1 - a_k^\dagger a_k$. In the second term a_k may again commute over to the vacuum to give 0 . Thus we are left with

$$a_k |n_1, \ldots, n_k, \ldots, n_\infty\rangle = \begin{cases} 0 & n_k = 0 \\ (-1)^{\sum_k |n_1, \ldots, n_k - 1, \ldots, n_\infty\rangle} & n_k = 1 \end{cases} \tag{1.108}$$

Similarly,

$$a_k^\dagger |n_1, \ldots, n_k, \ldots, n_\infty\rangle = \begin{cases} (-1)^{\sum_k} |n_1, \ldots, n_k + 1, \ldots, n_\infty\rangle & n_k = 0 \\ 0 & n_k = 1 \end{cases} \tag{1.109}$$

Since it can be shown that $a_k^\dagger a_k$ has the eigenvalue n_k, we can simplify these results by writing

$$\begin{aligned} a_k |n_1, \ldots, n_k, \ldots, n_\infty\rangle &= (-1)^{\sum_k \sqrt{n_k} |n_1, \ldots, n_k - 1, \ldots, n_\infty\rangle}, \\ a_k^\dagger |n_1, \ldots, n_k, \ldots, n_\infty\rangle &= (-1)^{\sum_k \sqrt{n_k + 1}} |n_1, \ldots, n_k + 1, \ldots n_\infty\rangle . \end{aligned} \tag{1.110}$$

Having developed the properties of the Hilbert space, we now use the expansion coefficients $f(n_1, \ldots, n_\infty, t)$ to define the abstract state vector

$$|\psi(t)\rangle = \sum_{n_1, \cdots, n_\infty} f(n_1, \ldots, n_\infty t) |n_1, \ldots, n_\infty\rangle \tag{1.111}$$

The reason for this becomes clear when we consider the equation of motion of this state vector. Taking the time derivative, we have

$$i\hbar \frac{\partial |\psi(t)\rangle}{\partial t} = i\hbar \sum_{n_1, \cdots, n_\infty} \frac{\partial f(n_1, \ldots, n_\infty, t)}{\partial t} |n_1, \ldots, n_\infty\rangle . \tag{1.112}$$

To evaluate this we go back to the real-space Schrödinger equation,

$$\begin{aligned} & i\hbar \frac{\partial |\psi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N, t)\rangle}{\partial t} \\ &= i\hbar \sum_{(n_1, \cdots, n_\infty)} \frac{\partial (n_1, \ldots, n_\infty, t)}{\partial t} \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_{E_1}(\boldsymbol{r}_1) \cdots \varphi_{E_1}(\boldsymbol{r}_N) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \varphi_{E_N}(\boldsymbol{r}_1) \cdots \varphi_{E_N}(\boldsymbol{r}_N) \end{vmatrix}, \\ &= \mathcal{H} \sum_{n_1, \cdots, n_\infty} f(n_1, \ldots, n_\infty, t) \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_{E_1}(\boldsymbol{r}_1) \cdots \varphi_{E_1}(\boldsymbol{r}_N) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \varphi_{E_N}(\boldsymbol{r}_1) \cdots \varphi_{E_N}(\boldsymbol{r}_N) \end{vmatrix} \end{aligned} \tag{1.113}$$

We now multiply through from the left by the conjugate Slater determinant appropriate for some particular set of occupation numbers $\{n_1, \ldots, n_\infty\}$. The left-hand side just gives

$$i\hbar \frac{\partial f\left(n_1, \ldots, n_\infty, t\right)}{\partial t} \tag{1.114}$$

Consider the one-particle terms of the right-hand side of (1.113), which arise from the $T\left(\dot{\boldsymbol{r}}_i\right)$ term in (1.97). We write the Slater determinant as

$$\frac{1}{\sqrt{N!}} \sum_p (-1)^p P \varphi_{E_1}\left(\boldsymbol{r}_1\right) \varphi_{E_2}\left(\boldsymbol{r}_2\right) \cdots \varphi_{E_N}\left(\boldsymbol{r}_N\right) \tag{1.115}$$

where P is an operator which permutes the order of the electrons and p is the number of such permutations. Then the matrix element becomes

$$\begin{aligned} &\frac{1}{N!} \sum_i \sum_{n_1', \cdots, n_\infty'} \sum_{p, p'} (-1)^{p+p'} f\left(n_1', \ldots, n_\infty', t\right) \\ &\times \int P \varphi_{E_1}^*\left(\boldsymbol{r}_1\right) \cdots T\left(\dot{\boldsymbol{r}}_i\right) P' \varphi_{E_1'}\left(\boldsymbol{r}_1\right) \cdots dr_1 \cdots dr_N \end{aligned} \tag{1.116}$$

Since $T\left(\dot{\boldsymbol{r}}_i\right)$ is a one-particle operator, the set of numbers $\{n_1', \ldots, n_\infty'\}$ cannot differ from the particular set $\{n_1, \ldots, n_\infty\}$ by more than two numbers. In particular, let the set $\{n_1', \ldots, n_\infty'\}$ contain the state E_i and the set $\{n_1, \ldots, n_\infty\}$ contain the state E_k. The sums over i, p, and p' give $N!$, leaving us with

$$\sum_{k,l} (-1)^{\sum_k + \sum_\ell \int dr \varphi_{E_k}^*(\boldsymbol{r}) T(\dot{\boldsymbol{r}}) \varphi_{E_l}(\boldsymbol{r}) f(n_1, \ldots, n_k - 1, n_l + 1, \ldots, n_\infty, t)} \tag{1.117}$$

Therefore

$$\begin{aligned} i\hbar \frac{\partial |\psi(t)\rangle}{\partial t} = &\sum_{n_1, \cdots, n_\infty} \sum_{k,l} \langle k|T|l\rangle \\ &\times f\left(n_1, \ldots, n_k = 0, \ldots, n_l = 1, \ldots, n_\infty, t\right) (-1)^{\sum_k + \sum_l} \\ &\times \left|n_1, \ldots, n_k = 1, \ldots, n_l = 0, \ldots, n_\infty\right\rangle + (\text{ interaction terms }) \end{aligned} \tag{1.118}$$

We now recall from above that

$$\begin{aligned} &(-1)^{\sum_k + \sum_l |n_1, \ldots, n_k, \ldots n_l, \ldots, n_\infty\rangle} \\ &= a_k^\dagger a_l \left|n_1, \ldots, n_k - 1, \ldots, n_l + 1, \ldots, n_\infty\right\rangle \end{aligned} \tag{1.119}$$

Substituting this into the equation for $\partial|\psi(t)\rangle/\partial t$, we see that the sum over $\{n_1, \ldots, n_\infty\}$ just gives $|\psi(t)\rangle$. Carrying through the same arguments for the two-particle interaction terms, we find

$$i\hbar \frac{\partial |\psi(t)\rangle}{\partial t} = \mathcal{H}|\psi(t)\rangle, \tag{1.120}$$

where

$$\mathcal{H} = \sum_{k,l} \langle k|T|l\rangle a_k^\dagger a_l + \frac{1}{2} \sum_{k,l,s,t} \langle kl|V|st\rangle a_k^\dagger a_l^\dagger a_t a_s. \tag{1.121}$$

Thus we have the important result that in this occupation-number space the state vector $|\psi(t)\rangle$, as defined above, also satisfies a Schrödinger-like equation, with the Hamiltonian expressed in this second-quantized form. It is easy to show that the matrix elements of such second-quantized operators between occupation-number states are the same as the matrix elements of "first-quantized" operators between the usual states.

Since we shall often have occasion to express an operator in secondquantized form, let us develop a prescription for doing this. For this purpose it is convenient to define what is called the field operator in our Hilbert space,

$$\psi(\boldsymbol{r}) = \sum_k \varphi_k(\boldsymbol{r}) a_k \tag{1.122}$$

Here again the $\varphi_k(\boldsymbol{r})$ are a complete set of single-particle states characterized by the quantum numbers k, and a_k is the fermion operator introduced above. To second quantize a one-particle operator such as $T(\dot{\boldsymbol{r}}_i)$ we write $\boldsymbol{r}_i \to \boldsymbol{r}$, sandwich this operator between $\psi^\dagger(\boldsymbol{r})$ and $\psi(\boldsymbol{r})$, and integrate over all space. For a two-particle operator such as $V(\boldsymbol{r}_i, \boldsymbol{r}_j)$ we let $\boldsymbol{r}_i \to \boldsymbol{r}$ and $\boldsymbol{r}_j \to \boldsymbol{r}'$, sandwich it between $\psi^\dagger(\boldsymbol{r})\psi^\dagger(\boldsymbol{r}')$ and $\psi(\boldsymbol{r}')\psi(\boldsymbol{r})$, and integrate over $d\boldsymbol{r}$ and $d\boldsymbol{r}'$.

1.4.1 Example: The Degenerate-Electron Gas

As an example of the use of this prescription let us second quantize the Hamiltonian for a gas of electrons moving in the field of a uniform positive charge distribution. The total Hamiltonian is

$$\mathcal{H} = \mathcal{H}_{e1-e1} + \mathcal{H}_{e1-n} + \mathcal{H}_{n-n} \tag{1.123}$$

The electron-electron Hamiltonian is

$$\mathcal{H}_{e1-e1} = \sum_i \frac{p_i^2}{2m} + \frac{e^2}{2} \sum_{i \neq j} \frac{\exp(-\mu |\boldsymbol{r}_i - \boldsymbol{r}_j|)}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \tag{1.124}$$

where a screening factor has been inserted for mathematical convenience. The interaction of the electrons with the positive background due to the nuclei is

$$\mathcal{H}_{e1-n} = -e^2 \sum_i \int \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}_i - \boldsymbol{r}'|} \exp(-\mu |\boldsymbol{r}_i - \boldsymbol{r}'|)\, d\boldsymbol{r}' \tag{1.125}$$

where $\rho(\boldsymbol{r})$ is the positive charge density, which for a uniform distribution is

$$\rho(\boldsymbol{r}) = \frac{N}{V} \tag{1.126}$$

Thus $\mathcal{H}_{e1-n}$ becomes

$$\mathcal{H}_{e1-n} = -e^2 \left(\frac{N}{V}\right) \sum_i \int \frac{\exp(-\mu |\boldsymbol{r}_i - \boldsymbol{r}'|)}{|\boldsymbol{r}_i - \boldsymbol{r}'|} d\boldsymbol{r}' \tag{1.127}$$

If μ^{-1} is much smaller than L, where L is the sample dimension, we may replace the integral over $d\boldsymbol{r}'$ by one over $d(\boldsymbol{r}' - \boldsymbol{r}_i)$, which gives

$$\mathcal{H}_{e1-n} = -e^2 \frac{N^2}{V} \frac{4\pi}{\mu^2} \tag{1.128}$$

Finally, the self-energy of the background charge is

$$\mathcal{H}_{n-n} = -\frac{1}{2} e^2 \int \frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} e^{-\mu|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r} d\boldsymbol{r}' = \frac{e^2}{2} \frac{N^2}{2} \frac{4\pi}{\mu^2} \tag{1.129}$$

We must now decide what functions to use as a basis for our field operator. Since the eigenfunctions for a gas of free electrons are plane waves, we shall use these as our basis. These states are characterized by their wave vector $\boldsymbol{k}$ and spin quantum number σ. Thus

$$\psi(\boldsymbol{r}) = \sum_{\boldsymbol{k},\sigma} \frac{1}{\sqrt{V}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \eta_\sigma a_{\boldsymbol{k}\sigma} \tag{1.130}$$

where

$$\eta_\uparrow = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \eta_\downarrow = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{1.131}$$

Notice that

$$\int \psi^\dagger(\boldsymbol{r})\psi(\boldsymbol{r}) d\boldsymbol{r} = \sum_{\boldsymbol{k}\sigma} a^\dagger_{\boldsymbol{k}\sigma} a_{\boldsymbol{k}_\sigma} = N \tag{1.132}$$

Since the terms $\mathcal{H}_{e1-n}$ and $\mathcal{H}_{n-n}$ do not involve any electron coordinates, they are carried over directly into our Hilbert space. The kinetic energy of the electron-electron Hamiltonian becomes

$$\sum_{\boldsymbol{k}\sigma} \frac{\hbar^2 k^2}{2m} a^\dagger_{\boldsymbol{k}\sigma} a_{\boldsymbol{k}\sigma} \tag{1.133}$$

For the Coulomb interaction we have

$$\begin{aligned} &\frac{e^2}{2}\frac{1}{V^2} \sum_{\boldsymbol{k}\sigma} \sum_{\boldsymbol{k}'\sigma'} \sum_{\boldsymbol{k}'',\sigma''} \sum_{\boldsymbol{k}'''\cdot\sigma'''} \int d\boldsymbol{r} \int d\boldsymbol{r}' e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}+\boldsymbol{k}'\cdot\boldsymbol{r}')} \frac{e^{-\mu|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|} e^{i(\boldsymbol{k}'''\cdot\boldsymbol{r}'+\boldsymbol{k}''\cdot\boldsymbol{r})} \\ &\quad \times \eta^\dagger_\sigma(\boldsymbol{r}) \eta^\dagger_{\sigma'}(\boldsymbol{r}') \eta_{\sigma'''}(\boldsymbol{r}') \eta_{\sigma''}(\boldsymbol{r}) a^\dagger_{\boldsymbol{k}\sigma} a^\dagger_{\boldsymbol{k}'\sigma'} a_{\boldsymbol{k}'''\sigma'''} a_{\boldsymbol{k}''\sigma''} \\ &= \frac{e^2}{2}\frac{1}{V^2} \sum_{\boldsymbol{k},\sigma} \sum_{\boldsymbol{k}',\sigma'} \sum_{\boldsymbol{k}'',\sigma''} \sum_{\boldsymbol{k}''',\sigma'''} \delta_{\sigma,\sigma''}\delta_{\sigma',\sigma'''} \int d\boldsymbol{r} \int d\boldsymbol{r}' e^{-i(\boldsymbol{k}+\boldsymbol{k}'-\boldsymbol{k}''-\boldsymbol{k}''')\cdot\boldsymbol{r}} \\ &\quad \times \frac{e^{i(\boldsymbol{k}'-\boldsymbol{k}''')\cdot(\boldsymbol{r}-\boldsymbol{r}')} e^{-\mu|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|} a^\dagger_{\boldsymbol{k}\sigma} a^\dagger_{\boldsymbol{k}'\sigma'} a_{\boldsymbol{k}'''\sigma'''} a_{\boldsymbol{k}''\sigma''} \end{aligned} \tag{1.134}$$

If we again treat $\boldsymbol{r}$ and $\boldsymbol{r}-\boldsymbol{r}'$ as independent variables, the integrations may be carried out separately to give

$$\begin{aligned} &\frac{e^2}{2}\frac{1}{V} \sum_{\boldsymbol{k},\sigma} \sum_{\boldsymbol{k}',\sigma'} \sum_{\boldsymbol{k}'',\sigma''} \sum_{\boldsymbol{k}''',\sigma'''} \delta_{\sigma,\sigma'''}\delta_{\sigma',\sigma'''} \Delta\left(\boldsymbol{k}+\boldsymbol{k}'-\boldsymbol{k}''-\boldsymbol{k}'''\right) \\ &\quad \times \frac{4\pi}{\left(\boldsymbol{k}'-\boldsymbol{k}'''\right)^2+\mu^2} a^\dagger_{\boldsymbol{k}\sigma} a^\dagger_{\boldsymbol{k}'\sigma'} a_{\boldsymbol{k}'''\sigma'''} a_{\boldsymbol{k}''\sigma''} \end{aligned} \tag{1.135}$$

If we define $\boldsymbol{k}'-\boldsymbol{k}''' \equiv \boldsymbol{q}$ and collect all the terms, we obtain the total secondquantized Hamiltonian,

$$\begin{aligned} \mathcal{H} = &\sum_{\boldsymbol{k}\sigma} \frac{\hbar^2 k^2}{2m} a^\dagger_{\boldsymbol{k}\sigma} a_{\boldsymbol{k}\sigma} + \frac{e^2}{2V} \sum_{\boldsymbol{k}} \sum_{\boldsymbol{k}'} \sum_{\boldsymbol{q}} \sum_{\sigma,\sigma'} \frac{4\pi}{\boldsymbol{q}^2+\mu^2} \\ &\times a^\dagger_{\boldsymbol{k}-\boldsymbol{q},\sigma} a^\dagger_{\boldsymbol{k}'+\boldsymbol{q},\sigma'} a_{\boldsymbol{k}',\sigma'} a_{\boldsymbol{k},\sigma} - \frac{e^2 N^2}{2V}\frac{4\pi}{\mu^2} \end{aligned} \tag{1.136}$$

It is convenient to introduce an equivalent electron radius $\boldsymbol{r}_0$ by

$$\frac{4}{3}\pi r_0^3 = \frac{V}{N} \tag{1.137}$$

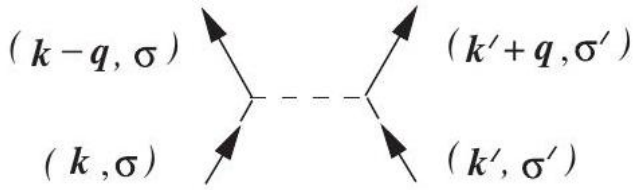
This is made dimensionless by dividing it by the Bohr radius $a_0 = \hbar^2/me^2$. We also define $r_0/a_0 = r_s$ and $V/r_0^3 \equiv \Omega$. Since 1 Rydberg $= me^4/\hbar^2 = e^2/2a_0$, the Hamiltonian expressed in Rydbergs is

$$\begin{aligned}\mathcal{H} =& \frac{1}{r_s^2}\sum_{\boldsymbol{k},\sigma}\left(r_0\boldsymbol{k}\right)^2 a^\dagger_{\boldsymbol{k}\sigma}a_{\boldsymbol{k}_\sigma} + \frac{1}{r_s\Omega}\sum_{\boldsymbol{k},\sigma}\sum_{\boldsymbol{k}',\sigma'}\sum_{\boldsymbol{q}}\frac{4\pi}{r_0^2\left(\boldsymbol{q}^2+\mu^2\right)} \\ &\times a^\dagger_{\boldsymbol{k}-\boldsymbol{q},\sigma}a^\dagger_{\boldsymbol{k}'+\boldsymbol{q},\sigma'}a_{\boldsymbol{k}',\sigma'}a_{\boldsymbol{k},\sigma} - \frac{N^2}{r_s\Omega}\frac{4\pi}{\left(r_0\mu\right)^2}\end{aligned} \tag{1.138}$$

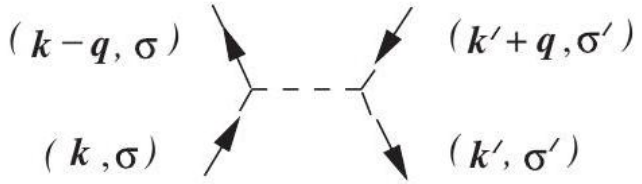
Notice that the electron-electron interaction for $\boldsymbol{q} = 0$ gives a term proportional to N^2, which cancels the contribution from the positive background,
plus a term proportional to N, which vanishes by virtue of the condition $\mu^{-1} \ll L$. Thus we finally have

$$\mathcal{H} = \sum_{\boldsymbol{k},\sigma}\varepsilon_{\boldsymbol{k}}a^\dagger_{\boldsymbol{k},\sigma}a_{\boldsymbol{k},\sigma} + \sum_{\boldsymbol{k},\sigma}\sum_{\boldsymbol{k}',\sigma'}\sum_{\boldsymbol{q}\neq 0}V(\boldsymbol{q})a^\dagger_{\boldsymbol{k}-\boldsymbol{q},\sigma}a^\dagger_{\boldsymbol{k}'+\boldsymbol{q},\sigma'}a_{\boldsymbol{k}',\sigma'}a_{\boldsymbol{k},\sigma}, \tag{1.139}$$

where $\epsilon_k \equiv \left(r_0k/r_s\right)^2$ and $V(q) \equiv 4\pi/\left[r_s\Omega\left(r_0q\right)^2\right]$.
Aside from its advantages for handling particle statistics, the secondquantization formalism also lends itself readily to graphical interpretation. For example, the interaction term above corresponds to the destruction of two particles in states ($\boldsymbol{k},\sigma$) and ($\boldsymbol{k}',\sigma$) and the creation of two particles in states ($\boldsymbol{k}'+\boldsymbol{q},\sigma'$) and ($\boldsymbol{k}-\boldsymbol{q},\sigma$). This may be represented as



In general, we shall also have particle-hole interactions which have the form



We shall see that long-range magnetic order may be characterized as a coherent electron-hole state just as superconductivity is characterized as a coherent electron-electron state.

1.4.2 Example: The Zeeman Interaction

Finally, let us apply this second quantization prescription to the interaction of an electron spin with a magnetic field $H\cos(\boldsymbol{q}\cdot\boldsymbol{r})$. In the next chapter we shall show that this interaction has the form

$$\mathcal{H} = \mu_B\sigma_z H\cos(\boldsymbol{q}\cdot\boldsymbol{r}),$$

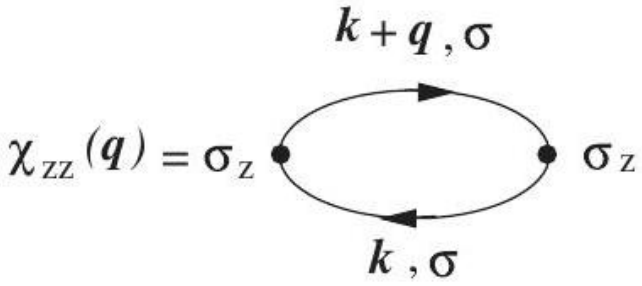
where μ_B is the Bohr magneton and σ_z the Pauli matrix

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

In terms of the field operators (1.122) this becomes

$$\mathcal{H} = \frac{1}{2}\mu_B H \sum_k \left(a^\dagger_{k+q,\uparrow} a_{k\uparrow} - a^\dagger_{k+q,\downarrow} a_{k\downarrow} + a^\dagger_{k-q,\uparrow} a_{k\uparrow} + a^\dagger_{k-q,\downarrow} a_{k\downarrow} \right).$$

Since $\mathcal{H}_0 = 0$ in this simple example, $\sigma_z(t) = \sigma_z$ and the calculation of the longitudinal susceptibility using (1.81) simply involves terms of the form $\left\langle a^\dagger_{k\uparrow} a_{k+q\uparrow} a^\dagger_{k+q,\uparrow} a_{k\uparrow} \right\rangle$. This has the diagammatic representation



In more complex systems it is convenient to characterize various approximations in terms of their diagrammatic representations.

Problems

1.1. The current density j associated with a wave function ψ is given by

$$j = eRe\left(\psi^* \frac{\hbar}{im} \nabla \psi \right).$$

Assuming a hydrogenic wave function of the form

$$\psi(r - R) = R_{n\ell}(r - R) Y^m_\ell(\theta, \varphi)$$

calculate the magnetic moment from (1.6).

1.2. Consider a two-dimensional harmonic oscillator with the Hamiltonian

$$\mathcal{H} = \frac{1}{2m}\left(p_x^2 + m^2\omega^2 x^2\right) + \frac{1}{2m}\left(p_y^2 + m^2\omega^2 y^2\right).$$

Introduce the operators

$$a_\mu = \left(\frac{m\omega}{2\hbar}\right)^{\frac{1}{2}} \mu + i(2m\hbar\omega)^{-\frac{1}{2}} p_\mu$$
$$a^\dagger_\mu = \left(\frac{m\omega}{2\hbar}\right)^{\frac{1}{2}} \mu - i(2m\hbar\omega)^{-\frac{1}{2}} p_\mu, \quad (\mu = x, y)$$

(a) Compute the commutation relation $\left[a_\mu, a^\dagger_\mu\right]$ recalling $[\mu, p_\mu] = i\hbar$.

(b) Express the Hamiltonian in terms of the operators $a_\mu, a^\dagger_\mu$

(c) Let $|n_\mu\rangle$ be an eigenvector of $a^\dagger_\mu a_\mu$ with eigenvalue n_μ. The spectrum of these eigenvalues consists of the set of non-negative integers. Sketch the lowest lying eigenvalues of the oscillator indicating any degeneracies.

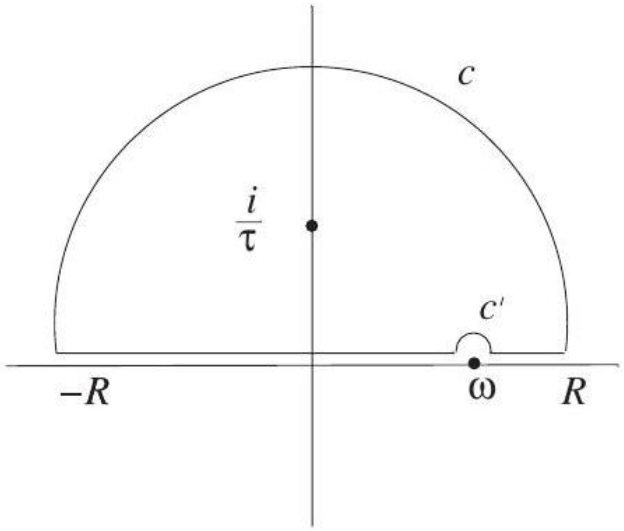
(d) Treat the second term in (1.25) as a perturbation and calculate its affect on the first excited state.

1.3. If $\chi'(\omega) = \frac{1}{1+\omega^2\tau^2}$ find $\chi''(\omega)$ from the relation (1.65).

Hint: The principal value is given by

$$\lim_{R\to\infty}\left[\int_{-R}^{\omega-\epsilon}\frac{\chi'}{\omega'-\omega}d\omega'+\int_{\omega+\epsilon}^{R}\frac{\chi'}{\omega'-\omega}d\omega'\right]$$
$$=\oint\frac{\chi'}{\omega'-\omega}d\omega'-\int_{c}\frac{\chi'}{\omega'-\omega}d\omega'-\int_{c'}\frac{\chi'}{\omega'-\omega}d\omega'$$

where the contour of integration is shown in the figure on page 31. Show that the form (1.66) gives the same answer.



1.4. Consider a gas of electrons. The particle density is

$$\rho(\boldsymbol{r})=\sum_i\delta\left(\boldsymbol{r}-\boldsymbol{r}_i\right)$$

The dynamic form factor is defined as

$$S(\boldsymbol{q},\omega)=\sum_n\left|\left\langle n\right|\rho_{\boldsymbol{q}}^{\dagger}\left|0\right\rangle\right|^2\delta\left(\omega-\omega_{n0}\right)$$

where the states $|n\rangle$ are the eigenstates of the electron system. We shall show later that the scattering of electrons or neutrons from a metal is proportional to this form factor.
(a) Second quantize

$$\rho_{\boldsymbol{q}}^{\dagger}=\int d^3r\rho(\boldsymbol{r})e^{-i\boldsymbol{q}\cdot\boldsymbol{r}}$$

in terms of normalized free electron plane wave states.
(b) Evaluate the form factor $S_0(\boldsymbol{q},\omega)$ for a free electron system characterized by a Fermi sphere of radius k_F.

10.3 The Magnetic Hamiltonian

Most of the magnetic properties that we shall consider arise from electrons. In this chapter we shall develop the Hamiltonian which pertains to the magnetic behavior of a system of electrons. It has been found experimentally that the electron possesses an intrinsic magnetic moment, or spin. The existence of such a moment is a consequence of relativistic considerations. Therefore it is essential that we look for a relativistic description of the motion of an electron. This is given by the Dirac wave equation. We shall limit our discussion of the Dirac equation to the origin of the spin and the form of the spin-orbit interaction (for a more thorough treatment see [6]).

10.3.1 2.1 The Dirac Equation

The objective in developing a relativistic quantum theory of the electron is to ensure that space coordinates and time enter the theory symmetrically. There are several ways of doing this. One way is to start with the general wave equation

$$i\hbar\frac{\partial\psi(\boldsymbol{r},t)}{\partial t}=\mathcal{H}\psi(\boldsymbol{r},t) \tag{2.1}$$

Since the first derivative with respect to time enters on the left, the Hamiltonian must contain a linear space derivative, that is, the Hamiltonian must be linear in the momentum, $\boldsymbol{p}=-i\hbar\boldsymbol{\nabla}$. Thus we assume that the Hamiltonian has the form

$$\mathcal{H}=c\boldsymbol{\alpha}\cdot\boldsymbol{p}+\beta mc^2 \tag{2.2}$$

where $\boldsymbol{\alpha}$ and β are arbitrary coefficients. By imposing certain requirements on the solutions of (2.1), such as that it gives the correct energy-momentum relation $E^2=p^2c^2+m^2c^4$ we obtain conditions on the quantities $\boldsymbol{\alpha}$ and β. These conditions may be satisfied by the 4×4 representations

$$\beta=\begin{bmatrix}\mathbf{1}&0\\0&-\mathbf{1}\end{bmatrix}\quad\text{and}\quad\boldsymbol{\alpha}=\begin{bmatrix}0&\boldsymbol{\sigma}\\\boldsymbol{\sigma}&0\end{bmatrix} \tag{2.3}$$

where the σ_i are the Pauli matrices

$$\sigma_x=\begin{bmatrix}0&1\\1&0\end{bmatrix},\quad\sigma_y=\begin{bmatrix}0&-i\\i&0\end{bmatrix},\quad\sigma_z=\begin{bmatrix}1&0\\0&-1\end{bmatrix},\quad\mathbf{1}=\begin{bmatrix}1&0\\0&1\end{bmatrix}. \tag{2.4}$$

Thus the wave function ψ must be a four-component object. Two of the components correspond to positive-energy solutions and the other two correspond to negative-energy solutions. Holes in the negative-energy spectrum correspond to positrons and require energies of the order of mc^2 for their production.

From a Lagrangian formulation we find that the effect of an external electromagnetic field described by the vector potential $\boldsymbol{A}$ and the scalar potential ϕ may be included by making the substitutions $\boldsymbol{p}\to\boldsymbol{p}-(e/c)\boldsymbol{A}$ and adding $e\phi$ to the Hamiltonian. Thus the Dirac equation becomes

$$i\hbar\frac{\partial\psi}{\partial t}=\left[c\boldsymbol{\alpha}\cdot\left(\boldsymbol{p}-\frac{e}{c}\boldsymbol{A}\right)+\beta mc^2+e\phi\right]\psi. \tag{2.5}$$

Since the energies encountered in magnetic phenomena are much smaller than mc^2, it is convenient to decouple the positive- and negative-energy solutions. This is accomplished by a canonical transformation due to Foldy and Wouthuysen [7]. The resulting Hamiltonian associated with the positive-energy solutions has the form

$$\begin{aligned}\mathcal{H}=&\left[mc^2+\frac{1}{2m}\left(\boldsymbol{p}-\frac{e}{c}\boldsymbol{A}\right)^2-\frac{p^4}{8m^3c^2}\right]+e\phi-\frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot\boldsymbol{H}\\&-i\frac{e\hbar^2}{8m^2c^2}\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\times\boldsymbol{E}-\frac{e\hbar}{4m^2c^2}\boldsymbol{\sigma}\cdot\boldsymbol{E}\times\boldsymbol{p}-\frac{e\hbar^2}{8m^2c^2}\boldsymbol{\nabla}\cdot\boldsymbol{E}.\end{aligned} \tag{2.6}$$

The interesting terms in this Hamiltonian are the last four. The term

$$-(e\hbar/2mc)\boldsymbol{\sigma}\cdot\boldsymbol{H}$$

corresponds to the interaction of the intrinsic spin of the electron with the external field $\boldsymbol{H}$. The next two terms are spin-orbit terms. In a stationary vector potential $\boldsymbol{\nabla} \times \boldsymbol{E} = 0$. And, if the scalar potential, $V(r)$, is spherically symmetric,

$$\boldsymbol{\sigma} \cdot \boldsymbol{E} \times \boldsymbol{p} = -\frac{1}{r}\frac{\partial V}{\partial r}\boldsymbol{\sigma} \cdot \boldsymbol{r} \times \boldsymbol{p} = -\frac{\hbar}{r}\frac{\partial V}{\partial r}\boldsymbol{\sigma} \cdot \boldsymbol{l},$$

where $\hbar\boldsymbol{l} = \boldsymbol{r} \times \boldsymbol{p}$. Thus, the spin-orbit terms reduce to

$$\frac{e\hbar^2}{4m^2c^2}\frac{1}{r}\frac{\partial V}{\partial r}\boldsymbol{\sigma} \cdot \boldsymbol{l}. \tag{2.7}$$

This is what would be expected for an electron spin interacting with the field produced by its orbital motion, except that it is reduced by a factor
of $\frac{1}{2}$ due to relativistic kinematics, also known as the Thomas precession. The last term in (2.6), the so-called Darwin term, represents a correction to the Coulomb interaction due to fluctuations (zitterbewegung) in the electron position arising from the presence of the negative-energy component in the wave function.

The term $p^4/8m^3c^2$ in (2.6) is very small, and along with the Darwin term, it may be neglected for our purposes. If we define the zero of energy as the rest-mass energy, the Hamiltonian which governs the magnetic behavior of an electron is

$$\mathcal{H} = \frac{1}{2m}\left(\boldsymbol{p} - \frac{e}{c}\boldsymbol{A}\right)^2 + e\phi - \frac{e\hbar}{2mc}\boldsymbol{\sigma} \cdot \boldsymbol{H} + \zeta\boldsymbol{l} \cdot \boldsymbol{\sigma}, \tag{2.8}$$

where we have introduced the spin-orbit parameter

$$\zeta = \frac{e\hbar^2}{4m^2c^2}\frac{1}{r}\frac{\partial V}{\partial r}$$

The wave functions associated with (2.8) have two components. These two component functions transform differently under rotations in 3-dimensional space, as we shall see below, than vectors. They are called spinors.

10.3.2 2.2 Sources of Fields

In developing the general Hamiltonian for a single electron we found that the interaction of an electron with its environment is described by the scalar potential ϕ and the vector potential $\boldsymbol{A}$. Both these potentials are functions of the position of the electron under consideration as well as of the coordinates and momenta of any other particles in the system, that is,

$$\phi\left(\boldsymbol{r}; \boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{p}_1, \boldsymbol{p}_2, \ldots\right) \quad \text{and} \quad \boldsymbol{A}\left(\boldsymbol{r}; \boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{p}_1, \boldsymbol{p}_2, \ldots\right).$$

In this section we shall investigate the form these potentials take in a crystalline solid. Our objective is to catalog all the interactions that enter into the magnetic properties of solids so that we shall be free to draw on these results later.

2.2.1 Uniform External Field

The simplest potentials are those arising from uniform external fields. For an electric field $\boldsymbol{E}$, uniform over all space, the interaction $e\phi$ becomes $-e\boldsymbol{r} \cdot \boldsymbol{E}$, where $e\boldsymbol{r}$ is the electric-dipole-moment operator.

For a uniform magnetic field $\boldsymbol{H}$ the magnetic vector potential is not uniquely defined. However, it is convenient to take $\boldsymbol{A} = \frac{1}{2}\boldsymbol{H} \times \boldsymbol{r}$. In this gauge $\nabla \cdot \boldsymbol{A} = 0$. Thus $(\boldsymbol{p} - e\boldsymbol{A}/c)^2/2m$ becomes

$$\frac{p^2}{2m} - \frac{e}{2mc}(\boldsymbol{r} \times \boldsymbol{p}) \cdot \boldsymbol{H} + \frac{e^2}{8mc^2}(\boldsymbol{H} \times \boldsymbol{r})^2. \tag{2.9}$$

The first term is the kinetic-energy term. The second term is a paramagnetic term. Since $\boldsymbol{r} \times \boldsymbol{p} = \hbar \boldsymbol{l}$ is related to the electron's orbital moment $\boldsymbol{\mu}_l$, by

$$\boldsymbol{\mu}_l = \frac{e}{2mc}\boldsymbol{r} \times \boldsymbol{p} = -\frac{|e|\hbar}{2mc}\boldsymbol{l} = -\mu_B \boldsymbol{l}, \tag{2.10}$$

where

$$\mu_B \equiv \frac{|e|\hbar}{2mc}$$

is the Bohr magneton, this second term may be written as $\mu_B \boldsymbol{l} \cdot \boldsymbol{H}$. The third term is a diamagnetic term. When $\boldsymbol{H}$ is in the z direction, that is, $\boldsymbol{H} = H\hat{\boldsymbol{z}}$, then this term reduces to $(e^2H^2/8mc^2)\,(x^2 + y^2)$. This gives, for the total Hamiltonian of an electron in a uniform magnetic field,

$$\mathcal{H} = \frac{p^2}{2m} + \mu_B(\boldsymbol{l} + \boldsymbol{\sigma}) \cdot \boldsymbol{H} + \frac{e^2H^2}{8mc^2} \cdot \left(x^2 + y^2\right) + \zeta \boldsymbol{l} \cdot \boldsymbol{\sigma}. \tag{2.11}$$

2.2.2 The Electric Quadrupole Field

Let us now look into the potentials the electron sees as it moves around or past a nucleus. If we assume that the electron remains outside the nuclear charge and current distributions, we may expand $|\boldsymbol{r} - \boldsymbol{r}_n|^{-1}$ in spherical harmonics, which results in a multipole expansion.

Let us first consider the charge distribution. If $\rho\,(\boldsymbol{r}_n)$ is the charge density at a point $\boldsymbol{r}_n$ inside the nucleus, the electrostatic potential becomes

$$\begin{aligned}\phi(\boldsymbol{r}) &= \int d\boldsymbol{r}_n \frac{\rho\,(\boldsymbol{r}_n)}{|\boldsymbol{r} - \boldsymbol{r}_n|} \\ &= 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \frac{Y_l^m(\theta, \varphi)}{(2l+1)r^{l+1}} \int d\boldsymbol{r}_n \rho\,(\boldsymbol{r}_n)\, Y_l^{m^*}\,(\theta_n, \varphi_n)\, r_n^l\end{aligned} \tag{2.12}$$

where $Y_l^m(\theta, \varphi)$ is the sperical harmonic (see Table 2.1). Writing out the first few terms explicitly, we have

$$\begin{aligned}\phi(\boldsymbol{r}) =& 4\pi \frac{Y_0^0(\theta, \varphi)}{r} \int d\boldsymbol{r}_n \rho\,(\boldsymbol{r}_n)\, Y_0^{0^*}\,(\theta_n, \varphi_n) \\ &+ 4\pi \sum_{m=-1}^{1} \frac{Y_1^m(\theta, \varphi)}{3r^2} \int d\boldsymbol{r}_n \rho\,(\boldsymbol{r}_n)\, r_n Y_1^{m^*}\,(\theta_n, \varphi_n) \\ &+ 4\pi \sum_{m=-2}^{2} \frac{Y_2^m(\theta, \varphi)}{5r^3} \int d\boldsymbol{r}_n \rho\,(\boldsymbol{r}_n)\, \boldsymbol{r}_n^2 Y_2^{m^*}\,(\theta_n, \varphi_n) + \ldots\end{aligned} \tag{2.13}$$

Table 2.1. Spherical Harmonics

Y_0^0	$\sqrt{(1/4\pi)}$	$\sqrt{(1/4\pi)}$
Y_1^{-1}	$\sqrt{(3/8\pi)}\sin\theta e^{-i\varphi}$	$\sqrt{(3/8\pi)}(x-iy)/r$
Y_1^0	$\sqrt{(3/4\pi)}\cos\theta$	$\sqrt{(3/4\pi)}(z/r)$
Y_1^1	$-\sqrt{(3/8\pi)}\sin\theta e^{i\varphi}$	$-\sqrt{(3/8\pi)}(x+iy)/r$
Y_2^{-2}	$\sqrt{(15/32\pi)}\sin^2\theta e^{-2i\varphi}$	$\sqrt{(15/32\pi)}\left[(x-iy)^2/r^2\right]$
Y_2^{-1}	$\sqrt{(15/8\pi)}\sin\theta\cos\theta e^{-i\varphi}$	$\sqrt{(15/8\pi)}\left[z(x-iy)/r^2\right]$
Y_2^0	$\sqrt{(5/16\pi)}\left(3\cos^2\theta-1\right)$	$\sqrt{(5/16\pi)}\left[\left(3z^2-r^2\right)/r^2\right]$
Y_2^1	$-\sqrt{(15/8\pi)}\sin\theta\cos\theta e^{i\varphi}$	$-\sqrt{(15/8\pi)}\left[z(x+iy)/r^2\right]$
Y_2^2	$\sqrt{(15/32\pi)}\sin^2\theta e^{2i\varphi}$	$\sqrt{(15/32\pi)}\left[(x+iy)^2/r^2\right]$

Since $Y_0^0(\theta,\varphi)=1/\sqrt{4\pi}$, the first term becomes

$$\frac{1}{r}\int d\boldsymbol{r}_n\rho\left(\boldsymbol{r}_n\right)=\frac{Ze}{r}, \tag{2.14}$$

which is just the field arising from a point charge at the origin. We can make use of the spherical-harmonic addition theorem

$$\frac{4\pi}{2l+1}\sum_{m=-l}^{l}Y_l^{m^*}\left(\theta_1,\varphi_1\right)Y_l^m\left(\theta_2,\varphi_2\right)=P_l\left(\cos\theta_{12}\right) \tag{2.15}$$

to write the second term as

$$\frac{\hat{\boldsymbol{r}}}{r^2}\cdot\int d\boldsymbol{r}_n\rho\left(\boldsymbol{r}_n\right)\boldsymbol{r}_n \tag{2.16}$$

The integral is the electric-dipole-moment operator of the nucleus. If the nuclear states have definite parity, the diagonal matrix elements of this operator vanish by symmetry. The third term is the quadrupole term.

Since we shall eventually be interested in computing matrix elements of the quadrupole term, as well as various other similar operators, it is appropriate to digress for a moment to develop a technique known as operator equivalents, for rewriting such operators in a form which greatly facilitates the evaluation of their matrix elements. This technique is based on the transformation properties of these operators.

Operator Equivalents. Suppose we consider a rotation of our coordinate system through some angle θ about an axis defined by $\hat{\boldsymbol{n}}$. Let this rotation be defined by a linear operator R which rotates any vector $\boldsymbol{r}$ into $R\boldsymbol{r}$. If θ is regarded as infinitesimally small, then

$$R\boldsymbol{r}=\boldsymbol{r}+\theta\hat{\boldsymbol{n}}\times\boldsymbol{r} \tag{2.17}$$

Under such a rotation a scalar wave function $\psi(\boldsymbol{r})$ is transformed into a new wave function $\psi'(R\boldsymbol{r})$. If the system is invariant under rotation, then

$$\psi'(R\boldsymbol{r})=\psi(\boldsymbol{r}). \tag{2.18}$$

This change may be characterized by a transformation $U_R(\theta)$ defined by

$$U_R(\theta)\psi(\boldsymbol{r})=\psi'(\boldsymbol{r}). \tag{2.19}$$

Using (2.18) this becomes

$$U_R(\theta)\psi(\boldsymbol{r}) = \psi\left(R^{-1}\boldsymbol{r}\right). \tag{2.20}$$

Expanding the right side,

$$\begin{aligned} U_R(\theta)\psi(\boldsymbol{r}) = \psi(\boldsymbol{r} - \theta\hat{\boldsymbol{n}} \times \boldsymbol{r}) &\simeq \psi(\boldsymbol{r}) - (\theta\hat{\boldsymbol{n}} \times \boldsymbol{r}) \cdot \boldsymbol{\nabla}\psi(\boldsymbol{r}) \\ &\simeq \psi(\boldsymbol{r}) - \frac{i}{\hbar}(\theta\hat{\boldsymbol{n}} \times \boldsymbol{r}) \cdot \boldsymbol{p}\psi(\boldsymbol{r}). \end{aligned}$$

Therefore

$$U_R(\theta) \simeq 1 - i\theta\hat{\boldsymbol{n}} \cdot \boldsymbol{l} \tag{2.21}$$

where $\boldsymbol{l}$ is the orbital angular momentum in units of $\hbar$. For this reason we refer to $\boldsymbol{l}$ as the "generator" of infinitesimal rotations. This argument may be extended to finite rotations with the result

$$U_R(\theta) = \exp(-i\theta\hat{\boldsymbol{n}} \cdot \boldsymbol{l}) \tag{2.22}$$

Since the spinor wave function associated with the Dirac equation is defined with respect to a definite axis, a rotation in ordinary three-dimensional space will also transform the components of this wave function. Rotation in the complex two-dimensional space is described by 2×2 complex unitary matrices. For example, the matrix for a clockwise rotation of θ about the z-axis is [13], p. 109

$$\begin{pmatrix} \cos\frac{\theta}{2} - i\sin\frac{\theta}{2} & 0 \\ 0 & \cos\frac{\theta}{2} + i\sin\frac{\theta}{2} \end{pmatrix} = \mathbf{1}\cos\frac{\theta}{2} - i\sigma_z \sin\frac{\theta}{2}.$$

Since $\sigma_z^{2n} = 1$ and $\sigma_z^{2n+1} = \sigma_z$, where n is an integer, this matrix expression is equivalent to $\exp\left(-\frac{1}{2}i\theta\sigma_z\right)$ which may be generalized to

$$\exp(-i\theta\hat{\boldsymbol{n}} \cdot \boldsymbol{\sigma}/2) \tag{2.23}$$

Since $\boldsymbol{\sigma}/2$ is therefore the generator for rotations in spinor space we identify it as the spin angular momentum, s. The total angular momentum then becomes

$$j = l + s$$

and the transformation generalizes to

$$U_R(\theta) = \exp(-i\theta\hat{\boldsymbol{n}} \cdot \boldsymbol{j}) \tag{2.24}$$

The orbital part, $\boldsymbol{l}$, acts on the $\boldsymbol{r}$-dependence of the wave function while s rearranges the components of the wave function.

For many-electron systems, such as atoms or ions, the total angular momentum is the vector sum of the angular momenta of the individual electrons,

$$\boldsymbol{L} = \sum_i l_i, \quad \boldsymbol{S} = \sum_i s_i \tag{2.25}$$

and $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$.
Let $|JM\rangle$ be an eigenfunction of $\boldsymbol{J}^2$ and J_z. (Since $\boldsymbol{J}^2$ commutes with J_z one can form a complete set of common eigenfunctions of $\boldsymbol{J}^2$ and J_z.) Then

$$U_R(\theta)|JM\rangle = \exp(-i\theta\hat{\boldsymbol{n}} \cdot \boldsymbol{J})|JM\rangle \tag{2.26}$$

Inserting the identity in the form

$$\sum_{M'} |JM'\rangle \langle JM'| = 1$$

on the right gives
$U_R(\theta)|JM\rangle = \sum_{M'} \langle JM'| \exp(-i\theta \hat{\boldsymbol{n}} \cdot \boldsymbol{J})|JM\rangle \, |JM'\rangle = \sum_{M'} D^j_{MM'}(\alpha\beta\gamma) \, |JM'\rangle$.

Thus the rotation operator transforms the function $|JM\rangle$ into a linear combination of the states $|JM'\rangle$ whose coefficients are the matrix elements of the rotation operator itself, $D^J_{MM'}(\alpha\beta\gamma)$, where α, β, and γ are the Euler angles that specify the rotation.

Under the rotation U_R an operator O is transformed into $U_R O U_R^{-1}$. If the operator O consists of $2J+1$ functions $T_{JM}(M = -J, -J+1, \ldots, J)$, and if it transforms according to

$$U_R T_{JM} U_R^{-1} = \sum_{M'} D^J_{MM'}(\alpha\beta\gamma) T_{JM'} \tag{2.28}$$

then it is called an irreducible tensor operator of rank J. This may seem a rather restrictive definition. However, it turns out that many operators encountered in physical situations are, in fact, tensor operators. For example, a vector is a tensor of rank 1 ; moments of inertia and quadrupole moments are tensors of rank 2 . An example of an operator which may not be a tensor is the density operator discussed earlier.

Tensors have their own algebra, including various theorems. One of the most useful of these for our purposes is the Wigner-Eckart theorem. This states
that the matrix element of a tensor operator may be factored into a part which involves the projection quantum numbers and is independent of the tensor itself, and a part not involving the projection quantum numbers, called the reduced matrix element. The first part is, in fact, just the Clebsch-Gordan coefficient encountered in the coupling of angular momenta. Thus we have

$$\langle J'M'| \, T_{J''M''} |JM\rangle = C\left(JJ''J'; MM''M'\right) \langle J' \, \| T_{J''} \| \, J\rangle \, , \tag{2.29}$$

where $C\left(JJ''J'; MM''M'\right)$ is the Clebsch-Gordan coefficient and $\langle J' \, \| T_{J''} \| \, J\rangle$ is the reduced matrix element. Notice that if T' is also a tensor operator of the same rank as T, then the matrix elements of T are proportional to those of T'. This result has immense practical application to our magnetic Hamiltonian.

Let us return now to the quadrupole terms in $\phi(\boldsymbol{r})$. Writing the nuclear charge density as

$$\rho\left(\boldsymbol{r}_n\right) = e \sum_i \delta\left(\boldsymbol{r}_n - \boldsymbol{r}_i\right) \tag{2.30}$$

where $\boldsymbol{r}_i$ is the coordinate of a proton, the quadrapole moment Q_2^M becomes

$$Q_2^M = \sqrt{\frac{4\pi}{5}} \sum_i e r_i^2 Y_2^M\left(\theta_i, \varphi_i\right) \tag{2.31}$$

Since the quadrupole moment operators Q_2^M are proportional to the spherical harmonics, they are tensor operators of rank 2 . We can also form a tensor of rank 2 from the components of the total nuclear angular momentum $\boldsymbol{I}$. Thus if

$$T_2^{+1} = Q_2^{+1} = \frac{\sqrt{6}}{4} \sum_i z_i \left(x_i + i y_i\right) \tag{2.32}$$

this suggests that we form

$$\left(T'\right)_2^{+1} = \frac{\sqrt{6}}{4} \left(I_z I^+ + I^+ I_z\right) \tag{2.33}$$

Notice that (2.33) is written in the symmetrized form. The reason for this is that the coordinates entering (2.32) commute with each other, whereas the angular momenta do not. Therefore, in order to preserve this symmetry, we must symmetrize the operator equivalent. By the Wigner-Eckart theorem, the matrix elements of these two operators must be proportional. Thus

$$\langle IM'|\, Q_2^{+1}|IM\rangle = \alpha\, \langle IM'|\, \frac{\sqrt{6}}{4}\left(I_z I^+ + I^+ I_z\right)|IM\rangle, \tag{2.34}$$

or,

$$\langle IM'|\, Q_2^{+1}|IM\rangle = \alpha\frac{\sqrt{6}}{4}(2M+1)\sqrt{(I-M)(I+M+1)}\delta_{M',M+1} \tag{2.35}$$

It is customary to define the particular matrix element $\langle II|Q_2^0|II\rangle$ as eQ. The proportionality constant α then becomes $eQ/I(2I-1)$.

As long as we remain within a manifold in which I is a good quantum number, we may also equate the operators themselves. Thus

$$Q_2^{+1} = \frac{eQ}{I(2I-1)}\frac{\sqrt{6}}{4}\left(I_z I^+ + I^+ I_z\right) \tag{2.36}$$

with similar expressions for the other operators. The corresponding term in the quadrupole potential is therefore

$$-\sqrt{\frac{4\pi}{5}}\frac{Y_2^{-1}}{r^3}\frac{eQ}{I(2I-1)}\frac{\sqrt{6}}{4}\left(I_z I^+ + I^+ I_z\right). \tag{2.37}$$

It is obvious that we could now write the electron part in terms of the total orbital angular momentum of the electron state [8, 9]. Thus,

$$e\phi(\boldsymbol{r}) = -\frac{Ze^2}{r} + e^2 Q\xi\left[3(\boldsymbol{l}\cdot\boldsymbol{I})^2 + 3/2(\boldsymbol{l}\cdot\boldsymbol{I}) - l(l+1)I(I+1)\right],$$

where ξ is a constant that is proportional to the reduced matrix element of the electronic angular momentum. Notice that an s-state electron is not affected by the quadrupole field of the nucleus. The quadrupole field is, in general, small compared with other fields acting on the electron. From the point of view of the nucleus, however, this interaction is very important. If the electron is in a nondegenerate state characterized by the orbital quantum numbers l, m_l, and the coordinates are chosen to lie along the principal axes of the tensor $l_\mu l_\nu$, then the nuclear Hamiltonian becomes

$$\mathcal{H}_{\mathcal{Q}} = \frac{e^2 qQ}{4I(2I-1)}\left[3I_z^2 - I(I-1) + \frac{1}{2}\eta\left(I_+^2 + I_-^2\right)\right] \tag{2.38}$$

where

$$q = \xi\left\langle l_z^2\right\rangle \quad \text{and} \quad \eta = \frac{\left(\left\langle l_x^2\right\rangle - \left\langle l_y^2\right\rangle\right)}{\left\langle l_z^2\right\rangle}.$$

The same expression also characterizes the interaction with a more general surrounding charge distribution. In this case q is proportional to the electric field gradient produced by this charge distribution.

2.2.3 The Magnetic Dipole (Hyperfine) Field

The vector potential arising from the nuclear currents may also be expanded to yield

$$\begin{aligned}\boldsymbol{A}(\boldsymbol{r}) = & \int d\boldsymbol{r}_n \frac{\rho(\boldsymbol{r}_n)\,\boldsymbol{v}(\boldsymbol{r}_n)}{c\,|\boldsymbol{r}-\boldsymbol{r}_n|} \\ = & \frac{1}{cr}\int d\boldsymbol{r}_n \rho(\boldsymbol{r}_n)\,\boldsymbol{v}(\boldsymbol{r}_n) + \frac{1}{2cr^3}\int d\boldsymbol{r}_n \rho(\boldsymbol{r}_n) \\ & \times \{(\boldsymbol{r}\cdot\boldsymbol{r}_n)\,\boldsymbol{v}(\boldsymbol{r}_n) + [\boldsymbol{r}\cdot\boldsymbol{v}(\boldsymbol{r}_n)]\,\boldsymbol{r}_n\} \\ & - \frac{\boldsymbol{r}}{r^3}\times\frac{1}{2c}\int d\boldsymbol{r}_n \rho(\boldsymbol{r}_n)\,[\boldsymbol{r}_n\times\boldsymbol{v}(\boldsymbol{r}_n)] + \ldots \end{aligned} \tag{2.39}$$

If the current distribution is stationary with respect to the angular-momentum axis, then the first two terms vanish, leaving only the third. The integral in this term is the nuclear magnetic dipole moment $\boldsymbol{\mu}_I$, which is related to the nuclear angular momentum by

$$\boldsymbol{\mu}_I = g_N \mu_N \boldsymbol{I} = \gamma_N \hbar \boldsymbol{I}$$

where μ_N is the nuclear magneton, γ_N is the nuclear gyromagnetic ratio, and g_N is the nuclear g value.

Notice that the nuclear angular momentum $\boldsymbol{I}$ is in units of $\hbar$. Thus

$$\boldsymbol{A}(\boldsymbol{r}) = \boldsymbol{\mu}_I \times \frac{\boldsymbol{r}}{r^3} \tag{2.40}$$

Substituting this into the expression

$$\frac{1}{2m}\left(\boldsymbol{p} - \frac{e}{c}\boldsymbol{A}\right)^2 - \frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\times\boldsymbol{A} \tag{2.41}$$

and recognizing that $\boldsymbol{\nabla}\cdot\boldsymbol{A}(\boldsymbol{r}) = 0$, we obtain

$$\frac{\boldsymbol{p}^2}{2m} - \frac{e}{mc}\frac{(\boldsymbol{\mu}_I\times\boldsymbol{r})\cdot\boldsymbol{p}}{r^3} + \frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot\left[\frac{\boldsymbol{\mu}_I}{r^3} - 3\frac{(\boldsymbol{r}\cdot\boldsymbol{\mu}_I)\,\boldsymbol{r}}{r^5}\right]. \tag{2.42}$$

Interchanging the dot and cross products in the second term gives the orbital hyperfine interaction

$$2\mu_B \frac{\boldsymbol{\mu}_I\cdot\boldsymbol{\ell}}{r^3} \tag{2.43}$$

The last term in (2.42) containing the square brackets is the dipolar hyperfine interaction.

If the electron is in an s state, then the matrix elements of the orbital hyperfine interaction will clearly vanish. Similarly, the matrix elements of the dipolar hyperfine interaction also vanish for an s-state electron. However, there is an additional interaction for s-state electrons that is not included in expansion (2.39), since it is valid only for charge distributions which vanish at the nucleus. To obtain this additional interaction we consider the matrix element of the hyperfine interaction for an electron orbital state $\psi(\boldsymbol{r})$:

$$\begin{aligned} & -\frac{e\hbar}{2mc}\int_{\text{all space}} d\boldsymbol{r}\psi^*(\boldsymbol{r})\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\times\boldsymbol{A}(\boldsymbol{r})\psi(\boldsymbol{r}) \\ & = -\frac{e\hbar}{2mc}\int_{r<R} d\boldsymbol{r}\psi^*(\boldsymbol{r})\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\times\boldsymbol{A}(\boldsymbol{r})\psi(\boldsymbol{r}) \\ & -\frac{e\hbar}{2mc}\int_{r>R} d\boldsymbol{r}\psi^*(\boldsymbol{r})\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\times\boldsymbol{A}(\boldsymbol{r})\psi(\boldsymbol{r}). \end{aligned} \tag{2.44}$$

The radius R defines a sphere which encloses the nucleus. Outside this sphere $\boldsymbol{A}(\boldsymbol{r})$ has the form $(\boldsymbol{\mu}_I \times \boldsymbol{r})/r^3$. The second term in (2.44) gives the dipolar hyperfine interaction derived previously. The first term is the additional interaction, which may be rewritten as

$$-\frac{e\hbar}{2mc}\int_{r<R} d\boldsymbol{r}\boldsymbol{\nabla}\cdot(\boldsymbol{A}\times\boldsymbol{\sigma})|\psi(\boldsymbol{r})|^2 = -\frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot\int d\boldsymbol{S}\times\boldsymbol{A}|\psi(\boldsymbol{r})|^2. \tag{2.45}$$

Because the sphere of integration has been chosen to lie outside the nucleus, $\boldsymbol{A}(\boldsymbol{r})$ has the form $(\boldsymbol{\mu}_I \times \boldsymbol{r})/R^3$. Since $\psi(\boldsymbol{r})$ is essentially constant over this surface and is equal to $\psi(0)$, the interaction becomes

$$\begin{aligned}-\frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot\int\frac{\boldsymbol{r}\times(\boldsymbol{\mu}_I\times\boldsymbol{r})}{R^2}d\Omega|\psi(0)|^2 &= -\frac{8\pi}{3}\frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot\boldsymbol{\mu}_I|\psi(0)|^2\\ &= \frac{16\pi}{3}g_N\mu_B\mu_N\boldsymbol{I}\cdot\boldsymbol{\sigma}|\psi(0)|^2.\end{aligned} \tag{2.46}$$

This is the contact hyperfine interaction, often written as the operator

$$(8\pi/3)g_N\mu_B\mu_N\boldsymbol{I}\cdot\boldsymbol{\sigma}\delta(\boldsymbol{r}).$$

Combining these results gives us the total hyperfine interaction,

$$\begin{aligned}\mathcal{H}_{\text{hyper}} =& 2g_N\mu_B\mu_N\frac{\boldsymbol{l}\cdot\boldsymbol{I}}{r^3} - g_N\mu_B\mu_N\boldsymbol{\sigma}\cdot\left[\frac{\boldsymbol{I}}{r^3} - 3\frac{(\boldsymbol{r}\cdot\boldsymbol{I})\boldsymbol{r}}{r^5}\right]\\ &+\frac{8\pi}{3}g_N\mu_B\mu_N\boldsymbol{\sigma}\cdot\boldsymbol{I}\delta(\boldsymbol{r})\end{aligned} \tag{2.47}$$

The Hamiltonian (2.11) plus the interactions (2.38), (2.47) determine the behavior of a single electron in the presence of a nucleus.

2.2.4 Other Electrons on the Same Ion

Let us now consider the effect of other electrons. One of the most important sources of the electric field felt by an ionic electron is the Coulomb field arising from the other electrons on the same ion,

$$\phi(\boldsymbol{r}) = \sum_i \frac{e}{|\boldsymbol{r}-\boldsymbol{r}_i|}. \tag{2.48}$$

In ionic materials this interaction leads to the term levels (the determination of these many-electron states in terms of their Coulomb integrals is discussed in [10]). In itinerant-electron materials it is often assumed that the electrons experience a Coulomb repulsion only if they both happen to be in the same ionic cell. We shall consider the corresponding Hamiltonian for this situation later.

2.2.5 Crystalline Electric Fields

The Coulomb interactions between each electron and all the charges external to the ion are described by the electrostatic potential $V(\boldsymbol{r})$. In the case of irongroup ions the magnetic electrons (the $3d$ electrons) are outermost and hence are strongly affected by such a potential. In the case of rare-earth ions the magnetic $4f$ electrons are shielded by the $5s^25p^6$ shells and are less affected.

Since the charge distribution associated with neighboring ions may overlap that of the electron in question, the full treatment of this problem is very complex. The external charge

distributions of these neighboring ions are called ligands, and their effects are computed by means of ligand field theory [11]. However, for our purposes, it will be sufficient to treat the neighboring ions as point charges; the problem may then be handled by crystal field theory. The advantage of using point charges is that $V(\boldsymbol{r})$ satisfies Laplace's equation and may be expanded in spherical harmonics as

$$V(r,\theta,\varphi) = \sum_{L'} \sum_{M'} A_{L'}^{M'} r^{L'} Y_{L'}^{M'}(\theta,\varphi) \tag{2.49}$$

The number of terms that need be considered is greatly reduced, for the following reasons. Suppose we consider an iron-group ion in a crystal field. Then we shall eventually be interested in matrix elements of the form $\int \chi^* V \psi d\tau$, where χ and ψ are d-electron wave functions. Since the density $\chi^*\psi$, when expanded in spherical harmonics will not contain terms with $L' > 4$, the integrals with $L' > 4$ will vanish by orthogonality of the spherical harmonics. Similarly, the integral vanishes for all terms in V which have L' odd. The term for $L' = 0$ is usually dropped because it is an additive constant. If we are considering several $3d$ electrons within a term, then $L'_{\max}$ is determined by the L value of this term (for example, $L' \leq 6$ for an F-state ion, i.e., $L' \leq 2L$).

The potential energy of a charge q' at (r, θ, φ), in a potential due to charges q at a distance d from the origin and arranged in a cubic coordination, is

$$\begin{aligned} V_c(r,\theta,\varphi) =& D'_4 \left\{ Y_4^0(\theta,\varphi) + \sqrt{\frac{5}{14}} \left[Y_4^4(\theta,\varphi) + Y_4^{-4}(\theta,\varphi) \right] \right\} \\ &+ D'_6 \left\{ Y_6^0(\theta,\varphi) - \sqrt{\frac{7}{2}} \left[Y_6^4(\theta,\varphi) + Y_6^{-4}(\theta,\varphi) \right] \right\} \end{aligned} \tag{2.50}$$

where, for example, $D'_4 = +\frac{7}{3}\sqrt{\pi} q' r^4/d^5$ for sixfold coordination. The coefficients in such expansions have been tabulated by Hutchings [12]. The potential may also be expressed in terms of cartesian coordinates (see Table 2.1). Thus the above potential may be written as

$$\begin{aligned} V_c(x,y,z) =& C_4 \left[\left(x^4 + y^4 + z^4\right) - \frac{3}{5} r^4 \right] + C_6 \left[\left(x^6 + y^6 + z^6\right) \right. \\ &\left. + \frac{15}{4} \left(x^2 y^4 + x^2 z^4 + y^2 x^4 + y^2 z^4 + z^2 x^4 + z^2 y^4\right) - \frac{15}{14} r^6 \right] \end{aligned} \tag{2.51}$$

where $C_4 = +\frac{35}{4} q q'/d^5$ for sixfold coordination.

We are now faced with the problem of calculating the matrix elements of this potential. This is easily accomplished by the operator-equivalent method, which makes use of the fact that the matrix elements of operators involving x, y, and z within a given L or J manifold are proportional to those of L_x, L_y, and L_z or J_x, J_y, and J_z. As pointed out earlier, the fact that the angularmomentum operators do not commute necessitates some care in constructing the operator equivalents. Fortunately there are tables for these (a good source is [12]). For example, within a manifold where L is constant the sum of the potential energies of all the electrons contributing to L is

$$\sum\left(x^4+y^4+z^4-\frac{3}{5}r^4\right)$$
$$\Rightarrow\frac{\beta\overline{r^4}}{8}\left[35L_z^4-30L(L+1)L_z^2+25L_z^2-6L(L+1)+3L^2(L+1)^2\right]$$
$$+\frac{\beta\overline{r^4}}{8}\left[\left(L^+\right)^4+\left(L^-\right)^4\right]\equiv\frac{\beta\overline{r^4}}{20}O_4^0+\frac{\beta\overline{r^4}}{4}O_4^4=B_4^0O_4^0+B_4^4O_4^4 \tag{2.52}$$

where $\overline{r^4}$ is the average value of the fourth power of the electron radius. The operators O_n^m appear frequently in the literature. The ground state β is a constant which depends on the term; for a ^2D or a ^5D term $\beta=\frac{2}{63}$.

Consider a single $3d$ electron. This has the term ^2D, which is fivefold orbitally degenerate, with states $^2D\,(L_z,S_z)$. The matrix elements of V_c are [12]

$$\begin{matrix}{}^2D(2) & {}^2D(1) & {}^2D(0) & {}^2D(-1) & {}^2D(-2)\end{matrix}$$
$$\frac{\beta\overline{r^4}}{20}\begin{bmatrix}12 & 0 & 0 & 0 & 60\\ 0 & -48 & 0 & 0 & 0\\ 0 & 0 & 72 & 0 & 0\\ 0 & 0 & 0 & -48 & 0\\ 60 & 0 & 0 & 0 & 12\end{bmatrix}. \tag{2.53}$$

The eigenvalues and eigenvectors are easily found to be

Energy	Eigenfunctions
$\frac{12}{5}\beta\overline{r^4}$	$^2D\,(1,S_z)$
	$^2D\,(-1,S_z)$
	$\frac{1}{\sqrt{2}}\left[^2D\,(2,S_z)-{}^2D\,(-2,S_z)\right]$
$-\frac{18}{5}\beta\overline{r^4}$	$^2D\,(0,S_z)$
	$\frac{1}{\sqrt{2}}\left[^2D\,(2,S_z)+{}^2D\,(-2,S_z)\right].$

The quantity 6β is often written as Δ. Therefore we find that the ^2D term is split into two states separated by $C_4\overline{r^4}\Delta$. Notice that C_4 can be positive or negative, depending on the coordination. This is illustrated in Fig. 2.1.

The nature of such splittings obviously depends on the symmetry of the crystal field. For this reason group theory is a powerful tool in determining the degeneracies associated with various symmetries. Group theory as applied to crystal wave functions is discussed fully elsewhere [13], Chap. 4, but it is worth our while to digress again briefly to introduce some of the group-theory terminology and notation which will enter our discussions from time to time.

Symmetry Representations. The symmetry of a system is generally specified by those operations which leave its physical appearance unchanged. For example, the symmetry operations which leave an equilateral triangle unchanged are listed in Table 2.2. A collection of symmetry operations that satisfies certain conditions is called a group. In order to take advantage of the powerful

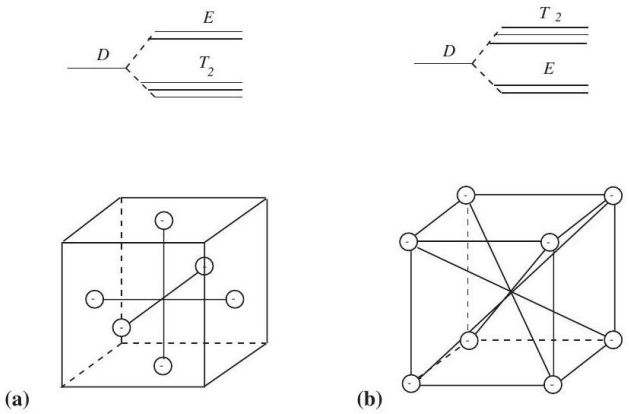
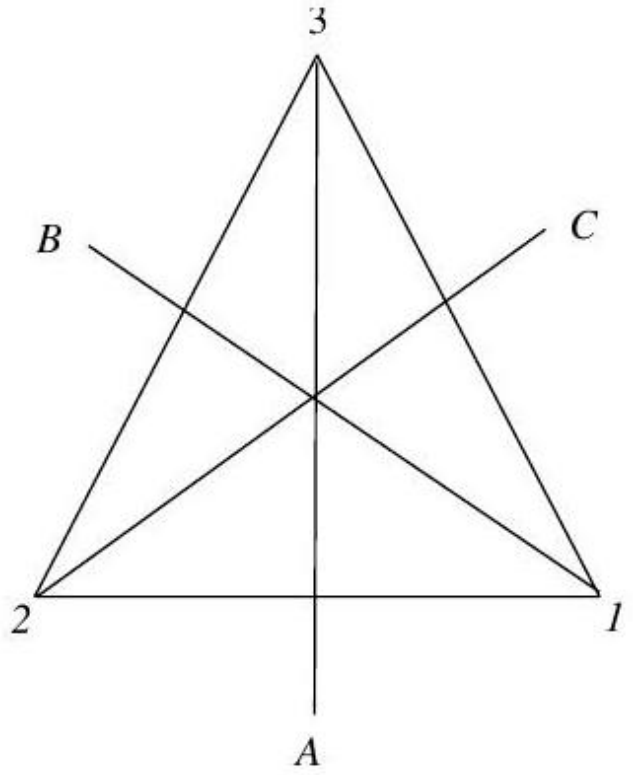


Fig. 2.1. Splitting of a D state in a cubic crystal field for (a) sixfold coordination and (b) eightfold coordination

Table 2.2. Symmetry operations associated with an equilateral triangle



E = identity
C_2^A = rotation of π around the A-axis
C_2^B = rotation of π around the B-axis
C_2^C = rotation of π around the C-axis
C_3^z = rotation of $2\pi/3$ around the z-axis
$(C_3^z)^2$ = rotation of $4\pi/3$ around the z-axis

theorems associated with group theory we always work with those symmetry operations which do, in fact, constitute a group. These operations may be specifically represented by matrices which describe how a coordinate point transforms under the particular symmetry operation. Thus, if $\boldsymbol{r}$ represents a rotation and $\boldsymbol{t}$ represents a translation, the most general coordinate transformation is $\boldsymbol{x}' = \boldsymbol{r}\boldsymbol{x}+\boldsymbol{t}$. Such a collection of operations is called a space group. The rotational part, obtained by setting $\boldsymbol{t} = 0$, itself forms a group, called the point group. When we are dealing with noninteracting ions, the point group is sufficient to characterize

the properties of the system. However, for interacting systems the full space group must be employed. Fortunately, since the point groups in a crystal must be compatible with translational symmetry, there are only 32 such groups [13], p. 55 . Our equilateral triangle is characterized by the point group labeled D_3 in the so-called Schöflies notation, or 32 in the "international" notation.

Now, let us consider a function whose form depends on the arrangement of the system. For example, suppose that three protons are located at the vertices of an equilateral triangle. The energy of an electron in such an environment depends on the positions of the protons, but this energy is unchanged under any permutation of the protons. Notice that there are 3 !, or six, such permutations. These are just the result of the six symmetry operations which leave the triangle invariant. If a figure is defined by some arrangement of identitical particles, the operations which leave it invariant also leave the interaction energy between these two particles and other particles invariant. Thus it is convenient to introduce a new group, isomorphic to the coordinate-transformations group, in which the group elements are operators which operate on functions rather than on coordinates. These operators are defined by

$$P_R f(\boldsymbol{x}) \equiv f\left(\boldsymbol{r}^{-1}\boldsymbol{x}\right). \tag{2.54}$$

The particular function with which we shall be concerned is the energy in its operator form, the Hamiltonian $\mathcal{H}$. Those symmetry operations which leave the Hamiltonian invariant comprise the "group of the Schrödinger equation".

If P_R leaves $\mathcal{H}$ invariant, then it must commute with $\mathcal{H}$. Therefore

$$P_R\mathcal{H}\psi_n = \mathcal{H}P_R\psi_n = E_nP_R\psi_n \tag{2.55}$$

Thus any function $P_R\psi_n$ obtained by operating on an eigenfunction ψ_n with a symmetry operator of the group of the Schrödinger equation will also be an eigenfunction having the same energy. Suppose that the state n is l_n-fold degnerate. Then the function $P_R\psi_n$ must be a linear combination of l_n orthonormal eigenfunctions, which we shall denote $\psi_\mu^{(n)}$ $(\mu = 1, \ldots, l_n)$. Therefore

$$P_R\psi_\mu^{(n)} = \sum_{\nu=1}^{l} \boldsymbol{\Gamma}^{(n)}(\boldsymbol{r})_{\mu\nu}\psi_\nu^{(n)} \tag{2.56}$$

The transformation coefficients constitute a set of matrices which form an irreducible representation of the group of the Schrödinger equation. Furthermore, we see that the nth representation is associated with the nth eigenstate, and the dimensionality of the representation is equal to the degeneracy of this eigenstate. This representation is irreducible, since there is always an operator in the group that will transform each function into any other function. If this were not true, we could construct smaller sets of states which would, in general, have different eigenvalues, contradicting our original hypothesis.

Because of the relation (2.56), we speak of the $\psi^{(n)}$ as "transforming according to $\Gamma^{(n)}$ ". For this reason energy eigenstates are labeled by their irreducible representations. Also, since the representations are generated from the eigenfunctions, we say that the l_n degenerate eigenfunctions $\psi_\mu^{(n)}$ form a basis for an l_n-dimensional representation $\Gamma^{(n)}$ of the group of the Schrödinger equation.

The number and nature of the irreducible representations associated with the various symmetry groups have all been tabulated in what is known as a character table. The character $\chi_n(\boldsymbol{r})$ associated with the operation $\boldsymbol{r}$ belonging to the nth irreducible representation is merely the trace of the matrix of that representation, that is,

$$\chi_n(\boldsymbol{r}) = \sum_{\mu} \Gamma^{(n)}(\boldsymbol{r})_{\mu\mu} \tag{2.57}$$

One of the powerful features of group theory is that it enables us to determine the irreducible representations and all their characters without ever having to know specifically the basis functions. The character table for the equilateraltriangle symmetry group D_3 is given in Table 2.3. To see what this character table implies, let us suppose that we have a single electron, bound, say, to some ionic core giving rise to certain eigenstates. Since this system has complete rotational symmetry, these states are labeled by the familiar s, p, d, etc. Let us surround the system by three protons located at the vertices of an equilateral triangle. The symmetry of this system is D_3. The character table for D_3 tells us that the eigenfunctions of the electron are now labeled by the irreducible

Table 2.3. Character table for the point group D_3

Symmetry Group	Operations		
D_3	E	$2C_3$	$3C_3'$
Irreducible representations	Characters		
A_1	1	1	1
A_2	1	1	-1
E	2	-1	0

representations A_1, A_2, or E. The character associated with the identity operation, E, tells us the degeneracy of the various representations. For example, A_1 and A_2 are nondegenerate states, and E is doubly degenerate.

Exactly how the original states decompose into the new types of states depends on the original states themselves and on the symmetry group characterizing the environment. Fortunately, such decompositions have been tabulated for a great number of situations. For example, in Fig. 2.1 we saw that a D state, when exposed to a cubic crystal field, splits into a doublet labeled by E and a triplet labeled by T_2. Group theory, however, does not tell us the ordering of the states or their relative separations. Such specific information can be obtained only by doing a calculation, as we did at the beginning of this section.

Quenching. At this point it is convenient to introduce a general property of angular momentum. This might be stated as a theorem:

The matrix element of the orbital angular momentum between non-degenerate states has an arbitrary phase. In particular, it may be pure real or pure imaginary.

To prove this let us consider the time-reversal operator T acting on a state ψ [13], p. 141. If we neglect the spin, then $T\psi = \psi^*$. Furthermore, if the Hamiltonian of the system is Hermitian, then ψ^* has the same eigenvalue as ψ, but if ψ is nondegenerate, then ψ and ψ^* must be linearly dependent. That is, $\psi^* = c\psi$, where c is a coefficient of proportionality. Operating on this relation with T gives $\psi = |c|^2\psi$ which requires that $|c|^2 = 1$ or $c = e^{i\varphi}$, where φ is a real quantity.

Now consider the matrix element $\langle n|\boldsymbol{L}|m\rangle$. Inserting the identity operator $T^{-1}T$, we may write this as $\langle n|T^{-1}T\boldsymbol{L}T^{-1}T|m\rangle$. Under time reversal the angular momentum changes sign, $T\boldsymbol{L}T^{-1} = -\boldsymbol{L}$. Also, since T satisfies $\langle n|T^{-1}|m\rangle = \langle Tn \mid m\rangle^*$, we obtain

$$\langle n|\boldsymbol{L}|m\rangle = -\exp\left[i\left(\varphi_n - \varphi_m\right)\right]\langle n|\boldsymbol{L}|m\rangle^*.$$

Since the phases are arbitrary we could, for example, choose $\varphi_n - \varphi_m = 0$, in which case this matrix element would be the negative of its complex conjugate,
which would make it pure imaginary. If $\varphi_n - \varphi_m = \pi/2$, then it would be pure real.

This theorem has an important corollary:
The expectation value of $\boldsymbol{L}$ for any nondegenerate state is 0 .
In the theorem, if $|m\rangle = |n\rangle$, then $\langle n|\boldsymbol{L}|n\rangle$ must be pure imaginary. But $\langle n|\boldsymbol{L}|n\rangle$ is a physical observable. Therefore it must be 0.

Thus if the crystal field has sufficiently low symmetry to remove all the orbital degeneracy, then, to lowest order, the orbital angular momentum is 0 , and we say that the crystal field has completely quenched it. For this reason the static susceptibility of iron-group salts is found experimentally to arise predominantly from the spin.

2.2.6 Dipole-Dipole Interaction

The magnetic neighbors surrounding a given ion will contribute to the vector potential a term similar to that which we found for the electron-nucleus magnetic coupling, (2.42). If the ions have moments $\boldsymbol{\mu}_i$, the dipole-dipole interaction has the form

$$\mathcal{H}_{\text{dip}} = \sum_{\substack{i,j\neq j \\ i}} \frac{1}{r_{ij}^3} \left[\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j - 3\left(\boldsymbol{\mu}_i \cdot \hat{\boldsymbol{r}}_{ij}\right)\left(\boldsymbol{\mu}_j \cdot \hat{\boldsymbol{r}}_{ij}\right)\right] \tag{2.58}$$

It is convenient to separate this into various terms, the meaning of which will become evident later. Assuming the moments arise from spin, $\boldsymbol{\mu}_i = g\mu_B \boldsymbol{S}_i$, and (2.58) becomes

$$\begin{aligned}
\mathcal{H}_{\text{dip}} =& g^2\mu_B^2 \sum_{i>j} \Bigg\{ -\frac{3\cos^2\theta_{ij} - 1}{r_{ij}^3} S_i^z S_j^z \\
&+ \frac{3\cos^2\theta_{ij} - 1}{4r_{ij}^3}\left(S_i^+ S_j^- + S_i^- S_j^+\right) \\
&- \frac{3}{2}\frac{\sin\theta_{ij}\cos\theta_{ij}\exp\left(-i\varphi_{ij}\right)}{r_{ij}^3}\left(S_i^z S_j^+ + S_i^+ S_j^z\right) \\
&- \frac{3}{2}\frac{\sin\theta_{ij}\cos\theta_{ij}\exp\left(-i\varphi_{ij}\right)}{r_{ij}^3}\left(S_i^z S_j^- + S_i^- S_j^z\right) \\
&- \frac{3}{2}\frac{\sin^2\theta_{ij}}{r_{ij}^3}\left[\exp\left(-2i\varphi_{ij}\right) S_i^+ S_j^+ + \exp\left(2i\varphi_{ij}\right) S_i^- S_j^-\right] \Bigg\}
\end{aligned} \tag{2.59}$$

where θ_{ij} and φ_{ij} are the angles that $\boldsymbol{r}_{ij}$ makes with the fixed coordinate system.

2.2.7 Direct Exchange

The exchange energy is the contribution to the interaction energy of a system of electrons which arises from the use of antisymmetrized wave functions, as opposed to single products of one-electron wave functions. Under certain conditions the same effect may be achieved with single-product wave functions if an exchange-interaction term is added to the Hamiltonian. This effect was discovered simultaneously and independently by Dirac and Heisenberg in 1926. Since then a great deal of work has been done on this subject, particularly in developing the appropriate Hamiltonian [14]. In this section we shall discuss the origin of exchange and indicate the approximations under which it may be represented by an effective interaction Hamiltonian.

Let us begin by considering two electrons interacting with each other and a fixed positive point charge Ze. Let us assume that we know the eigenfunctions of the one-electron Hamiltonian

$\mathcal{H}_0(\boldsymbol{r},\boldsymbol{\sigma})$. For the time being let us assume that this does not include the spin-orbit interaction. Then $\mathcal{H}_0(\boldsymbol{r},\boldsymbol{\sigma}) = \mathcal{H}_0(\boldsymbol{r})$ and we may write the eigenfunctions as products of an orbital function $\varphi_n(\boldsymbol{r})$ and a spinor $\eta_\mu(\boldsymbol{\sigma})$. We shall consider the modifications introduced by the spin-orbit interaction later.

The two-electron Hamiltonian is

$$\mathcal{H} = \mathcal{H}_0(\boldsymbol{r}_1) + \mathcal{H}_0(\boldsymbol{r}_2) + \frac{e^2}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} \tag{2.60}$$

Let us assume that the electron-electron interaction is, on average, smaller than $\mathcal{H}_0$, so that it may be treated by perturbation theory. We must now determine what functions to use as the basis for computing the matrix elements of the electron-electron interaction. The fact that the Hamiltonian without the electron-electron interaction is separable suggests that we try product wave functions. Thus, if electron 1 is in an orbital state n with spin up and electron 2 is in an orbital state m, also with spin up, we might try $\varphi_n(\boldsymbol{r}_1)\,\alpha(\boldsymbol{\sigma}_1)\,\varphi_m(\boldsymbol{r}_2)\,\alpha(\boldsymbol{\sigma}_2)$, where α is the spin-up spinor. However, the Pauli exclusion principle requires that the wave functions be antisymmetric with respect to particle interchanges. This condition may be satisfied by writing the wave function as a normalized Slater determinant. If the single-electron wave functions are orthogonal, the appropriate determinantal wave function is

$$\frac{1}{\sqrt{2}} \begin{vmatrix} \varphi_n(\boldsymbol{r}_1)\,\alpha(\boldsymbol{\sigma}_1) & \varphi_n(\boldsymbol{r}_2)\,\alpha(\boldsymbol{\sigma}_2) \\ \varphi_m(\boldsymbol{r}_1)\,\alpha(\boldsymbol{\sigma}_1) & \varphi_m(\boldsymbol{r}_2)\,\alpha(\boldsymbol{\sigma}_2) \end{vmatrix}. \tag{2.61}$$

Since there are an infinite number of orbital states, we could construct an infinite number of such Slater determinants. The general wave function would be a linear combination of such determinants. However, if the electron-electron interaction is small, we may neglect this admixture of other orbital states. In particular, let us assume that electron 1 has a low-lying nondegenerate
orbital state φ_a with energy E_a and electron 2 has a similar low-lying nondegenerate orbital state φ_b with energy E_b. If both spin functions are up, the determinantal wave function becomes

$$\psi_1 = \frac{1}{\sqrt{2}} \begin{vmatrix} \varphi_a(\boldsymbol{r}_1)\,\alpha(\boldsymbol{\sigma}_1) & \varphi_a(\boldsymbol{r}_2)\,\alpha(\boldsymbol{\sigma}_2) \\ \varphi_b(\boldsymbol{r}_1)\,\alpha(\boldsymbol{\sigma}_1) & \varphi_b(\boldsymbol{r}_2)\,\alpha(\boldsymbol{\sigma}_2) \end{vmatrix}. \tag{2.62}$$

If the spin function associated with orbital a is down, then

$$\psi_2 = \frac{1}{\sqrt{2}} \begin{vmatrix} \varphi_a(\boldsymbol{r}_1)\,\beta(\boldsymbol{\sigma}_1) & \varphi_a(\boldsymbol{r}_2)\,\beta(\boldsymbol{\sigma}_2) \\ \varphi_b(\boldsymbol{r}_1)\,\alpha(\boldsymbol{\sigma}_1) & \varphi_b(\boldsymbol{r}_2)\,\alpha(\boldsymbol{\sigma}_2) \end{vmatrix}. \tag{2.63}$$

There are two additional possible spin configurations which lead to wave functions ψ_3 and ψ_4. These four functions form a complete orthonormal set and therefore constitute an appropriate bais with which to evaluate the matrix elements of $\mathcal{H}$. The result is

$$H = \begin{bmatrix} E_a + E_b + K_{ab} - J_{ab} & 0 & 0 & 0 \\ 0 & E_a + E_b + K_{ab} - J_{ab} & 0 & \\ 0 & -J_{ab} & E_a + E_b + K_{ab} & 0 \\ 0 & 0 & 0 & E_a + E_b + K_{ab} - J_{ab} \end{bmatrix} \tag{2.64}$$

where

$$K_{ab} = \iint d\boldsymbol{r}_1 d\boldsymbol{r}_2 \frac{e^2}{r_{12}} |\varphi_a(\boldsymbol{r}_1)|^2 |\varphi_b(\boldsymbol{r}_2)|^2 \tag{2.65}$$

and

$$J_{ab} = \iint d\boldsymbol{r}_1 d\boldsymbol{r}_2 \varphi_a^*(\boldsymbol{r}_1)\,\varphi_b^*(\boldsymbol{r}_2)\,\frac{e^2}{r_{12}}\varphi_b(\boldsymbol{r}_1)\,\varphi_a(\boldsymbol{r}_2) \tag{2.66}$$

Diagonalizing this matrix gives a singlet with energy

$$E_s = E_a + E_b + K_{ab} + J_{ab} \tag{2.67}$$

and a triplet with energy

$$E_t = E_a + E_b + K_{ab} - J_{ab} \tag{2.68}$$

Since J_{ab} is the self-energy of the charge distribution $e\varphi_a^*(\boldsymbol{r})\varphi_b(\boldsymbol{r})$, it is positive definite. Therefore the triplet always has a lower energy than the singlet. This is the origin of Hund's rule, which says that the ground state of an atom has maximum multiplicity.

Dirac noticed that the eigenvalues (2.67), (2.68) could be obtained with a basis consisting only of products of spin functions if an exchange interaction were added to the Hamiltonian. To obtain the form of this effective interaction
term we notice that just as any 2×2 matrix may be expressed as a linear combination of Pauli matrices plus the unit matrix, any 4×4 matrix may be written as a quadratic function of direct products of Pauli matrices [13], p. 320. For example, if

$$\sigma_{1x} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad \sigma_{2x} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{2.69}$$

then

$$\sigma_{1x} \otimes \sigma_{2x} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}. \tag{2.70}$$

We are particularly interested in that quadratic form which gives three equal eigenvalues. Such a form is

$$\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{2.71}$$

Therefore the Hamiltonian which will produce in a spinor basis the same eigenvalues as (2.52) evaluated in a fully antisymmetrized basis is

$$\overline{\mathcal{H}} = \frac{1}{4}(E_s + E_t) - \frac{1}{4}(E_s - E_t)\,\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 = \text{ const. } - \frac{1}{4}J\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2. \tag{2.72}$$

Thus the exchange interaction, which is a purely electrostatic effect, may be expressed as a spin-spin interaction. The exchange parameter J is $E_s - E_t$. If J is positive, we say that the interaction is ferromagnetic.

In obtaining the exchange interaction (2.72) we have made two important assumptions. The first was that we could restrict ourselves to a certain subset of nondegenerate orbital states. There is no real justification for this, as the Coulomb interaction does, in fact, couple different orbital states. We shall see this more clearly in our discussion of exchange in the N-electron system. The second assumption was that the orbital functions were orthogonal. When we are dealing with wave functions that have a common origin, as in an atom, this is usually the case.

However, as soon as we begin talking about electrons centered at different sites the problem becomes very complex.

The hydrogen molecule is perhaps the simplest example of such a twocenter problem. This was first considered by Heitler and London in 1927 [15]. In the limit of infinite separation we shall assume we have two neutral hydrogen atoms. The Hamiltonian for a single hydrogen atom located at $\boldsymbol{r}_a$ is

$$\mathcal{H}_\infty(\boldsymbol{r}) = -\frac{\hbar^2\nabla^2}{2m} - \frac{e^2}{|\boldsymbol{r}-\boldsymbol{r}_a|}. \tag{2.73}$$

The corresponding orbital eigenfunctions are $\varphi(\boldsymbol{r}-\boldsymbol{r}_a) \equiv a(r)$. The Hamiltonian for the two-proton system is

$$\begin{aligned}\mathcal{H} = &-\frac{\hbar^2\nabla_1^2}{2m} - \frac{\hbar^2\nabla_2^2}{2m} - \frac{e^2}{|\boldsymbol{r}_1-\boldsymbol{r}_a|} - \frac{e^2}{|\boldsymbol{r}_1-\boldsymbol{r}_b|} - \frac{e^2}{|\boldsymbol{r}_2-\boldsymbol{r}_a|} - \frac{e^2}{|\boldsymbol{r}_2-\boldsymbol{r}_b|} \\ &+ \frac{e^2}{|\boldsymbol{r}_1-\boldsymbol{r}_2|} + \frac{e^2}{|\boldsymbol{r}_a-\boldsymbol{r}_b|}.\end{aligned} \tag{2.74}$$

Heitler and London assumed basis functions of the form (2.61) but rather than taking eigenfunctions of noninteracting electrons in the field of two nuclei, they took the eigenfunctions of isolated free atoms. Thus, if both electrons are "up",

$$\psi_1 = \frac{1}{\sqrt{2-2\ell^2}} \begin{vmatrix} a(1)\alpha(1) & a(2)\alpha(2) \\ b(1)\alpha(1) & b(2)\alpha(2) \end{vmatrix}, \tag{2.75}$$

where $\ell = \int d\boldsymbol{r} a^*(r)b(r)$ is the overlap integral. If one spin is "up" and the other "down" there are two Slater determinants

$$\phi_1 = \begin{vmatrix} a(1)\alpha(1) & a(2)\alpha(2) \\ b(1)\beta(1) & b(2)\beta(2) \end{vmatrix} \tag{2.76}$$

and

$$\phi_2 = \begin{vmatrix} a(1)\beta(1) & a(2)\beta(2) \\ b(1)\alpha(1) & b(2)\alpha(2) \end{vmatrix}. \tag{2.77}$$

Due to the overlap, these are not orthogonal, i.e.,

$$\langle \phi_1 \mid \phi_2 \rangle = -2\ell^2. \tag{2.78}$$

An orthogonalized pair of basis functions is

$$\psi_2 = \frac{1}{2\sqrt{1-\ell^2}}(\phi_1 + \phi_2) \tag{2.79}$$

$$\psi_3 = \frac{1}{2\sqrt{1-\ell^2}}(\phi_1 - \phi_2). \tag{2.80}$$

The fourth basis function is identical to (2.75) with β in place of α. In this basis the Hamiltonian matrix is, in fact, diagonal:

$$H = \begin{pmatrix} E_t & 0 & 0 & 0 \\ 0 & E_t & 0 & 0 \\ 0 & 0 & E_s & 0 \\ 0 & 0 & 0 & E_t \end{pmatrix}, \tag{2.81}$$

where

$$E_t = E_a^0 + E_b^0 + \frac{e^2}{|\boldsymbol{r}_a - \boldsymbol{r}_b|} - \frac{K - \Delta_a - \Delta_b - 2\ell\Delta_{ab} - J}{1 - \ell^2}, \tag{2.82}$$

$$E_s = E_a^0 + E_b^0 + \frac{e^2}{|\boldsymbol{r}_a - \boldsymbol{r}_b|} + \frac{K - \Delta_a - \Delta_b - 2\ell\Delta_{ab} + J}{1 + \ell^2}. \tag{2.83}$$

Here Δ_a and Δ_b represent shifts in the distant-atom eigenvalues, E_a^0 and E_b^0, respectively, due to the other core. Δ_{ab} is the interaction of the "overlap charge density", $a(r)b(r)$, with a core, and

$$K = \iint d\boldsymbol{r}_1 d\boldsymbol{r}_2 \left|a\left(r_1\right)\right|^2 \left|b\left(r_2\right)\right|^2 \frac{e^2}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} \tag{2.84}$$

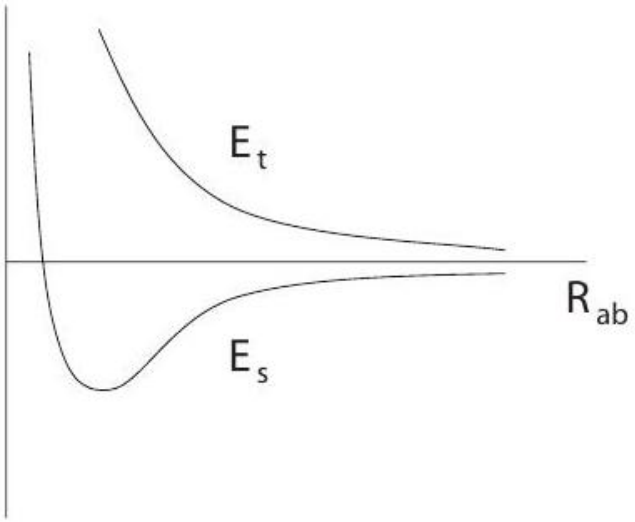
and

$$J = \iint d\boldsymbol{r}_1 d\boldsymbol{r}_2 a^*\left(r_1\right) b^*\left(r_2\right) \frac{e^2}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} b\left(r_1\right) a\left(r_2\right) \tag{2.85}$$

The difference between the singlet and the triplet is

$$E_s - E_t = -2\frac{\left(K - \Delta_a - \Delta_b - 2\ell\Delta_{ab}\right)\ell^2 - J}{1 - \ell^4}. \tag{2.86}$$

Notice that this may be positive or negative, depending upon the relative sizes of the various parameters. Thus, it is not obvious whether the ground state will be ferromagnetic or antiferromagnetic. Actual evaluation shows that for realistic separations the singlet lies lowest. As the nuclei are brought together, and ℓ increases, the denominators in E_t and E_s cause E_t to increase and E_s to decrease. Eventually, internuclear repulsion also causes E_s to increase:



Equation (2.86) can be evaluated exactly for the case of hydrogenic wave functions. It is found that for very large separations the triplet has lower energy. This cannot be, for, as Herring pointed out, the lowest eigenvalue of a semibounded Sturm-Liouville differential operator, such as (2.74), must be free of nodes. This means it must always be a singlet. This would be a good description of the chemical bond, but would not explain magnetism. The problem arises from the oversimplified nature of the Heitler-London states. The exchange coupling measures the rate at which two identifiable electrons
exchange places by tunneling through the barrier separating them. In the Heitler-London approximation this tunneling is uncorrelated. In reality, however, the two electrons will tend to avoid one another. This reduces the amplitude of the wave function for configurations with

both electrons close to the internuclear line. Fortunately, however, (2.85) is a reasonably good approximation for separations generally encountered.

In considering more than two electrons we find that difficulties arise from the nonorthogonality of the wave functions we have been using thus far. When the Heitler-London method is applied to a very large system the nonorthogonality integrals enter the secular equation with high powers and lead to an apparent divergence. This "nonorthogonality catastrophy" is a purely mathematical difficulty [14], and Herring has reviewed various treatments that show that even for large systems the energies and eigenstates are given by the exchange interaction with exchange constants having the same values as for a two-site system.

One approach to the problem of exchange among many sites is to give up our well-defined but nonorthogonal functions and work with functions which are orthogonal. An example of such a set of orthogonal functions are Wannier functions. The Wannier function $\phi_{n\lambda}(\boldsymbol{r}-\boldsymbol{r}_a)$ resembles the nth atomic orbital with spin λ near the α th lattice site, but it falls off throughout the crystal in such a way that it is orthogonal to similar functions centered at other sites. Since the exchange interaction is essentially a quantum-statistical effect the technique of second quantization discussed in Chap. 1 is very convenient for obtaining the exchange Hamiltonian.

Let us consider N electrons reasonably localized on N lattice sites. The Hamiltonian for such a system is

$$\mathcal{H}=\sum_i \frac{p_i^2}{2m}-\sum_{i,a}\frac{Ze^2}{|\boldsymbol{r}_i-\boldsymbol{r}_a|}+\frac{1}{2}\sum_{i,j}\frac{e^2}{|\boldsymbol{r}_i-\boldsymbol{r}_j|}. \tag{2.87}$$

In Chap. 1 we found that the prescription for second quantizing such a Hamiltonian entailed introducing a field operator which could be expanded in terms of a complete set of single-particle wave functions. In this case the appropriate set of functions are the Wannier functions. Thus the field operator (1.122) becomes

$$\psi(\boldsymbol{r})=\sum_{\alpha,n,\lambda}\phi_{n\lambda}(\boldsymbol{r}-\boldsymbol{r}_\alpha)\,a_{n\lambda}(\boldsymbol{r}_\alpha) \tag{2.88}$$

where $a_{n\lambda}(\boldsymbol{r}_\alpha)$ annihilates an electron in orbital state n and spin state λ at the lattice site α. The interaction part of the Hamiltonian (2.87) becomes

$$\begin{aligned}\frac{1}{2}\sum_{\substack{\alpha_1,\alpha_2,\alpha_3,\alpha_4\\ n_1,n_2,n_3,n_4\\ \lambda_1,\lambda_2}} &\langle\alpha_1 n_1;\alpha_2 n_2|\,V\,|\alpha_3 n_3;\alpha_4 n_4\rangle\\ &\times a^\dagger_{n_1\lambda_1}(\boldsymbol{r}_{\alpha_1})\,a^\dagger_{n_2\lambda_2}(\boldsymbol{r}_{\alpha_2})\,a_{n_4\lambda_2}(\boldsymbol{r}_{\alpha_4})\,a_{n_3\lambda_1}(\boldsymbol{r}_{\alpha_3})\end{aligned} \tag{2.89}$$

Since the Wannier states are localized to within a unit cell, the main contributions to (2.89) arise from those terms in which $\alpha_3=\alpha_1$ and $\alpha_4=\alpha_2$ or $\alpha_3=\alpha_2$ and $\alpha_4=\alpha_1$. The remaining terms involve various orbital excitations induced by the Coulomb interaction. These lead to off-diagonal exchange. Just as in the two-electron case, we shall restrict each electron to a definite orbital state. That is, we shall keep only those terms in which $n_3=n_1$ and $n_4=n_2$ or $n_3=n_2$ and $n_4=n_1$. If, for simplicity, we also neglect orbital-transfer terms, in which two electrons interchange orbital states, then (2.89) reduces to

$$\frac{1}{2}\sum_{\substack{\alpha,\alpha'\\ n,n'\\ \lambda,\lambda'}}\Big[\langle\alpha,n;\alpha',n'|\,V\,|\alpha,n;\alpha',n'\rangle\, a^{\dagger}_{n\lambda}\left(\boldsymbol{r}_{\alpha}\right) a^{\dagger}_{n'\lambda'}\left(\boldsymbol{r}_{\alpha'}\right) a_{n'\lambda'}\left(\boldsymbol{r}_{\alpha'}\right) a_{n\lambda}\left(R_{\alpha}\right)$$
$$+\langle\alpha,n;\alpha',n'|\,V\,|\alpha',n';\alpha,n\rangle\, a^{\dagger}_{n\lambda}\left(\boldsymbol{r}_{\alpha}\right) a^{\dagger}_{n'\lambda'}\left(\boldsymbol{r}_{\alpha'}\right) \alpha_{n,\lambda'}\left(\boldsymbol{r}_{\alpha}\right) a_{n'\lambda}\left(\boldsymbol{r}_{\alpha'}\right)\Big].$$

The first term is called the direct term and the second is the exchange term. Using the fermion anticommutation relation

$$\left\{a^{\dagger}_{n\lambda}\left(\boldsymbol{r}_{\alpha}\right), a_{n'\lambda'}\left(\boldsymbol{r}_{\alpha'}\right)\right\} = \delta_{\alpha\alpha'}\delta_{nn'}\delta_{\lambda\lambda'} \tag{2.91}$$

we may write the exchange term as

$$-\frac{1}{2}\sum_{\substack{\alpha,\alpha'\\ n,n'\\ \lambda,\lambda'}} J_{nn'}\left(\boldsymbol{r}_{\alpha},\boldsymbol{r}_{\alpha'}\right) a^{\dagger}_{n\lambda}\left(\boldsymbol{r}_{\alpha}\right) a_{n\lambda'}\left(\boldsymbol{r}_{\alpha}\right) a^{\dagger}_{n'\lambda'}\left(\boldsymbol{r}_{\alpha'}\right) a_{n'\lambda}\left(\boldsymbol{r}_{\alpha'}\right) \tag{2.92}$$

When the spin sum is expanded, we obtain four terms. These may be written in a particularly revealing way by noting the following. First of all, if we allow only one electron to occupy each orbital, then

$$N_{n\uparrow}\left(\boldsymbol{r}_{\alpha}\right) + N_{n\downarrow}\left(\boldsymbol{r}_{\alpha}\right) = 1, \tag{2.93}$$

where $N_{n\lambda}\left(\boldsymbol{r}_{\alpha}\right) = a^{\dagger}_{n\lambda}\left(\boldsymbol{r}_{\alpha}\right) a_{n\lambda}\left(\boldsymbol{r}_{\alpha}\right)$ is the number operator associated with the nth orbital with spin λ at the site α. We also have

$$N_{n\uparrow}\left(\boldsymbol{r}_{\alpha}\right) - N_{n\downarrow}\left(\boldsymbol{r}_{\alpha}\right) = \sigma_{z}\left(\boldsymbol{r}_{\alpha}\right) \tag{2.94}$$

Combining these two relations gives

$$N_{n\uparrow}\left(\boldsymbol{r}_{\alpha}\right) N_{n'\uparrow}\left(\boldsymbol{r}_{\alpha'}\right) + N_{n'\downarrow}\left(\boldsymbol{r}_{\alpha}\right) N_{n'\downarrow}\left(\boldsymbol{r}_{\alpha'}\right) = \frac{1}{2}\sigma_{z}\left(\boldsymbol{r}_{\alpha}\right)\sigma_{z}\left(\boldsymbol{r}_{\alpha'}\right) + \frac{1}{2} \tag{2.95}$$

We also note that

$$a^{\dagger}_{n\uparrow}\left(\boldsymbol{r}_{\alpha}\right) a_{n\downarrow}\left(\boldsymbol{r}_{\alpha}\right) = \frac{1}{2}\sigma^{+}\left(\boldsymbol{r}_{\alpha}\right) \tag{2.96}$$

and

$$a^{\dagger}_{n\downarrow}\left(\boldsymbol{r}_{\alpha}\right) a_{n\uparrow}\left(\boldsymbol{r}_{\alpha}\right) = \frac{1}{2}\sigma^{-}\left(\boldsymbol{r}_{\alpha}\right) \tag{2.97}$$

With these results the exchange interaction (2.92) becomes

$$\sum_{\substack{\alpha\alpha'\\ nn'}} J_{nn'}\left(\boldsymbol{r}_{\alpha},\boldsymbol{r}_{\alpha'}\right)\left[\frac{1}{4} + \frac{1}{4}\boldsymbol{\sigma}\left(\boldsymbol{r}_{\alpha}\right)\cdot\boldsymbol{\sigma}\left(\boldsymbol{r}_{\alpha'}\right)\right]. \tag{2.98}$$

This is the many-electron generalization of our earlier result for two electrons. Since the wave functions used in $J_{nn'}$ are orthogonal, this exchange is always ferromagnetic. In the next section on superexchange we shall see that when we allow for the fact that an electron can hop onto a neighboring site giving it a double occupancy, the result is an exchange of the form (2.98) but with an antiferromagnetic coupling.

In most practical situations we do not have just N electrons each localized on one of N lattice sites, but rather Nh electrons, where h is the number of unpaired electrons on each ion. If these h electrons all have the same exchange integrals with all the other electrons, then the interaction may be expressed in terms of the total ionic spin,

$$\mathcal{H}_{\text{ex}} = -\sum_{\substack{\alpha,\alpha' \\ n,n'}} J_{nn'}\left(\boldsymbol{r}_\alpha, \boldsymbol{r}_{\alpha'}\right)\left[\frac{1}{4} + \boldsymbol{S}\left(\boldsymbol{r}_\alpha\right)\cdot\boldsymbol{S}\left(\boldsymbol{r}_{\alpha'}\right)\right] \tag{2.99}$$

As Van Vleck has pointed out, this form is also valid if the unfilled shells of each ion are half full and the atom is in its state of maximum multiplicity [16]. Such a situation is always true for S-state ions but may also arise as a result of quenching by crystal fields. We shall consider an example of this in our discussion on effective exchange. Equation (2.99), usually referred to as the Heisenberg exchange interaction, often forms the starting point for discussions of ferromagnetism or antiferromagnetism in insulators. The fact that it is valid only under certain conditions does not seem to deter its application. In fact, it works surprisingly well, as we shall see in later chapters.

2.2.8 Superexchange

The exchange constant in (2.99) involves essentially an eigenfunction of the whole crystal. Needless to say, such a function is difficult to obtain. Let us therefore consider an approach which has proven useful in discussing exhange in insulators. The transition-metal flourides MnF_2, FeF_2, and CoF_2 are all observed to be antiferromagnets at low temperatures, with crystal structure and spin configuration indicated in Fig.2.2. This exchange is difficult

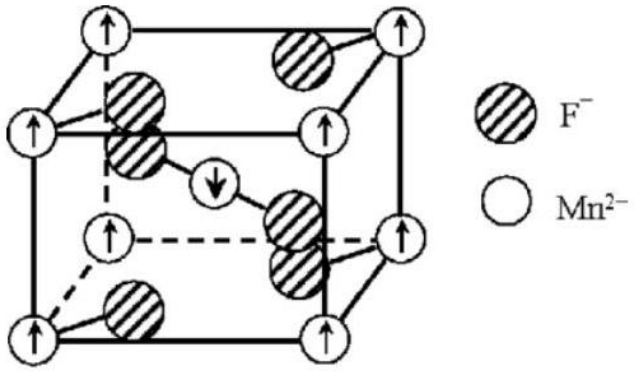


Fig. 2.2. Spin configuration of the transition-metal flourides

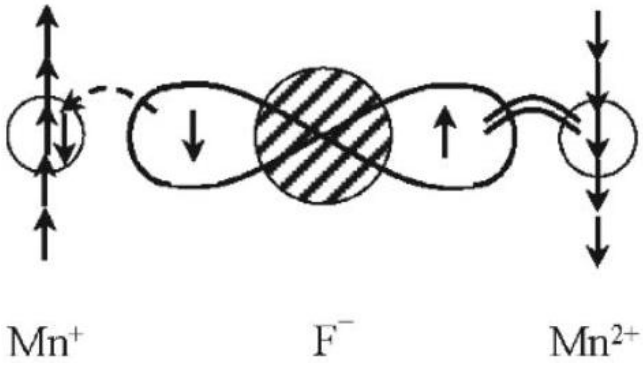


Fig. 2.3. Schematic representations of one of the intermediate states in superexchange

to understand in terms of direct exchange between the cations because of the intervening flourine anions. A similar situation arises in the case of magnetic oxides. In 1934 Kramers proposed the explanation that the cation wave functions were being strongly admixed with the flourine wave functions, enabling the cations to couple indirectly with each other. Kramers applied perturbation theory to obtain the effective exchange resulting from this mechanism. Let us consider two Mn^{2+} ions and an intervening F^{-}ion. Because of the overlap of their wave functions, one of the p electrons from the F^{-}hops over to one of the Mn^{2+} ions. The remaining unpaired p electron on the F then enters into a direct exchange with the other Mn^{2+} ion. This excited state is illustrated in Fig. 2.3 for the case in which the exchange between the unpaired p electron and the Mn^{2+} is antiferromagnetic. By using such excited states in a perturbation calculation of the total energy of the system we obtain an effective exchange between the Mn^{2+} ions. The sign of this exchange depends on the nature of the orbitals involved. However, a number of general features that have evolved through the work of Goodenough [17] and Kanamori [18] enable us to qualitatively predict the nature of the superexchange. Two such features are that the electron transfer can take place only if the cation and anion orbitals are nonorthogonal, and that if the cation-anion orbitals are orthogonal, the direct exchange referred to above is positive (ferromagnetic); otherwise it is negative (antiferromagnetic).

As an example of the application of these rules, let us consider the antiferromagnet $CaMnO_3$. In this material the manganese occurs in the tetravelent state Mn^{4+}, which means that we have three d electrons. The crystal field at
the Mn^{4+} sites is cubic. The d electrons are strongly affected by this crystal field. In fact, the approximation is often made that the crystal field is stronger than the intraionic Coulomb interaction, so that the latter may be neglected. Then each electron may be considered separately. The effect of a cubic field on the fivefold orbitally degenerate d state of a single electron is to split this state into a threefold degenerate state labeled t_{2g} and a twofold degenerate state labeled e_g. This splitting with the associated wave functions is shown in Fig. 2.4. There is, in fact, some intraionic Coulomb interaction, which leads to a Hund's rule coupling of the spins. Therefore, since the t_{2g} state lies lowest for the particular coordination in $CaMnO_3$, the three electrons will each go into one of the t_{2g} orbitals with their spins up.

The superexchange in this case involves the p electrons of the O^{2-}. The p orbitals are illustrated in Fig. 2.5. Examination of the wave functions in Fig. 2.4, shows that the p_z orbital is orthogonal to all the cation orbitals except $d_{x^2-y^2}$. Therefore, if a p_z electron hops over to a Mn^{4+}, it must go into this e_g orbital. Since Hund's rule requires that the total spin be a maximum, it is the up spin from the p_z orbital that transfers if the spins of the Mn^{4+} to which it is going are up. The remaining p_z down spin, since it is orthogonal to e_g orbitals, couples ferromagnetially to the other Mn^{4+}. As a result, we find that the p_z orbital has produced a net antiferromagnetic coupling between the cations themselves. It turns out that the contributions from the p_x orbitals are much smaller.

The Goodenough-Kanamori rules predict that 180° superexchange between electronic configurations $d^3 - d^5$ bridged via an oxygen is ferromagnetic.

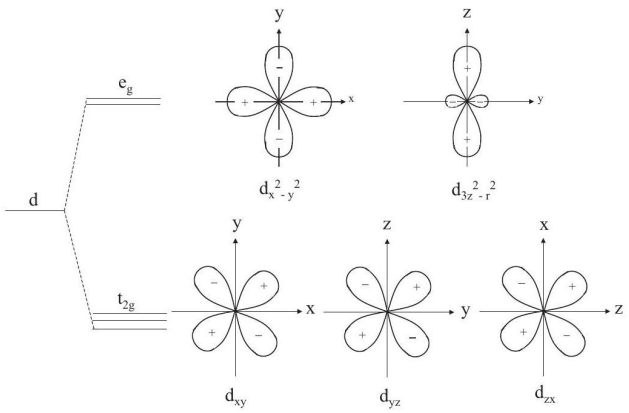


Fig. 2.4. Representations of the eigenvalues and eigenfunctions of a d electron in a cubic crystal field

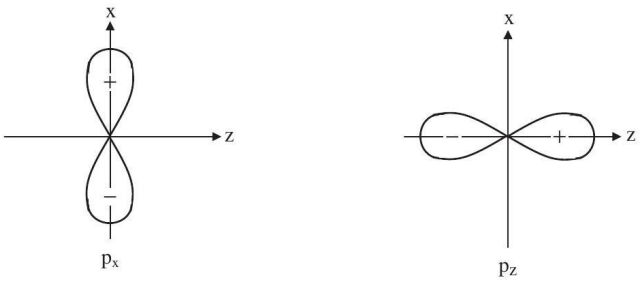


Fig. 2.5. Representations of the p orbitals

An example is Fe^{3+} and Cr^{3+} introduced alternately on the B sites of the perovskites, ABO_3. Attempts to synthesize a ferromagnet with $LaCr_{0.5}Fe_{0.5}O_3$ originally failed because the material phase separates into the antiferromagnets $LaCrO_3$ and $LaFeO_3$. However, with new thin film fabrication techniques, such as laser molecular beam epitaxy, it is possible to deposit alternating monolayers of $LaCrO_3$ and $LaFeO_3$. The resulting superlatttice is indeed a ferromagnet with a Curie temperature of 375 K [19].

Anderson [20] reformulated Kramer's theory in an attempt to avoid the high order perturbation expansion, and his results suggest that antiferromagnetism may be more common than Kramer's theory implies. Anderson worked in a basis of ligand wave functions which are a covalent admixture of cation and anion functions. In Anderson's theory magnetism is the result of the interplay between two effects-the hopping of electrons between ligand complexes, characterized by a hopping matrix element $t_{\alpha\alpha'}$, and an average Coulomb interaction U between electrons on the same complex. In the limit where the hopping may be treated as a perturbation, Anderson found that the super-exchange interaction has the same spin dependence as (2.99) but with a coefficient $-t^2_{\alpha\alpha'}/U$.

In recent years chemists have recast superexchange in terms of molecular orbitals. Hay et al. [21] for example, have shown that the exchange interaction between two spin $\frac{1}{2}$ ions can be expressed as

$$J = J_{ab} - \frac{(\epsilon_1 - \epsilon_2)^2}{K_{aa} - K_{bb}}. \tag{2.100}$$

The exchange J_{ab} involves orthogonalized molecular orbitals and is therefore inherently ferromagnetic. The energies ϵ_1 and ϵ_2 are the bonding and antibonding energies associated

with the molecular orbitals on the two metal ions. The K 's are defined in terms of molecular orbitals analogous to (2.84). Since $K_{aa} > K_{bb}$ the nature of the exchange depends upon the magnitude of $\epsilon_1 - \epsilon_2$. This quantity turns out to be a sensitive function of the metal-ligand-metal angle. Willett [22] and his co-workers have exploited this fact to synthesize a

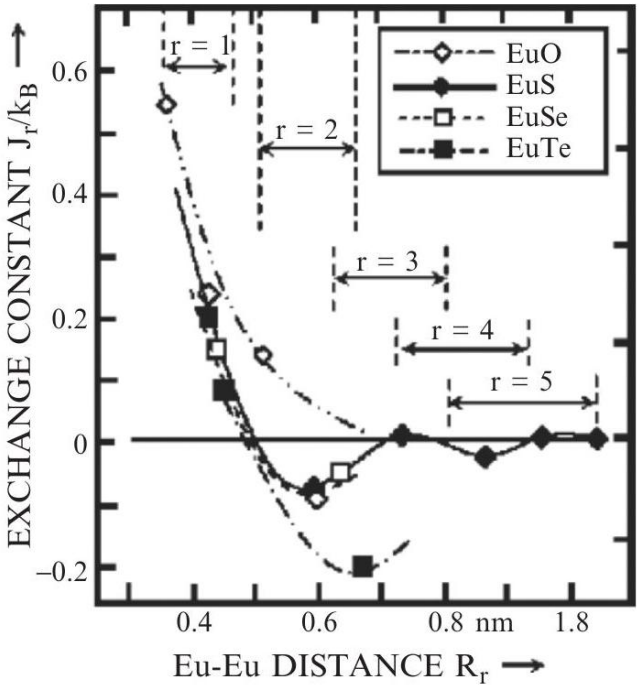


Fig. 2.6. Exchange interactions between Eu-Eu pairs in Eu monochalcogenides versus the pair separation R_{τ}. The horizontal arrows indicate the range of R_τ values from the oxides to the tellurides. Open symbols indicate results deduced assuming only second-nearest-neighbor coupling [23]. Units are degrees K

variety of pseudo-one-dimensional magnetic systems based on copper dimeric species.

In semiconductors the anion energies form bands. As a result the superexchange can extend to distant neighbors as illustrated in Fig. 2.6. These results were deduced from the measured spin-wave dispersion relations just as atomic force constants are deduced from phonon dispersion relations. We shall discuss spin-waves in Chap. 8.

2.2.9 Molecular Magnets

Ordinarily we do not associate magnetism with organic or molecular-based materials. Recently, however, chemists have synthesized materials with interesting magnetic properties. One approach has been to synthesize organic molecules with large numbers of unpaired spins that are in a high-spin state. Examples include a hexacarbene with $S = 6$ and a nonocarbene with $S = 9$. Incorporating these molecules into solids is still a challenge.

A more successful approach has been the synthesis of organometallic solids comprised of linear chains of alternating metallocenium donors and cyanocarbon acceptors. One such system is $[FeCp_2^*]^+ [TCNE]^-$based on decamethyferrocene ($Cp^* = C_5 (CH_3)_5$) and tetracyanoethylene, shown in Fig. 2.7.

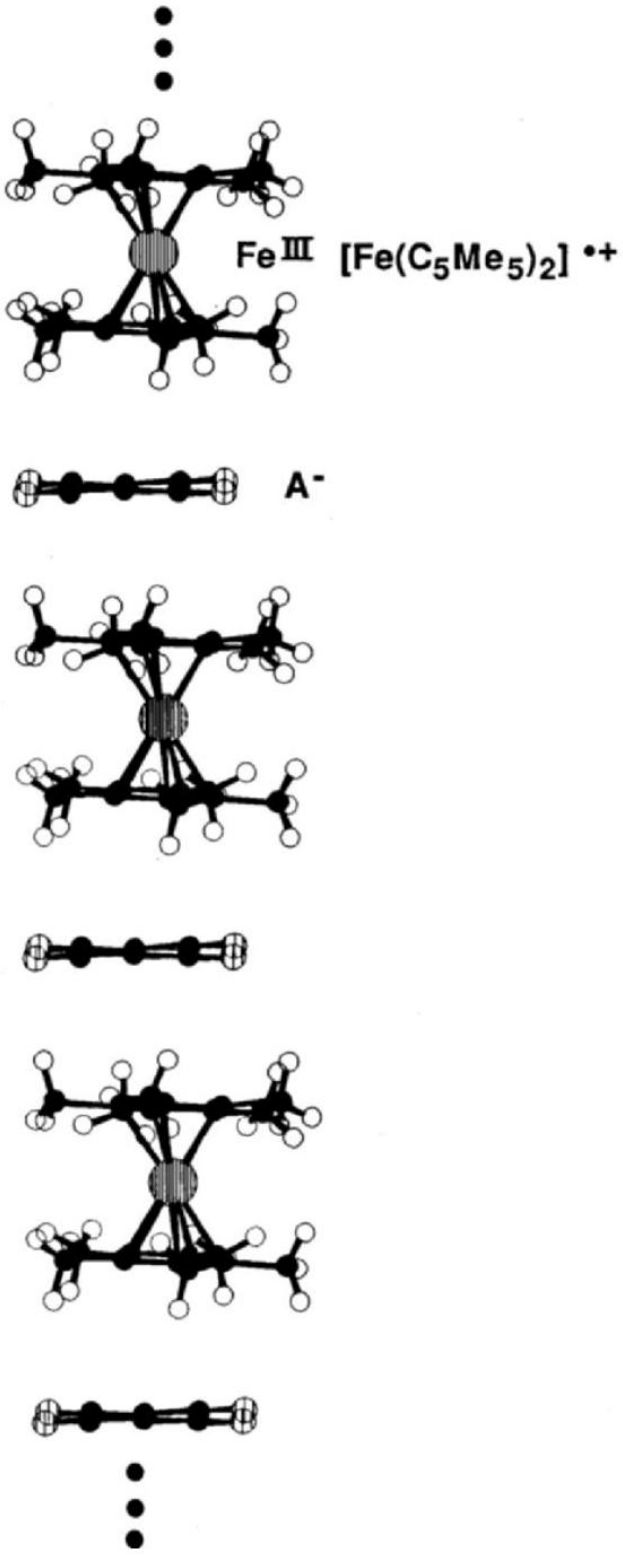
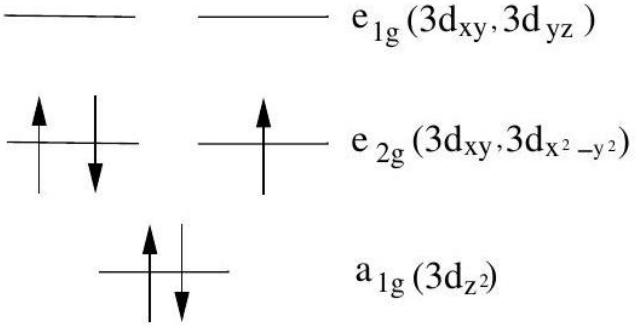


Fig. 2.7. Alternating donor/acceptor linear chain structure observed for $\left[F e C p_{2 {*}\right]{+}[T C N E]^{-}$. Open circles = H, solid circles = C, small hatched circles = N, and larged hatched circles = Fe [J. Miller]}

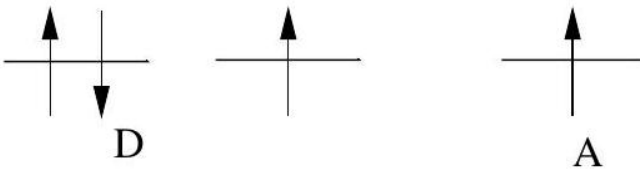
There are many variations in which the iron is replaced by other transition metals ions and the radical anion is replaced by other radical anions such as TCNQ (tetracyano-p-quinodimethanide) or DDQ (dichloro-5,6dicyanobenzoquinone). As their long names suggest, the magnetic species are widely separated which means that the

magnetization, or magnetic moment per unit volume is small. The magnetization of $[FeCp_2^*]^+ [TCNE]^-$
is 47emu/cm^3 which is to be compared with that of iron of 1715emu/cm^3. The large separation between cations is also responsible for relatively weak exchange interactions as reflected in low ordering temperatures. The Curie temperature of our example Fig. 2.7 is 4.8 K . It may be surprising that these materials show any coupling, considering that the transition metal ion sits in the center of a relatively large molecule. The point is, however, that the wavefunctions of these molecules are molecular orbitals with charge density and spin density that extends throughout the molecule. These molecular orbitals have been calculated for the similar, but simpler, molecule, ferrocene, $Fe(C_5H_5)_2$, usually written $Fe(Cp)_2$. The structure consists of two planar C_5H_5 rings with the iron sandwiched in between. There are low-lying bonding states associated with the C_5 rings. The states involving the Fe^{III} are:

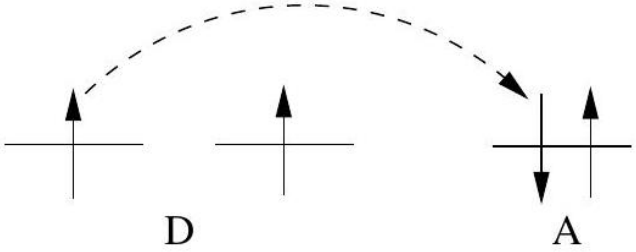


Fe^{III} has the low spin d^5 configuration. These five electrons fill the orbitals as shown, giving a molecular spin of 1/2. The TCNE $^-$also has one unpaired electron.

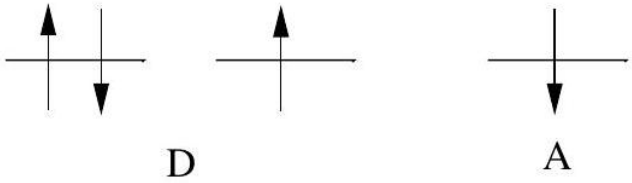
There is some overlap of the cation and anion orbitals that gives rise to an exchange interaction. McConnell [24] was the first to point out that such ionic crystals could have ferromagnetic exchange. The mechanism McConnell proposed is based on the virtual charge transfer. Suppose, for example, the anion (A = acceptor) and cation (D = donor) spins are ferromagnetically aligned:



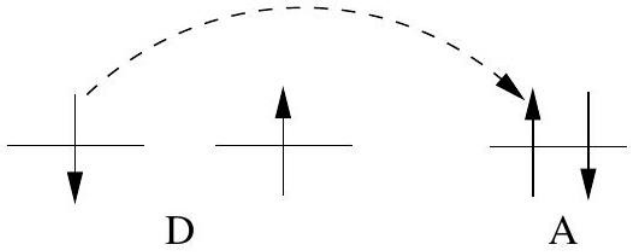
Then the charge transfer $D^+A^- \rightarrow D^{+2}A^{-2}$ leads to an intermediate state, in which there are two electrons in orthogonal orbitals with parallel spin on the cation:



Had we started with an antiferromagnetic alignment:



then the intermediate state would have antiparallel spins on the cation:



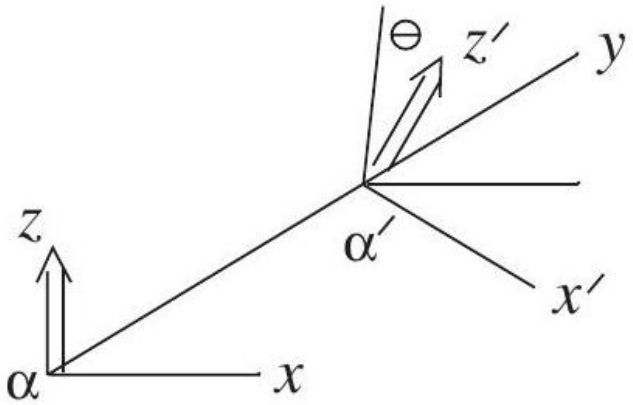
The parallel configuration has the lower energy (Hund's rule). Therefore the charge transfer favors the ferromagnetic state.

The unpaired electron on the anion could also transfer back to the cation (i.e., $D^+A \rightarrow D^\circ A^\circ$). This favors antiferromagnetic coupling. Consistant with the observed data the exchange is governed more by the $D \rightarrow A$ transfer. There are also various interchain exchange interactions that could influence the nature of the 3 -dimensional ordering. However, it appears that the ordering in these salts is governed by the intrachain exchange.

Much of the recent work on these molecular magnets has been done by Miller and Epstein [25] who have also discovered a polymeric magnet involving vanadium with a Curie temperature about 400 K .

2.2.10 Double Exchange

In Anderson's formulation of superexchange, the electron hopping is a virtual process. If, however, we actually have carriers in the magnetic material, then the hopping matrix element $t_{\alpha\alpha'}$ enters directly into the band energy in a tight-binding approximation. If we assume that the intra-atomic exchange is so strong that the spin of a carrier is parallel to the local ionic spin then the band energy is dependent upon the ionic spin configuration. This correlation between the magnetism and the conductivity was first addressed by Zener [26] and subsequently in more detail by Anderson and Hasegawa [27]. To determine how the magnetic order affects a conduction electron, we recall our discussion of how the spinor part of the electron wavefunction transforms under rotations. Let us consider the following geometry:



The ionic spin is assumed to be rotated by an angle θ about the y-axis. If the spinor at site α is $\binom{1}{0}$, then at site α', it is given by the matrix for a rotation about the y-axis,

$$\begin{pmatrix} \cos\frac{\theta}{2} & \sin\frac{\theta}{2} \\ -\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\frac{\theta}{2} \\ -\sin\frac{\theta}{2} \end{pmatrix}.$$

Since we assume electron hopping can only occur if the electron hops into the $\binom{1}{0}$ state, the orbital transfer integral $b_{\alpha\alpha'}$ is multiplied by $\cos\frac{\theta}{2}$. Thus,

$$t_{\alpha\alpha'} = b_{\alpha\alpha'} \cos\frac{\theta}{2}. \tag{2.101}$$

If the ionic spins have an exchange, J, then the tipping between α and α' gives an exchange energy proportional to $J\cos\theta$. DeGennes has shown [28] that minimizing the kinetic energy associated with the hopping and the exchange energy can lead to a canted spin structure. Distortion of the spin lattice around the hopping electron results in a magnetic polaron. This is discussed in Sect. 10.4.

2.2.11 Exchange on a Surface

Recent developments in nanotechnology enable one to experimentally study the exchange interaction between individual atoms. Figure 2.8 shows the tip on a scanning tunneling microscope which is used to position Mn atoms next to one another (0.36 nm apart) on an insulating surface. The tip can also be

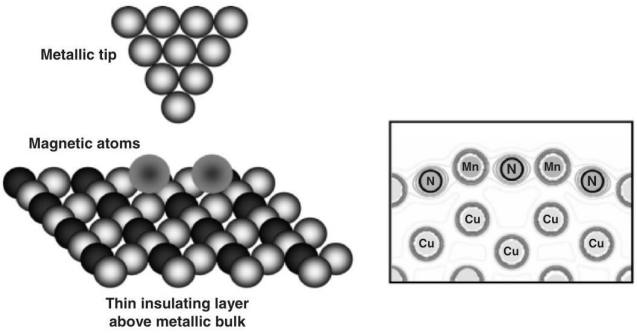


Fig. 2.8. CuN surface showing how a scanning tunneling microscope can be used to position two manganese atoms next to one another (C.F. Hirjibehedin)

used to measure the conductance of these artificially created Mn clusters. The inelastic electron tunneling spectra for a dimer shows structure that has been indentified as an excitation between an antiparallel ground state involving two spin 5/2 moments and a parallel excited state ref [29]. The voltage at which this structure occurs enables a determination of the exchange interaction between the two Mn moments. This has the values 2.7 meV or 6.4 meV at 0.6 K depending upon where the dimer sits on the CuN surface. The insert in Fig. 2.8 shows the cross-section of the charge density as calculated from a "first-principles" calculation. The Mn 3d electrons couple to this surrounding charge density through an "s-d" exchange interaction, $J_{sd}\mathbf{S}$.s. This provides an effective exchange between the two Mn moments of the order of J_{sd}^2/E_g, where E_g is the bandgap of the CuN . Since J_{sd} is of order 0.1 eV and E_g is several eV the $\mathrm{Mn}-\mathrm{Mn}$ exchange is of order of several meV as observed.

10.3.3 2.3 The Spin Hamiltonian

The Hamiltonian developed in the preceding section is completely general, and a knowledge of its eigenvalues would accurately describe the magnetic properties of any material. Unfortunately, because of the large number of particles involved, such a knowledge is beyond us at this time. Therefore we try to project out of the Hamiltonian those terms which adequately describe the situation and yet are amenable to calculation. Experimentalists in particular often propose "phenomenological" Hamiltonians to explain certain observations, leaving to theoreticians the job of establishing the legitimacy of such forms. In the remainder of this chapter we shall indicate the origin of such phenomenological Hamiltonians.

2.3.1 Transition-Metal Ions

The first row of transition-metal ions and their electronic configurations are listed in Table 2.4. The important feature about transition-metal ions is that the magnetic, or unpaired, electrons lie in the outermost shell of the ion. Therefore they are easily influenced by any external field produced by neighboring ligands. That is, the crystal field is likely to be one of the largest terms in the Hamiltonian. Thus we might expect that the contributions to the Hamiltonian, in order of descending strength, are

$$\mathcal{H} = \mathcal{H}_{\text{intraatomic Coulomb}} + \mathcal{H}_{\text{crystal field}} + \mathcal{H}_{\text{spin-orbit}} + \mathcal{H}_{\text{Zeeman}} . \tag{2.102}$$

Of course, depending on the situation, we may have to consider additional terms, such as the hyperfine interaction. However, let us consider the eigenstates of the Hamiltonian (2.102). First of all, the intraatomic Coulomb interaction leads to spectroscopic energy levels, the lowest of which is determined by Hund's rules. This ground state is indicated in the last column of Table 2.4.

For most magnetic properties it is sufficient to consider only the lowest term. This is because term energies are of the order of tens of thousands of wave numbers, whereas magnetic energies are at most tens of wave numbers.

The behavior of a given term in the crystal field may be calculated by the technique developed in the last section. Cubic symmetry is the predominant symmetry encountered in most crystals, and Fig. 2.9 indicates how group theory predicts splitting of the various terms of Table 2.4 in such a case. The eigenfunctions of $\mathcal{H}_{\text{intra}} + \mathcal{H}_{\text{cryst}}$ will be denoted as $|\Gamma, \gamma; S, M_S\rangle$, where Γ is the irreducible representation of the point-group symmetry.

Consider, for example, the $3d^3$ configuration. From Hund's rules, the ground state is

$$|L, M_L; S, M_S\rangle = \left|3, M_L; \frac{3}{2}, M_S\right\rangle , \tag{2.103}$$

which is $(2L+1)(2S+1) = 28$-fold degenerate. In the presence of a cubic crystal field this state splits as shown in Fig. 2.9. Thus the ground state would be denoted as $\left|A_2, \gamma; \frac{3}{2}, M_S\right\rangle$.
The g Tensor. Now consider the spin-orbit and Zeeman terms. Since we are considering matrix elements only within a given LS term, the matrix elements of

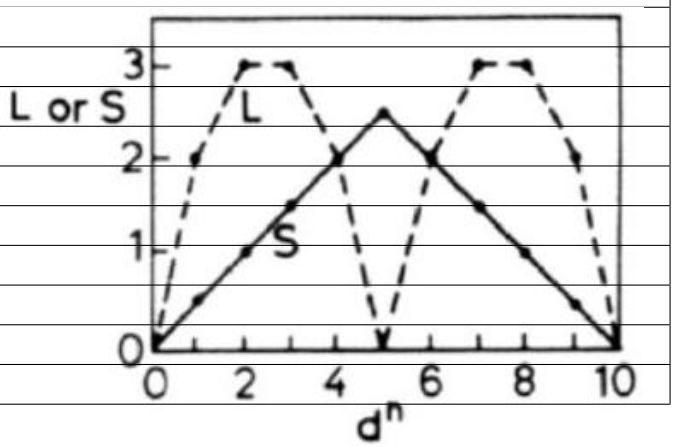
$$\sum_i \xi\left(\boldsymbol{r}_i\right) \boldsymbol{l}_i \cdot \boldsymbol{s}_i$$

are proportional to those of $\boldsymbol{L} \cdot \boldsymbol{S}$, by the Wigner-Eckhart theorem. Thus the spin-orbit Hamiltonian may be written

$$\mathcal{H}_{\text{sp-orb}} = \lambda \boldsymbol{L} \cdot \boldsymbol{S}, \tag{2.104}$$

Table 2.4. Configuration of the iron-group ions

Ti^{3+}, V^{4+}	$3d^1$	^2D	
V^{3+}	$3d^2$	^3F	
Cr^{3+}, V^{2+}	$3d^3$	^4F	
Mn^{3+}, Cr^{2+}	$3d^4$	^5D	
Fe^{3+}, Mn^{2+}	$3d^5$	^{6} S	
Fe^{2+}	$3d^6$	^5D	
Co^{2+}	$3d^7n$ ^5D		
Co^{2+}	$3d^7n$ ^5D		
Co^{2+}	$3d^7n$ ^4F		
Ni^{2+}	$3d^8$	^3F	
Cu^{2+}	$3d^9$	^2D	



where λ is the spin-orbit parameter. Similarly, the Zeeman Hamiltonian is

$$\mathcal{H}_Z = \mu_B(\boldsymbol{L} + 2\boldsymbol{S}) \cdot \boldsymbol{H}. \tag{2.105}$$

We now transform $\mathcal{H}_{\text{sp-orb}} + \mathcal{H}_Z$ into the so-called spin Hamiltonian by a method proposed by Pryce, in which we project out the orbital dependence. Since neither $\mathcal{H}_{\text{intra}}$ nor $\mathcal{H}_{\text{cryst}}$ has mixed orbital and spin states, our eigenfunctions are products of the form $|\Gamma, \gamma\rangle\,|S, M_S\rangle$. Let us evaluate the expectation value of $\mathcal{H}_{\text{sp-orb}} + \mathcal{H}_Z$ for an orbitally nondegenerate ground state $|\Gamma, \gamma\rangle$. To second order in perturbation theory,

$$\begin{aligned}\mathcal{H}_{\text{eff}} &= \langle\Gamma, \gamma|\mathcal{H}_{\text{sp-orb}} + \mathcal{H}_Z|\Gamma, \gamma\rangle \\ &= 2\mu_B \boldsymbol{H} \cdot \boldsymbol{S} - \sum_{\Gamma', \gamma'} \frac{|\langle\Gamma', \gamma'|\,\mu_B \boldsymbol{H} \cdot \boldsymbol{L} + \lambda \boldsymbol{L} \cdot \boldsymbol{S}|\,\Gamma, \gamma\rangle\,|^2}{E_{\Gamma', \gamma'} - E_{\Gamma, \gamma}}\end{aligned} \tag{2.106}$$

Expanding the square gives

$$\mathcal{H}_{\text{eff}} = 2\mu \boldsymbol{H} \cdot \boldsymbol{S} - 2\mu_B \lambda \sum_{\mu,\nu} \Lambda_{\mu\nu} S_\mu H_\nu - \lambda^2 \sum_{\mu,\nu} \Lambda_{\mu\nu} S_\mu S_\nu - \mu_B^2 \sum_{\mu,\nu} H_\mu H_\nu, \tag{2.107}$$

where

$$\Lambda_{\mu\nu} = \sum_{\Gamma'\gamma'} \frac{\langle\Gamma, \gamma|L_\mu\,|\Gamma', \gamma'\rangle\,\langle\Gamma', \gamma'|\,L_\nu|\Gamma, \gamma\rangle}{E_{\Gamma', \gamma'} - E_{\Gamma, \gamma}} \tag{2.108}$$

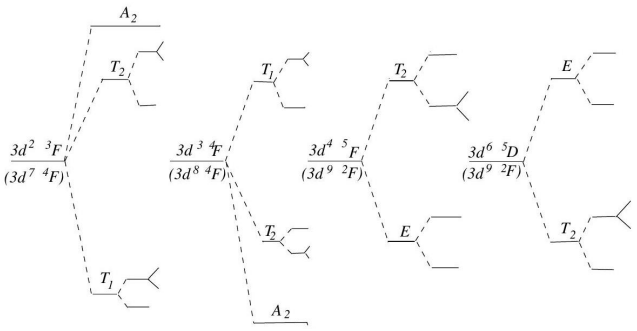


Fig. 2.9. Crystal-field splittings of the iron-group ion ground states. The first set of splittings are the result of a cubic crystal field. Subsequent splittings are due to an additional tetragonal distortion

This may be written as

$$\mathcal{H}_{\text{eff}} = \sum_{\mu,\nu} \left(\mu_B g_{\mu\nu} H_\mu S_\nu - \lambda^2 A_{\mu\nu} S_\mu S_\nu - \mu_B^2 \Lambda_{\mu\nu} H_\mu H_\nu \right), \tag{2.109}$$

where $g_{\mu\nu}$ is the g tensor

$$g_{\mu\nu} = 2\left(\delta_{\mu\nu} - \lambda \Lambda_{\mu\nu}\right). \tag{2.110}$$

The fact that $g_{\mu\mu}$ differs from 2 tells us that owing to the spin-orbit interaction, the magnetization is now no longer spin only. That is, a small amount of orbital angular momentum has been admixed back into the ground state.

The g value of the free electron is not precisely 2. There are quantum electrodynamic corrections which lead to the value $g = 2.002319$.

As an example of how this value is changed in a solid let us consider the semiconductor silicon. In the ideal crystalline state each silicon atom has four valence electrons which form bonds with four neighboring silicon atoms. If this ideal crystal is irradiated, with high-energy electrons, for example, defects are introduced. One of these is the divacancy illustrated in Fig. 2.10a. The removal of the two atoms A and B leaves six broken bonds, or unpaired electrons. Those electrons associated with atoms 2 and 3 and those associated with 5 and 6 reconstruct bonds as indicated in the figure. The remaining two electrons from atoms 1 and 4 then form an "extended" bond across the vacancy. If one of these latter two electrons is removed by some means, leaving the defect with a net positive charge, the remaining electron is found to have the g value [30]

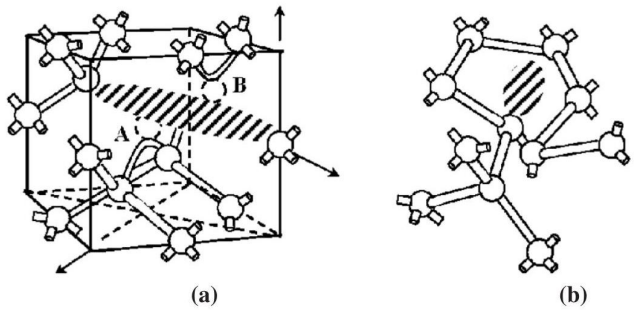


Fig. 2.10. Defects in (a) crystalline and (b) amorphous silicon having nearly free electron g values

$$g_1 = 2.0004$$
$$g_2 = 2.0020$$
$$g_3 = 2.0041$$

If the silicon is prepared as an amorphous film, by the decomposition of silane gas, for example, then the atomic disorder sometimes leaves a silicon atom with only three other silicon neighbors as illustrated schematically in Fig. 2.10b. The fourth unpaired electron is referred to as a "dangling" bond. This electron is found to have the average g value [31]

$$g = 2.0055$$

Although these g values differ by less than 1% these differences are easily measured, as we shall discuss in Chap. 5, and provide an extremely important characterization of the electronic center.

Anisotropy. The second term in (2.109) represents the fine-structure or singleion anisotropy. Notice that $\wedge_{\mu\nu}$ reflects the symmetry of the crystal. The spin Hamiltonian must also display this symmetry; for example, in a cubic crystal $\wedge_{xx} = \wedge_{yy} = \wedge_{zz}$. Thus the anisotropy term reduces to a constant. For axial symmetry $\wedge_{xx} = \wedge_{yy} = \wedge_{\perp}$ and $\wedge_{zz} = \wedge_{\|}$. Thus, if we neglect the last term, the effective axial Hamiltonian is

$$\begin{aligned}\mathcal{H}_{\text{eff}} =& g_{\|}\mu_B H_z S_z + g_{\perp}\mu_B \left(H_x S_x + H_y S_y\right) + D\left[S_z^2 - \frac{1}{3}S(S+1)\right] \\ &+ \frac{1}{3}S(S+1)\left(2\wedge_{\perp} + \Lambda_{\|}\right)\lambda^2 \end{aligned} \tag{2.111}$$

where $D = \lambda^2\left(\wedge_{\|} - \wedge_{\perp}\right)$. A Hamiltonian of this form based on the symmetry of the crystal is usually taken as the starting point in describing paramagnetic systems involving transition-metal ions. Thus such ions in crystals are characterized by their g, D, etc., parameters.

An important experimental fact is that these crystal-field parameters do not change appreciably for concentrated versions of the same salt. Hence, we shall find these paramagnetic states useful when we discuss ferromagnetism. For example, we can see from the discussion above that the anisotropy constants depend on the energy levels, which in turn depend on the positions of the neighboring ions. Thus, if these ions move because of the presence, say, of phonons, then we have a coupling between this motion and the spins.

The last term in the Hamiltonian (2.109) will survive the two field derivatives leading to the static susceptibility, resulting in the so-called Van Vleck susceptibility, which is temperature independent.

As an example of the application of the spin Hamiltonian, let us consider a spin $\frac{3}{2}$ in an axially symmetric system with an external field applied along the c axis. If we assume that the crystal field is of sufficiently low symmetry

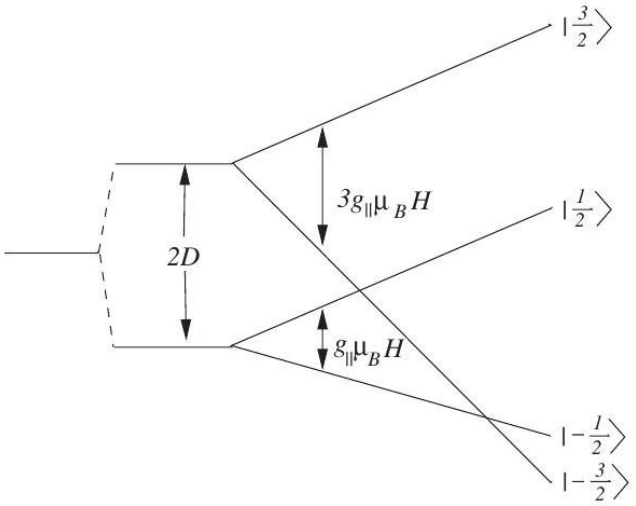


Fig. 2.11. Representation of the effect of the spin-orbit and Zeeman interactions on an orbital singlet with spin $\frac{3
{2}$

to remove any orbital degeneracy of the ground state, then (2.111) applies. Dropping the constant part, we have

$$\mathcal{H}_{\text{eff}} = g_{\|}\mu_B H S_z + D\left[S_z^2 - \frac{1}{3}S(S+1)\right]. \tag{2.112}$$

The matrix of $\mathcal{H}_{\text{eff}}$ in the basis $\left|\frac{3}{2}, M_S\right\rangle$ is

$$\mathcal{H}_{\text{eff}} = \begin{matrix} \left\langle -\frac{3}{2}\right| \\ \left\langle -\frac{1}{2}\right| \\ \left\langle \frac{1}{2}\right| \\ \left\langle \frac{3}{2}\right| \end{matrix} \begin{bmatrix} D - \frac{3}{2}g_{||}\mu_B H & \left|-\frac{1}{2}\right\rangle & \left|\frac{1}{2}\right\rangle & \left|\frac{3}{2}\right\rangle \\ 0 & -D - \frac{1}{2}g_{||}\mu_B H & 0 & 0 \\ 0 & 0 & -D + \frac{1}{2}g_{||}\mu_B H & 0 \\ 0 & 0 & 0 & D + \frac{3}{2}g_{||}\mu_B H \end{bmatrix}. \tag{2.113}$$

The eigenvalues are shown in Fig. 2.11. We see that there is a zero-field splitting of $2D$.

Effective Exchange. In our discussion of exchange in the last section we neglected spin-orbit effects as well as off-diagonal exchange effects. We found above that the spin-orbit interaction leads to important contributions to the spin Hamiltonian. It can also modify the form of the exchange interaction. To illustrate these effects, let us consider the exchange between a Mn^{2+} impurity and a neighboring Co^{2+} ion in $CoCl_2 \cdot 2H_2O$[32].Co^{2+} has seven d electrons which correspond to three holes in the d shell. These holes behave just as electrons would. Therefore we find that the intraatomic Coulomb interaction leads to a ${}^4F$ ground state with a spin of $\frac{3}{2}$. The crystal field in $CoCl_2 \cdot 2H_2O$ is predominantly cubic, with a small tetragonal distortion. The cubic part splits the ${}^4F$ into three states, the lowest being ${}^4T_1$, which has a threefold orbital degeneracy. As mentioned above, crystal-field effects on the transitionmetal ions are rather large. Therefore, in so far as we are concerned only with magnetic properties, we may restrict our considerations to this ${}^4T_1$ ground state.

We now investigate how the tetragonal crystal-field component and the spin-orbit interaction affect this level. Any threefold orbitally degenerate state behaves as though it had an effective quantum number $L = 1$. We saw earlier that if we restrict ourselves to a manifold of states in which, say, L is a good quantum number, then we may express the crystal-field Hamiltonian in powers of L_x, L_y, and L_z. The same argument applies even if L is only a good effective quantum number. Thus the effect of a tetragonal crystal field upon our ${}^4T_1$ ground state may be described by the Hamiltonian

$$\mathcal{H}_{\text{tetra}} = -\delta\left(L_z^2 - \frac{3}{2}\right) \tag{2.114}$$

where δ is a phenomenological crystal-field parameter and $L = 1$. The spinorbit interaction may also be expressed in terms of the effective orbital angular momentum,

$$\mathcal{H}_{\text{sp- orb}} = \lambda' \boldsymbol{L} \cdot \boldsymbol{S}_{\text{Co}} \tag{2.115}$$

where λ' is an effective spin-orbit parameter. The operators $L^+$and $L^-$are the raising and lowering operators for the components of the ${}^4T_1$, just as the corresponding real operators would connect the components of a real P state. This technique of writing Hamiltonians in terms of effective-angular-momentum operators is an extremely useful tool for understanding the qualitative features of the electromagnetic absorption spectra of magnetic systems.

The basis functions for our ${}^4T_1$ state have the form $|M_L, M_S\rangle$, where $M_L = -1, 0,$ or $+1$, and $M_S = \pm\frac{1}{2}$ or $\pm\frac{3}{2}$. The effect of $\mathcal{H}_{\text{tetra}} + \mathcal{H}_{\text{sp-orb}}$ is to split the ${}^4T_1$ into six doublets. Again, for magnetic considerations, we consider the lowest of these. The eigenfunctions have the form

$$\psi^{\pm} = a\left|\mp 1, \pm\frac{3}{2}\right\rangle + b\left|0, \pm\frac{1}{2}\right\rangle + c\left|\pm 1, \mp\frac{1}{2}\right\rangle \tag{2.116}$$

where $a, b,$ and c are certain mixing coefficients which may, in general, be complex. We now inquire how the exchange with the neighboring Mn^{2+} affects this doublet. Since it is a doublet, the exchange should be expressible as an effective spin of $\frac{1}{2}$ interacting with the Mn^{2+} spin. The purpose of this example is to show that when the exchange is in fact expressed in terms of an effective spin of $\frac{1}{2}$, it does not have the simple isotropic form we have so far been using.

The first thing we must do is establish an expression for the exchange which we know to be correct. Recall the the ${}^4T_1$ state has a threefold orbital degeneracy, and that the three d holes can be treated just as electrons. Therefore we have a situation in which there is only one state of maximum multiplicity, which means that the total spin is a good quantum number. Also, since
the ground state of Mn^{2+} is an S state, the total spin of this ion is a good quantum number. Therefore the exchange interaction may be written as

$$\mathcal{H}_{\text{ex}} = -\sum_{M_L M_L'} J\left(M_L, M_L'\right) \boldsymbol{S}_{\text{Mn}} \cdot \boldsymbol{S}_{\text{Co}} \tag{2.117}$$

Notice that we are explicitly including the possibility of off-diagonal exchange. Thus, in terms of actual electron interchange, the exchange integral $J\left(M_L, M_L'\right)$ would characterize the interchange in which an electron from a Co^{2+} in the state M_L jumps over to the Mn^{2+}, while an electron from the Mn^{2+} jumps back to the Co^{2+} and in so doing changes it to the M_L' state. These offdiagonal exchange integrals may be complex. As long as the orbital states so connected lie within the ${}^4T_1$ manifold, the scalar spin product may be used even to describe this off-diagonal exchange.

We now wish to know what form this exchange takes when expressed in terms of the effective spin $\frac{1}{2}$ of the Co^{2+} ground state. The most general interaction has the form

$$\mathcal{H}_{\text{eff}} = -\boldsymbol{S}_{\text{eff}} \cdot \mathcal{J} \cdot \boldsymbol{S}_{\text{Mn}}, \tag{2.118}$$

where $\mathcal{J}$ is an exchange dyadic. If we evaluate the matrix elements of (2.117) in the basis $|\psi^{\pm}, M_S\rangle$, where M_S is the Mn^{2+} spin quantum number ($S = \frac{5}{2}$), and compare them with the matrix elements of (2.118) evaluated in the basis $\left|\pm\frac{1}{2}, M_S\right\rangle$, we find that all the elements of $\mathcal{J}_{\mu\nu}$ are nonzero. For example,

$$\mathcal{J}_{xy} = -2c^2\,\mathrm{Im}\{J(1,-1)\} \tag{2.119}$$

The general matrix $\mathcal{J}_{\mu\nu}$ may be separated into a symmetric and an antisymmetric part. The antisymmetric exchange may be written as

$$\boldsymbol{D}\cdot\boldsymbol{S}_{\mathrm{eff}}\times\boldsymbol{S}_{\mathrm{Mn}} \tag{2.120}$$

where the vector coupling coefficient $\boldsymbol{D}$ is related to $\mathcal{J}_{\mu\nu}$. This exchange is also referred to as the Dzialoshinski-Moriya exchange [33]. If the elements of the symmetric exchange are different, we speak of this as anisotropic exchange. There are two limiting forms that are popularly used: the Ising model

$$\mathcal{H}_{\mathrm{Ising}} = -2J\sum_{i,\delta} S_i^z S_{i+\delta}^z$$

and the XY model

$$\mathcal{H}_{XY} = -2J\sum_{i,\delta}\left(S_i^x, S_{i+\delta}^x + S_i^y S_{i+\delta}^y\right)$$

Here δ refers to a nearest neighbor. Thus we see that when the exchange interaction is expressed in terms of the effective spin of the ground state, it may have a very general form owing to the presence of orbital effects.

Table 2.5. Configurations of the rare-earth ions

Ce^{3+}	$4f^1 5s^2 p^6$	$^2F_{5/2}$		
Pr^{3+}	$4f^2 5s^2 p^6$	$^3H_4$		
Nd^{3+}	$4f^3 5s^2 p^6$	$^4I_{9/2}$		
Pm^{3+}	$4f^4 5s^2 p^6$	$^6I_4$	L	or S
Sm^{3+}	$4f^5 5s^2 p^6$	$^6H_{5/2}$		
Eu^{3+}	$4f^6 5s^2 p^6$	$^7F_0$		
Gd^{3+}	$4f^7 5s^2 p^6$	$^8S_{7/2}$		
Tb^{3+}	$4f^8 5s^2 p^6$	$^7F_6$		
Dy^{3+}	$4f^9 5s^2 p^6$	$^6H_{15/2}$		
Ho^{3+}	$4f^{10} 5s^2 p^6$	$^5I_8$		
Er^{3+}	$4f^{11} 5s^2 p^6$	$^4I_{15/2}$		
Tm^{3+}	$4f^{12} 5s^2 p^6$	$^3H_6$		
Yb^{3+}	$4f^{13} 5s^2 p^6$	$^2F_{7/2}$		

2.3.2 Rare-Earth Ions

The electronic configurations of the rare-earth ions are listed in Table 2.5. For these ions we see that the unpaired electrons lie inside the $5s^2p^6$ shells. Consequently they are not very strongly affected by crystal fields, and we might expect the Hamiltonian to consist of the following terms, in order of descending strength:

$$\mathcal{H} = \mathcal{H}_{\text{intraatomic Coulomb}} + \mathcal{H}_{\text{spin-orbit}} + \mathcal{H}_{\text{crystal field}} + \mathcal{H}_{\text{Zeeman}}\,. \tag{2.121}$$

The intraatomic Coulomb interaction produces states characterized by L, M_L, S, and M_S. When the spin-orbit interaction is added, only the total angular momentum $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$ is conserved. Therefore the states have the form $|J, M_J; L, S\rangle$.

Let us consider, for the moment, the effect of the Zeeman term in the absence of any crystal field,

$$\mathcal{H}_Z = \mu_B(\boldsymbol{L} + 2\boldsymbol{S}) \cdot \boldsymbol{H}. \tag{2.122}$$

Since the states are characterized by the eigenvalues of $\boldsymbol{J}$, we rewrite this as

$$\mathcal{H}_{\text{Zeeman}} = g_J \mu_B \boldsymbol{J} \cdot \boldsymbol{H}, \tag{2.123}$$

where g_J is the Landé g value, defined by $g_J \boldsymbol{J} = \boldsymbol{L} + 2\boldsymbol{S}$. This has the value

$$g_J = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}. \tag{2.124}$$

Thus the state $|J, M_J\rangle$ is split into $2J+1$ equally spaced states with the separation $g_J \mu_B H$. For Ce^{3+}, for example, $g_J = \frac{6}{7}$.
In the presence of a crystal field the splitting of the state $|J, M_J\rangle$ is easily determined by expressing the crystal field in operator equivalents of $\boldsymbol{J}$. Thus for a crystal field of D_2 symmetry we have

$$\begin{aligned}\mathcal{H}_{\text{cryst}} =& B_2^0 O_2^0 + B_2^2 O_2^2 + B_4^0 O_4^0 + B_4^2 O_4^2 + B_4^4 O_4^4 \\ &+ B_6^0 O_6^0 + B_6^2 O_6^2 + B_6^4 O_6^4 + B_6^6 O_6^6\end{aligned} \tag{2.125}$$

where the operators 0_n^m were defined in (2.52).
The matrix of $\mathcal{H}_{\text{cryst}}$ in the basis $|J, M_J; L, S\rangle$ is easily constructed from tables in [12]. The result for Ce^{3+} is

$$\begin{array}{c} \left|\frac{5}{2}\right\rangle \\ \left\langle\frac{5}{2}\right| \\ \left\langle\frac{1}{2}\right| \\ \left\langle-\frac{3}{2}\right| \\ \left\langle-\frac{5}{2}\right| \\ \left\langle-\frac{1}{2}\right| \\ \left\langle\frac{3}{2}\right| \end{array} \tag{2.126}$$

$$\begin{array}{c} \\ \left|-\frac{3}{2}\right\rangle \\ \\ \\ \\ \\ \end{array} \begin{array}{|c} \\ \left|-\frac{5}{2}\right\rangle \quad \left|-\frac{1}{2}\right\rangle \quad \left|\frac{3}{2}\right\rangle \\ \hline \left[\begin{array}{ccc} B & 0 & \\ & & \\ & B & \\ & & \end{array}\right] \end{array}$$

where

$$B \equiv \begin{bmatrix} 10B_2^0 + 60_4^B & \sqrt{10}B_2^2 + 9\sqrt{10}B_4^2 & 12\sqrt{5}B_4^4 \\ \sqrt{10}B_2^2 + 9\sqrt{10}B_4^2 & -8B_2^0 + 120B_4^0 & 3\sqrt{2}B_2^2 - 15\sqrt{2}B_4^2 \\ 12\sqrt{5}B_4^4 & 3\sqrt{2}B_2^2 - 15\sqrt{2}B_4^2 & -2B_2^0 - 180B_4^0 \end{bmatrix}.$$

This gives three doublets which have the form

$$\psi_n^+ = a_n \left|\frac{5}{2}\right\rangle + b_n \left|\frac{1}{2}\right\rangle + c_n \left|-\frac{3}{2}\right\rangle \tag{2.127a}$$

$$\psi_n^- = a_n \left|-\frac{5}{2}\right\rangle + b_n \left|-\frac{1}{2}\right\rangle + c_n \left|\frac{3}{2}\right\rangle \tag{2.127b}$$

In terms of the states $|m_l m_z\rangle$, these may be written as

$$\psi_n^+ = \varphi_1 \alpha + \varphi_2 \beta \tag{2.128a}$$

$$\psi_n^- = \varphi_2^* \alpha - \varphi_1^* \beta \tag{2.128b}$$

where $\varphi_1^* = \varphi_1$ and $\varphi_2^* = -\varphi_2$. Under the time-reversal operator $T = iK\sigma_y$, where K is the conjugation operator, these states are related by

$$T\psi_n^+ = \psi_n^- \tag{2.129a}$$

$$T\psi_n^- = -\psi_n^+ \tag{2.129b}$$

Such a pair of states is said to form a Kramers doublet. The states (2.116) also formed a Kramers doublet.

Let us now consider the behavior of the lowest Kramers doublet in an external magnetic field. The Hamiltonian is

$$\mathcal{H}_Z = \mu_B(\boldsymbol{l} + 2\boldsymbol{s}) \cdot \boldsymbol{H} \tag{2.130}$$

Since we are within a manifold where J is a good quantum number, this may be written as

$$\mathcal{H}_Z = g_J \mu_B J_z H \tag{2.131}$$

The Zeeman matrix for the lowest doublet is

$$\begin{bmatrix} g_J\mu_B H\left(\frac{5}{2}a_1^2 + \frac{1}{2}b_1^2 - \frac{3}{2}c_1^2\right) & 0 \\ 0 & -g_J\mu_B H\left(\frac{5}{2}a_1^2 + \frac{1}{2}b_1^2 - \frac{3}{2}c_1^2\right) \end{bmatrix} \tag{2.132}$$

Therefore the doublet splits linearly in the field. An effective g value is often introduced by defining the splitting of two levels in a field as $g_{\text{eff}}\,\mu_B H$. In this case the effective g value would be

$$g_{\text{eff}} = g_J\left(5a_1^2 + b_1^2 - 3c_1^2\right) \tag{2.133}$$

Notice that the g value depends directly on the coefficients in the wave function (2.127). For this reason measurements of the g value provide a sensitive test of the ground-state wave function.

2.3.3 Semiconductors

It is interesting to carry out the calculation for the g value of an electron moving in the periodic potential, $V(\boldsymbol{r})$, of a crystal lattice. The Hamiltonian is

$$\mathcal{H} = \mathcal{H}_{\text{kinetic}} + \mathcal{H}_{\text{crystal}} + \mathcal{H}_{\text{spin-orbit}} + \mathcal{H}_{\text{Zeeman}} \tag{2.134}$$

The eigenfunctions of an electron moving in a periodic potential have the Bloch form $u_{nk}(\boldsymbol{r})\exp(i\boldsymbol{k}\cdot\boldsymbol{r})$ which give rise to energy bands characterized by band indices n and wave vector $\boldsymbol{k}$. In the absence of the Zeeman term the function $u_{nk}(\boldsymbol{r})$ satisfies

$$\left[\frac{1}{2m}(\boldsymbol{p}+\hbar\boldsymbol{k})^2 + V(\boldsymbol{r}) + \frac{\hbar}{4m^2c^2}\boldsymbol{\sigma}\times\nabla V(\boldsymbol{r})\cdot(\boldsymbol{p}+\hbar\boldsymbol{k})\right] u_{nk}(\boldsymbol{r}) = \epsilon_{nk}u_{nk}(\boldsymbol{r}). \tag{2.135}$$

In the case of the "three-five" semiconductors, which are composed of elements from the third and fifth columns of the period table (e.g., GaAs, InSb) the valence band is p-like while the conduction band is s-like. The band structure in the vicinity of $\boldsymbol{k}=0$ is illustrated in Fig. 2.12. The labels Γ_6, etc., indicate the irreducible representations according to which the wave functions at $\boldsymbol{k}=0$ transform. The first feature to note is that the spin-orbit interaction has split the threefold degeneracy of the p-like valence band. Secondly, the curvatures of these bands are different. This curvature is a measure of the effective mass of the electron. We may obtain an estimate of this mass by treating the term $(\hbar/m)\boldsymbol{k}\cdot\boldsymbol{p}$ in (2.135) as a perturbation.

Calculating the second-order correction to the energy of the conduction band leads to the effective mass tensor [34a],

$$\left(\frac{m}{m^*}\right)_{\mu\nu} = \delta_{\mu\nu} + \frac{2}{m}\sum_{\Gamma}\frac{\langle\Gamma_6|\,p_\mu|\Gamma\rangle\langle\Gamma|p_\nu\,|\Gamma_6\rangle}{\epsilon_{\Gamma_6}-\epsilon_\Gamma}. \tag{2.136}$$

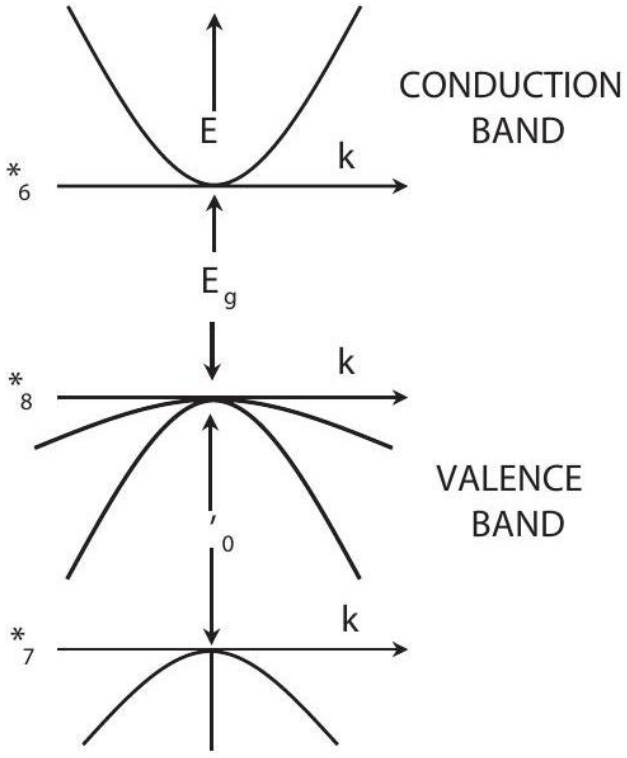


Fig. 2.12. Energy bands near $\boldsymbol{k}=0$ for zinc-blende crystals

The valence band wave functions are linear combinations of the usual x, y, and z components of the p state. Since the III-V compounds have the cubic zinc-blende structure, if we restrict our consideration only to the bands shown in Fig. 2.12 then there is only one parameter,

$$\frac{2}{m}\langle \Gamma, \mu | p_\mu | s \rangle \Big|^2 \equiv P^2$$

and the effective mass becomes

$$\frac{m}{m^*} = 1 + \frac{P^2}{3}\frac{3E_g + 2\Delta}{E_g\,(E_g + \Delta)} \tag{2.137}$$

Let us now consider the orbital moment associated with an electron at the bottom of the conduction band,

$$\langle \Gamma_6 |\, L_z \,| \Gamma_6 \rangle = \sum_\Gamma \left[\langle \Gamma_6 |\, x | \Gamma \rangle \langle \Gamma | p_y \,| \Gamma_6 \rangle - \langle \Gamma_6 |\, y | \Gamma \rangle \langle \Gamma | p_x \,| \Gamma_6 \rangle \right] \tag{2.138}$$

This may be rewritten by noting that

$$[x, \mathcal{H}] = \frac{i}{m} p_x + \frac{i\hbar}{4m^2c^2}(\boldsymbol{\sigma} \times \nabla V)_x \tag{2.139}$$

Neglecting the spin-orbit contribution, which can be shown to be small, the matrix elements of (2.139) satisfy

$$\langle \Gamma | x \,| \Gamma_6 \rangle = \frac{1}{m}\frac{\langle \Gamma | p_x \,| \Gamma_6 \rangle}{\epsilon_{\Gamma_6} - \epsilon_\Gamma} \tag{2.140}$$

This enables us to convert the orbital angular-momentum matrix element into a form similar to the effective mass. In particular, eliminating the linear-momentum matrix element between these two expressions leads to the relation [34b]

$$g^* = \frac{\mu^*}{\mu_B} = 1 + \langle \Gamma_6 |\, L_z \,| \Gamma_6 \rangle = 2\left[1 - \frac{\Delta}{3E_g + 2\Delta}\left(\frac{m}{m^*} - 1\right)\right] \tag{2.141}$$

In general, one must include additional bands, in which case the relationship between the effective g value and the effective mass is not as simple. Nevertheless, when the effective mass is very small the g value can become quite large. In InSb, for example, $m^*/m = -0.014$ and $g^* = -51.4$. Herring [35] has shown that the angular momentum (2.138) may be spatially decomposed into intraatomic and interatomic contributions. The large g values were shown to be associated with the latter, corresponding to interatomic circulating currents.

In this chapter we have seen the origins of various terms in the Hamiltonian that may influence the magnetic response of a system. In the remaining chapters we shall investigate how these terms manifest themselves when the system is excited with a space-and/or time-varying magnetic field.

Problems

2.1. While the discussion in Sect. 2.2.2 dealt mainly with the effect of the nuclear charge on the electron, the electron also affects the nucleus. Co^{57}, for example, is radio-active with a half-life of 270 days. It captures an electron to form Fe^{57} which decays by γ-radiation to an excited state with $I = 3/2$ which quickly (1.4×10^{-7} s) decays to the ground state with $I = 1/2$.

Ordinarily when the nucleus emits a γ-ray it recoils. However in 1958, R.L. Mossbauer discovered that nuclei embedded in solids can emit and absorb γ-rays without transferring recoil energy to the lattice. The basic experimental arrangement for measuring the γ-ray

energy consists of a moving source, an absorber, and a detector placed beyond the absorber. The transmitted intensity is plotted as a function of velocity. The 14.4KeVγ-ray has a natural linewidth of 5×10^{-9}eV which corresponds to a Doppler velocity of 1 mm/s.
(a) In Sect.2.2.2 we also assumed that there was no electron charge at the nucleus. But, in fact, s-electrons do have nonzero amplitude, $\psi(0)$, at the origin. Assuming the electron has a constant charge density, $e|\psi(0)|^2$, and interacts with the nuclear charge through a Coulomb potential Ze/r, out to a radius r_A for the excited state and r_B for the ground state, calculate the shift in the γ-ray energy due to this electronic charge density.
(b) If the emitter and the absorber have the same electronic structure their shifts will be the same. Therefore to observe this shift, called the isomer shift, the electronic charge density at the emitter nucleus, $e\,|\psi_e(0)|^2$ must be different from that at the absorber nucleus, $e\,|\psi_a(0)|^2$. Calculate the observed isomer shift, δ, in terms of these wavefunctions.
(c) Suppose the electronic charge distribution produces an electric field gradient at the Fe^{57} nucleus. Using (2.38), sketch the effect of this gradient on the energy levels in Fig. 2.13. Assuming dipole transition selection rules, i.e., $\Delta m_I = \pm 1$, sketch the Mossbauer spectrum.
(d) In order to simplify the Mossbauer spectrum, either the source or absorber should have an unsplit line. Co^{57} diffused into stainless steel is such a case. Figure 2.14 shows the absorption of Fe^{57} in Fe_2O_3 of γ-rays emitted from a stainless steel source. What is the effective magnetic field at the nucleus? Discuss the origin of this field in the context of (2.38).
2.2. As an example of the usefulness of character tables consider an atomic electron surrounded by an environment that has cubic symmetry. In particular, suppose it has sixfold coordination as shown in Fig. 2.1a. The symmetry elements which take this environment into itself are the identity (E), eight rotations by $2\pi/3$ about the cube body diagonals (C3), three rotations by π about the cube axis(C2), six rotations by $\pi/2$ about these same axes (C4) and six rotations by π about axes through the origin parallel to face diagonals (C2). These operations constitute a group, generally labelled O. The number of elements of the group, denoted by h, is 24 in this case. Notice that

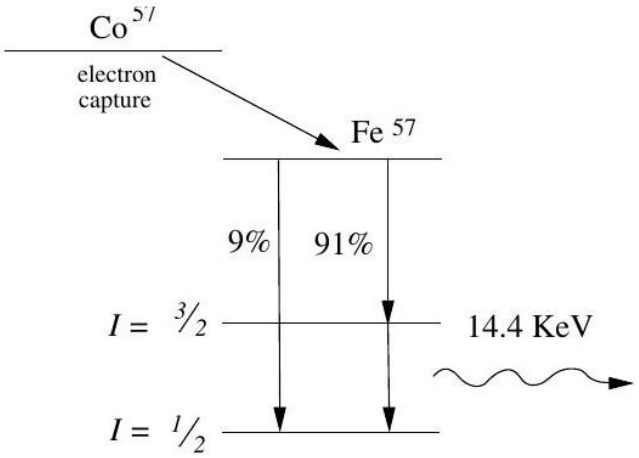


Fig. 2.13. Decay scheme of $\mathrm{Co^{57}}$ and Fe^{57}

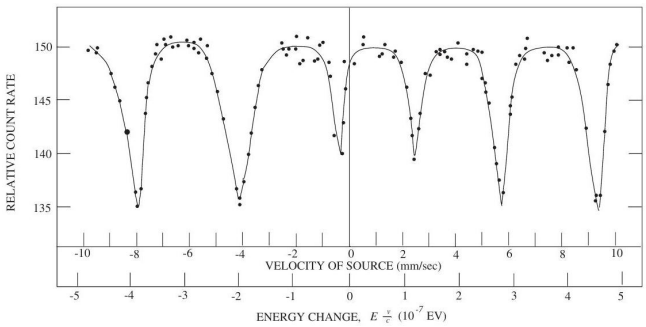


Fig. 2.14. The absorption of $\mathrm{Fe
^{57}$ bound in Fe_2O_3 of the 14.4 keV gamma ray emitted in the decay of Fe^{57} bound in stainless steel as a function of relative sourceabsorber velocity}

the elements group themselves into classes, G_k, with N_k elements. It can be shown that the number of irreducible representations is equal to the number of classes. Thus the character table will always be "square". The character table for this group is Table 2.6. An electron simply in a Coulomb potential has complete spherical symmetry. The set of rotations in three-dimensional space forms an infinite group. The eigenfunctions are spherical harmonics which transform into each other under rotation,

$$U_R(\alpha)Y_l^m = \sum \Gamma^{(l)}(\alpha)_{mm'}Y_\lambda^{m'},$$

where $\Gamma^l(\alpha)_{mm'} = e^{im\alpha}\delta_{m,m'}$. Therefore the character is

$$\chi^l(\alpha) = \frac{\sin(l+1/2)\alpha}{\sin(\alpha/2)}.$$

Notice that the character associated with the identity operation, E, i.e., $\alpha = 0$, is

$$\chi^{(l)}(0) = \lim_{\alpha\to 0}\frac{\sin(l+1/2)\alpha}{\sin(\alpha/2)} = 2l+1.$$

In the presence of the lower cubic symmetry this $(2l+1)$-dimensional representation is reducible to the irreducible representations of the O group. The decomposition of the reducible character is given by

$$\chi(R) = \sum_i a_i\chi^{(l)}(R),$$

where $\chi^l(R)$ is the character of the i^{th} representation for the group element R. The coefficient a_i is:

$$a_i = (1/h)\sum_k N_K\chi^{(i)}\left(G_K\right)^*\chi\left(G_K\right).$$

Suppose, for example, we have an s-state ($1 = 0$) electron in a cubic environment. Then the character $\chi^{(l)}(R) = 1$ for each of the five classes, and the representation is simply A_1 :

$$\frac{s}{\text{free space}} \quad \frac{A_1}{\text{cubic environment}}.$$

Calculate how the $2l+1$ degeneracies of p, d, and f electrons are split in a cubic environment. Which have the potential for complete quenching?

2.3. Consider the sixfold coordination shown in Fig. 2.1a. If we elongate the neighbors in the

z-direction, this corresponds to a tetragonal distortion and adds a term to the potential (2.50) proportional to Y_2^0. In cartesian coordinates this becomes

$$C_2\left(3z^2-r^2\right), \quad C_2 = qq'\left[\frac{1}{a_1^3}-\frac{1}{a_2^3}\right]$$

for the sixfold tetragonal coordination we are considering. The corresponding operator equivalent is

Table 2.6. Character table for the point group O

	E	$8C_3$	$3C_2 = 3C_4^2$	$6C_2$	$6C_4$
A_1	1	1	1	1	1
A_2	1	1	1	-1	-1
E	2	-1	2	0	0
T_1	3	0	-1	-1	1
T_2	3	0	-1	1	-1

$$\sum\left(3z^2-r^2\right) \Rightarrow \alpha\overline{r^2}\left[3L_z^2-L(L+1)\right]$$

where $\alpha = \pm 2/21$ for D states, and where the negative sign applies for T_2 states and the plus sign for E states.

Calculate the new eigenvalues and eigenfunctions resulting from this distortion, and plot the resulting splitting of the levels in Fig. 2.1 as a function of $C_2\alpha\overline{r^2}$.

2.4. The dipole-dipole interaction (2.58) can also give rise to anisotropy. Assuming all the spins are parallel and are characterized by direction cosines $\alpha_x, \alpha_y, \alpha_z$, express the anisotropy energy in terms of these cosines. Show that in a system with cubic anistropy the dipole interaction does not contribute to the anisotropy. In what sorts of materials would you expect to find dipolar anisotropy?

2.5. An alternative set of basis states for calculating the effect of electronelectron interactions are molecular orbitals. Consider the bonding and antibonding orbitals,

$$\varphi_b(\boldsymbol{r}) = 1\sqrt{2(1+l)}\left[\varphi_1(\boldsymbol{r})+\varphi_2(\boldsymbol{r})\right]$$
$$\varphi_a(\boldsymbol{r}) = 1/\sqrt{2(1-l)}\left[\varphi_1\right)(\boldsymbol{r})-\varphi_2(\boldsymbol{r})\Big].$$

Each orbital can hold two electrons. Assume that one electron is in the bonding orbital and the other in the antibonding orbital. Calculate the matrix analogous to (2.81) and diagonalize to obtain the singlet and the triplet eigenvalues. What is the difference between the singlet and triplet energies? What is the role of the "exchange" integral J_{12} ?

2.6. Using arguments similar to that used for $[\mathrm{FeCp}_2^*]^+[\mathrm{TCNE}]^-$, predict the magnetic order for the salts involving $\mathrm{Cr}, \mathrm{Mn}, \mathrm{Co}$, and Ni . The spin for these materials are $3/2, 1, 0$ and $1/2$, respectively, in accord with their electronic structures:

$$\mathrm{Cr} : a_{1g}e_{2g}^2$$
$$\mathrm{Mn} : a_{1g}e_{2g}^3$$
$$\mathrm{Co} : a_{1g}^2e_{2g}^4$$
$$\mathrm{Ni} : a_{1g}^2e_{2g}^4e_{1g}.$$

10.4 The Static Susceptibility of Noninteracting Systems

In this chapter we shall investigate the static susceptibility of systems described by Hamiltonians which may be written as the sum of noninteracting terms. Our approach here, and in later discussions, will be to perturb the system with a magnetic field and then compute the response to this field. In this chapter we shall discuss the effect of a time-independent field $\boldsymbol{H}$.

In the literature of magnetism a distinction is often made between insulators and conductors. This distinction is basically an operational one. That is, insulators are characterized by charge distributions that are reasonably well localized to unit cells; such systems may be described by localized effective spins. Metals are characterized by itinerant electrons which require an entirely different description. Of course, there are materials, such as metals containing rare-earth ions, in which both types of charge distributions coexist. For the present, however, we shall follow tradition and discuss the response to local moments and itinerant systems separately.

10.4.1 3.1 Localized Moments

Let us consider a system of N noninteracting identical ions or molecules each having h electrons. In the absence of an external field each ion or molecule is characterized by a Hamiltonian $\mathcal{H}_i^0$. We are interested in the response of this system to a small static magnetic field. Therefore we add a Zeeman term to the Hamiltonian to give a total Hamiltonian,

$$\mathcal{H} = \sum_i \mathcal{H}_i = \sum_i \left(\mathcal{H}_i^0 + \mathcal{H}_i^Z\right). \tag{3.1}$$

The magnetization is given by (1.49) as

$$\boldsymbol{M}(\boldsymbol{r}) = \mathrm{Tr}\{\rho \boldsymbol{\mathcal{M}}(\boldsymbol{r})\}, \tag{3.2}$$

where

$$\boldsymbol{\mathcal{M}}(\boldsymbol{r}) = \frac{1}{2}\sum_{i=1}^{N}\sum_{\alpha=1}^{h}\left[\boldsymbol{\mu}_{i\alpha}\delta\left(\boldsymbol{r}-\boldsymbol{r}_{i\alpha}\right) + \delta\left(\boldsymbol{r}-\boldsymbol{r}_{i\alpha}\right)\boldsymbol{\mu}_{i\alpha}\right]. \tag{3.3}$$

Here $\mu_{i\alpha}$ is the magnetic-moment operator associated with the α th electron on the i th ion. Since the ionic electrons interact with one another, the appropriate basis in which to evaluate (3.2) consists of the ionic eigenfunctions. Therefore we make use of the Wigner-Eckhart theorem and replace (3.3) by

$$\boldsymbol{\mathcal{M}}(\boldsymbol{r}) = \sum_{i=1}^{N}\boldsymbol{\mu}_i\delta\left(\boldsymbol{r}-\boldsymbol{r}_i\right). \tag{3.4}$$

We have assumed that the ionic moments have no spatial extent. As long as the wavelength of the applied field is larger than an atomic dimension, (3.4) is valid. However, when we talk about neutron scattering in Chap. 10 this assumption will have to be modified.

Since we are applying a static field, the density matrix is independent of time. Therefore $[\mathcal{H}, \rho] = 0$. The equilibrium solution of this equation may be written as [36]

$$\rho = \exp[\beta(F - \mathcal{H})], \tag{3.5}$$

where F is the free energy

$$F = -k_B T \ln Z \tag{3.6}$$

and Z is the partition function

$$Z = \mathrm{Tr}\{\exp(-\beta\mathcal{H})\}. \tag{3.7}$$

The solution (3.5) for the density matrix, known as the canonical distribution, is the quantum-mechanical generalization of the classical Gibbs distribution. Using (3.6), (3.7), we may write this as

$$\rho = \frac{\exp(-\beta\mathcal{H})}{\mathrm{Tr}\{\exp(-\beta\mathcal{H})\}}, \tag{3.8}$$

which, because of the form of the Hamiltonian (3.1), may be factored as

$$\rho = \frac{\prod_i \exp\left(-\beta\mathcal{H}_i\right)}{\mathrm{Tr}\left\{\prod_i \exp\left(-\beta\mathcal{H}_i\right)\right\}} = \prod_i \rho_i. \tag{3.9}$$

Therefore

$$\boldsymbol{M}(\boldsymbol{r}) = \sum_i \mathrm{Tr}\left\{\rho_i \boldsymbol{\mu}_i \delta\left(\boldsymbol{r} - \boldsymbol{r}_i\right)\right\}, \tag{3.10}$$

which defines the ionic moment

$$\boldsymbol{M}_i = \mathrm{Tr}\left\{\rho_i \boldsymbol{\mu}_i\right\}. \tag{3.11}$$

In writing (3.10) we have assumed that we may interchange the operations of taking the trace and summing over the ions. This neglects the possibility that an electron on one ion may hop over to a neighboring ion. If the electrons were distinct, as labeled by their positions in the lattice, such hopping would be forbidden. Therefore our assumption of interchangeability is, in this sense, equivalent to the assumption that the ionic electrons are distinguishable. The importance of this assumption will become evident in Sect.3.2.

So far we have not said anything about the spatial variation of the applied field. Let us assume that it has the wavelike form

$$\boldsymbol{H}(\boldsymbol{r}) = \boldsymbol{H}\cos\left(\boldsymbol{q}_0 \cdot \boldsymbol{r}\right) \tag{3.12}$$

We shall always assume that the wave vector $\boldsymbol{q}_0$ is one of the discrete wave vectors lying within the first Brillouin zone associated with the lattice under consideration. Let us keep in mind that we may have moments at only a few of the available lattice sites. If the applied field does not vary appreciably over a unit cell, the ionic eigenvalues will correspond to those of a uniform field with the value $\boldsymbol{H}_i = \boldsymbol{H}\cos\left(\boldsymbol{q}_0 \cdot \boldsymbol{r}_i\right)$. The corresponding moment is

$$\boldsymbol{\mu}_i = \frac{\partial\mathcal{H}_i}{\partial\boldsymbol{H}_i} \tag{3.13}$$

If the eigenvalues associated with the i th ion are E_n, then the ionic moment in the direction of the applied field is

$$M_i \equiv \mathrm{Tr}\left\{\rho_i \mu_i\right\} = \frac{\sum_n \exp\left(-\beta E_n\right)\left(\partial E_n/\partial H_i\right)}{\sum_n \exp\left(-\beta E_n\right)} \tag{3.14}$$

Since the eigenvalues E_n are functions of the field H_i, let us expand in powers of the field:

$$E_n = E_n^{(0)} + E_n^{(1)} H_i + E_n^{(2)} H_i^2 + \ldots \tag{3.15}$$

Since the magnetic field is only a probe, we may make it as small as we like. In particular, let us assume that the field-dependent corrections to the energy in (3.15) are smaller than the thermal energy. Then

$$\exp\left(-\beta E_n\right) = \exp\left(-\beta E_n^{(0)}\right)\left(1 - \beta E_n^{(1)} H_i - \ldots\right) \tag{3.16}$$

and (3.14) becomes

$$\boldsymbol{M}_i = \frac{\sum_n \exp\left(-\beta E_n^{(0)}\right)\left(1 - \beta E_n^{(1)} H_i - \ldots\right)\left(E_n^{(1)} + 2E_n^{(2)} H_i + \ldots\right)}{\sum_n \exp\left(-\beta E_n^{(0)}\right)\left(1 - \beta E_n^{(1)} H_i - \ldots\right)} \tag{3.17}$$

Assuming that the ion has no moment in the absence of the field and retaining only those terms linear in the field, we have

$$\boldsymbol{M}_i = \frac{\sum_i \left(\beta E_n^{(1)2} - 2E_n^{(2)}\right)\exp\left(-\beta E_n^{(0)}\right)}{\sum_n \exp\left(-\beta E_n^{(0)}\right)} H_i \equiv \chi_i H_i \tag{3.18}$$

which defines the quantity χ_i. Substituting this expression back into (3.10) and taking the spatial Fourier transform, we obtain

$$\boldsymbol{M}_z(\boldsymbol{k}) = \sum_i \boldsymbol{M}_i \exp\left(-i\boldsymbol{k}\cdot\boldsymbol{r}_i\right),$$

which may be written as

$$\boldsymbol{M}_z(\boldsymbol{k}) = \sum_i \left[\frac{1}{V}\sum_i \chi_i \exp[-i(\boldsymbol{k}-\boldsymbol{q})]\cdot\boldsymbol{r}_i\right] H(\boldsymbol{q}),$$

where

$$H(\boldsymbol{q}) = \frac{HV}{2}\left[\Delta\left(\boldsymbol{q}+\boldsymbol{q}_0\right) + \Delta\left(\boldsymbol{q}-\boldsymbol{q}_0\right)\right].$$

Therefore

$$\chi(\boldsymbol{k},\boldsymbol{q}) = \frac{1}{V}\sum_i \chi_i \exp\left[-i(\boldsymbol{k}-\boldsymbol{q})\cdot\boldsymbol{r}_i\right].$$

If there is a moment at each lattice site, then the sum in this expression gives $N\Delta(\boldsymbol{k}-\boldsymbol{q})$, and

$$\chi(\boldsymbol{q}) = \frac{N}{V}\frac{\sum_n \left(\beta E_n^{(1)2} - 2E_n^{(2)}\right)\exp\left(-\beta E_n^{(0)}\right)}{\sum_n \exp\left(-\beta E_n^{(0)}\right)} \equiv \chi. \tag{3.19}$$

The fact that this is independent of $\boldsymbol{q}$ is a consequence of our assumption that the magnetization is everywhere proportional to the local field. Thus it is not surprising that

$$\chi\left(\boldsymbol{r}-\boldsymbol{r}'\right) = \frac{1}{V}\sum_{\boldsymbol{q}} \exp\left[i\boldsymbol{q}\cdot\left(\boldsymbol{r}-\boldsymbol{r}'\right)\right]\chi(\boldsymbol{q}) = \frac{N}{V}\chi\delta\left(\boldsymbol{r}-\boldsymbol{r}'\right). \tag{3.20}$$

Let us now consider some specific applications of (3.19).

3.1.1 Diamagnetism

In Sect. 2.2 we found that the Zeeman Hamiltonian for an ionic or molecular system of electrons was

$$\mathcal{H}_i^Z = \sum_\alpha \left[\mu_B \boldsymbol{l}_\alpha \cdot \boldsymbol{H}_i + \frac{e^2 H_i^2}{8mc^2} \left(x_\alpha^2 + y_\alpha^2 \right) + \mu_B \boldsymbol{\sigma}_\alpha \cdot \boldsymbol{H}_i \right]. \tag{3.21}$$

The effect of this interaction on the eigenvalues may be obtained from perturbation theory, with the eigenfunctions of $\mathcal{H}_i^0$ as the unperturbed states.

A diamagnetic system is one in which all the matrix elements of the orbital and spin angular momenta are 0 . Therefore the only correction to the energy arises from the second term in (3.21). If the eigenfunctions of $\mathcal{H}_i^0$ are written as $|n\rangle$, then

$$E_n^{(2)} = \frac{e^2}{8mc^2} \langle n | \sum_\alpha \left(x_\alpha^2 + y_\alpha^2 \right) |n\rangle .. \tag{3.22}$$

Notice that the susceptibility involves the statistical average of

$$\sum_\alpha \left(x_\alpha^2 + y_\alpha^2 \right).$$

This is usually written as $\frac{2}{3}Z$ times the mean square radius $\langle r^2 \rangle$. Therefore the diamagnetic contribution to the quantity χ in the generalized susceptibility (3.19) is

$$\chi_{\text{diamag}} = -\frac{Ze^2}{6mc^2} \left\langle r^2 \right\rangle . \tag{3.23}$$

This negative response is just a manifestation of Lenz's law, which states that in the presence of a changing flux a system of charges will develop currents in such a direction as to oppose the flux change. These currents are the source of the diamagnetic moment.

In dielectric solids the molar susceptibility is of the order of -10^{-5} cm^3/ mole. This corresponds to an induced moment per atom of $10^{-8} H_i$ Bohr magneton, which is very small.

Notice that the diagmagnetic susceptibility appears to depend on the choice of the origin for computing $\langle r^2 \rangle$. This ambiguity is, unfortunately, real. Van Vleck has shown that the change in diamagnetic susceptibility resulting from a change in origin is just compensated for by a corresponding change in the Van Vleck paramagnetic susceptibility, to be discussed in the next section [37]. For a single atom or ion the origin is usually taken at the nucleus. In this case the Van Vleck contribution is small, and the diamagnetic susceptibility is well described by (3.23). However for polyatomic molecules both contributions to the susceptibility must be considered together.

3.1.2 Paramagnetism of Transition-Metal Ions

In the preceding discussion we assumed that all the matrix elements, both diagonal and off-diagonal, of the orbital and spin angular moments were 0 . The interesting aspects of magnetism arise, however, when these matrix elements are nonzero. Let us now examine the response of such ions to a static field. Since crystal fields have important effects on the orbital angular momentum, it is convenient to discuss transition-metal ions and rare-earth ions separately.

In Sect. 2.3 we found that a transition-metal ion could be described by a spin Hamiltonian. Let us consider the particular case where the crystal field has removed all the orbital degeneracy. The spin Hamiltonian which describes the lowest orbital state then has the form

$$\mathcal{H}_{\text{eff}} = g_\parallel \mu_B H S_z - \mu_B^2 \wedge_\parallel H^2. \tag{3.24}$$

If we take the eigenfunctions of S_z as our unperturbed states, then our orbital singlet splits into the $2S+1$ states characterized by the spin-projection quantum number M_S,

$$E_{M_S} = E^0_{M_S} + g_{\|}\mu_B M_S H - \mu_B^2 \wedge_{\|} H^2. \tag{3.25}$$

Making the identifications with (3.15) and using (3.14), we find that the ($\boldsymbol{k} = 0, \boldsymbol{q} = 0$) susceptibility may be written as the sum of the Langevin susceptibility and the Van Vleck susceptibility,

$$\chi(0,0) = \chi_{\text{Langevin}} + \chi_{VV}, \tag{3.26}$$

where

$$\chi_{\text{Langevin}} = \frac{N}{V}\frac{g_{\|}^2\mu_B^2}{k_BT}\frac{\sum_{M=-S}^{S} M_s^2}{2S+1} = \frac{N}{V}\frac{g_{\|}^2\mu_B^2 S(S+1)}{3k_BT} \equiv \frac{C}{T}, \tag{3.27}$$

which defines the Curie constant C, and

$$\chi_{VV} = \frac{N}{V} 2\mu_B^2 \wedge_{\|} \tag{3.28}$$

The Van Vleck susceptibility arises from that part of the orbital moment which has been admixed back into the ground state by the orbital Zeeman effect. The fact that this contribution is temperature independent is a result of the assumption that the energy difference between the ground state and the higher-lying orbital state entering $\wedge_{\|}$is larger than k_BT. If this condition does not hold, then the energy shift of this upper state must be considered, with the result that the Van Vleck susceptibility will acquire a temperature dependence.

The inverse temperature dependence of the Langevin susceptibility was first observed experimentally by Pierre Curie in 1895 and is referred to as

Table 3.1. Effective moments associated with transition-metal ions [38]

	Ti^{3+}	V^{3+}	Cr^{3+}	Mn^{3+}	Mn^{2+}, Fe^{3+}	Fe^{2+}	Co^{2+}	Ni^{2+}	Cu^{2+}
$2\sqrt{S(S+1)}$	1.73	2.83	3.87	4.00	5.92	4.90	3.87	2.83	1.73
μ_{eff} (exp)	1.80	2.80	3.80	4.90	5.90	5.40	4.80	3.20	1.90
$g_J\sqrt{J(J+1)}$	1.55	1.63	0.77	0	5.92	6.70	6.63	5.59	3.55
λ [cm^{-1}]	154	105	91	88	-	-103	-178	-325	-829

Curie's law. The Curie constant C is used experimentally to determine an effective moment in units of Bohr magnetons by the relation

$$\mu_{\text{eff}} = \sqrt{\frac{3k_BC}{(N/V)\mu_B^2}} \tag{3.29}$$

If the experimental values of μ_{eff} are compared with the spin-only theoretical values in Table 3.1, we see that the quenching is nearly complete for the first half of the periodic table. Orbital contributions become more important in the last half of the row because the spin-orbit coupling constant is becoming larger.

Because of the availability of high-static magnetic fields and low temperatures it is relatively easy to reach the nonlinear region of saturation. To describe this situation we must use the more general expression for the ionic moment given by (3.14). The Van Vleck susceptibility is unchanged. However, the Langevin susceptibility becomes

$$\chi_{\text{Langevin}} = -\frac{Ng_{\parallel}\mu_B S}{HV}\frac{\sum_{M_S=-S}^{S}\left[\exp\left(-\beta g_{\parallel}\mu_B H M_S\right)\right]M_S/S}{\sum_{M_S=-S}^{S}\exp\left(-\beta g_{\parallel}\mu_B H M_S\right)}$$
$$= \frac{Ng_{\parallel}\mu_B S}{HV}B_S\left(\frac{g_{\parallel}\mu_B SH}{k_B T}\right) \tag{3.30}$$

where $B_S\left(g_{\parallel}\mu_B SH/k_B T\right)$ defines the Brillouin function,

$$B_S(x) = \frac{2S+1}{2S}\coth\left(\frac{(2S+1)x}{2S}\right) - \frac{1}{2S}\coth\left(\frac{x}{2S}\right) \tag{3.31}$$

For $S = 1/2, B_{1/2} = \tanh(x)$.

It is customary for experimentalists to plot the inverse susceptibility against temperature. In Fig. 3.1 we have plotted $1/\chi$ for a spin of $\frac{3}{2}$ in a field of 30,000 Oe. Since the Curie constant is of the order of unity, the Van Vleck contribution becomes important only for $T > \chi_{VV}^{-1}$. Normally, the Van Vleck susceptibility is of the order of 10^{-5}. However, the larger value of 10^{-3} has been chosen in Fig. 3.1 to illustrate its effect.

The only approximation in obtaining (3.17) was that the Zeeman energy be smaller than the thermal energy. Therefore this expression also applies if the thermal energy exceeds the crystal-field energy. In this case we must be careful to use the proper energies E_n. We shall consider an example of how to include excited-state effects in the next section.

3.1.3 Paramagnetism of Rare-Earth Ions

The effect of crystal fields on rare-earth ions is generally quite small in comparison with the spin-orbit interaction. Let us therefore consider our rareearth ion as described by eigenstates characterized by the quantum numbers J, M_J, L, and S. The perturbing Zeeman Hamiltonian has the form

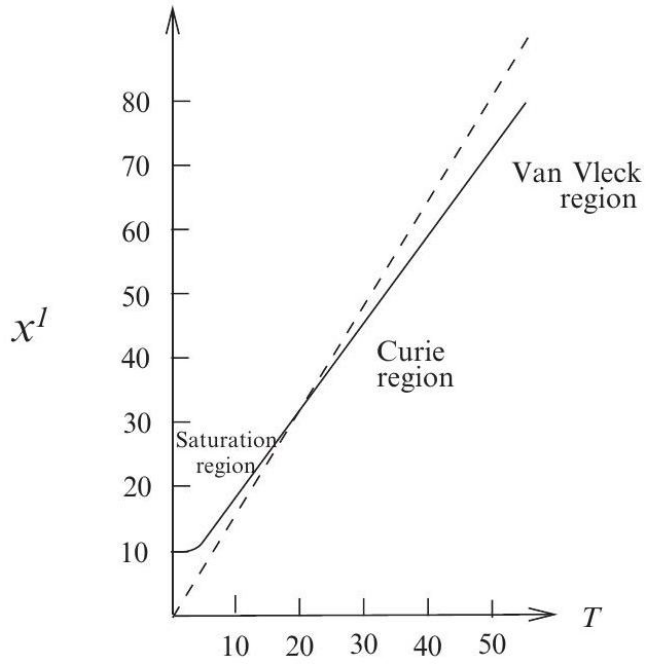


Fig. 3.1. Inverse susceptibility as a function of temperature for a spin $\frac{3
{2}$ in a field of 30,000 Oe}

$$\mathcal{H}_Z = \mu_B(\boldsymbol{L} + 2\boldsymbol{S}) \cdot \boldsymbol{H}, \tag{3.32}$$

If we restrict ourselves to the lowest J manifold and define the z direction as the direction of the applied field, then this Hamiltonian becomes

$$\mathcal{H}_Z = g_J \mu_B H J_z. \tag{3.33}$$

The effect of this interaction is to split the J state into its $2J+1$ Zeeman components, with energies

$$E_{J,M_J} = g_J \mu_B M_J H. \tag{3.34}$$

From (3.15), (3.19) we find

$$\chi(0,0) = \frac{N}{V} \frac{g_J^2 \mu_B^2 J(J+1)}{3k_B T}, \tag{3.35}$$

and from (3.29) we obtain the effective moment,

$$\mu_{\text{eff}} = g_J \sqrt{J(J+1)}. \tag{3.36}$$

Most of the rare-earth ions exhibit moments in good agreement with (3.36). However, a notable exception is Eu^{3+}. This has a $4f^6$ configuration, which corresponds to a $^7F_0$ ground state. Therefore $J = 0$, and μ_{eff} should also be 0 . Experimentally, however, a moment of 3.4 Bohr magnetons has been found at room temperature. To account for this difference we note that the $^7F_1$ excited state lies $\Delta/k_B \simeq 350$ K above the ground state. Therefore we

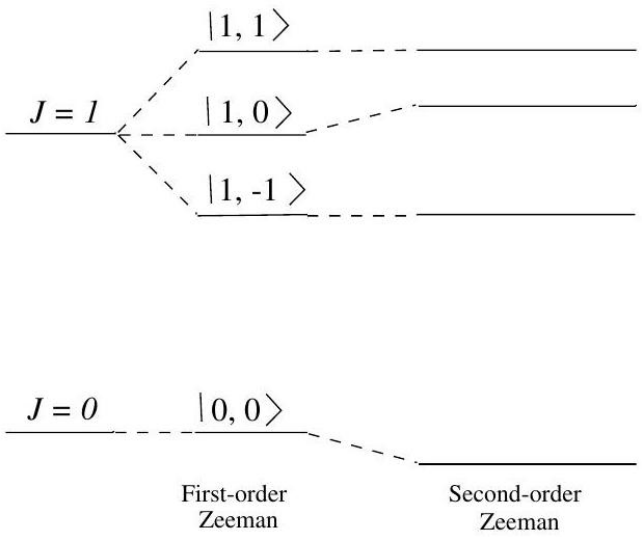


Fig. 3.2. Effect of the Zeeman interaction on the low-lying states of $\mathrm{Eu^{3+}}$

must include the influence of this state on our susceptibility. There are two effects to be considered, the addition of the terms in (3.19) associated with the excited state and the shift in the ground-state energy produced by the Zeeman coupling to the excited state.

From perturbation theory we find that the first-order corrections to the energies are

$$E^{(1)}_{J,M_J} H = g_J \mu_B H \langle J, M_J | J_z | J, M_J \rangle, \tag{3.37}$$

and the second-order corrections are

$$E^{(2)}_{J,M_J} H^2 = \mu_B^2 H^2 \sum_{J',M'_{J'}} \frac{|\langle J, M_J| \, L_z + 2S_z \, |J', M'_{J'}\rangle \, |^2}{E^{(0)}_{J,M_J} - E^{(0)}_{J',M'_{J'}}}. \tag{3.38}$$

From the properties of angular momentum [39] we find that the matrix element in (3.38) is nonvanishing only for $J' = J \pm 1$ and $M'_{J'} = M_J$. Therefore only the $|0,0\rangle$ and $|1,0\rangle$ states are coupled in second order. This matrix element is $\langle 0,0|L_z + 2S_z|1,0\rangle = 2$. The resulting eigenvalues are

$$\begin{aligned} E_{0,0} &= -\frac{4\mu_B^2 H^2}{\Delta} \\ E_{1,-1} &= \Delta - g_1 \mu_B H \\ E_{1,0} &= \Delta + \frac{4\mu_B^2 H^2}{\Delta} \\ E_{1,1} &= \Delta + g_1 \mu_B H \end{aligned} \tag{3.39}$$

where g_1 is the Landé g value (2.124) for $J = 1$. These eigenvalues are indicated in Fig. 3.2. Substituting these results into (3.19) leads to an effective moment at room temperature of 3.2 Bohr magnetons, in much better agreement with the experimental value.

In conclusion, we have found that the susceptibility of an insulator containing noninteracting ionic moments is

$$\chi(T) = \chi_{\text{diamag}} + \chi(T)_{\text{Langevin}} + \chi_{VV}. \tag{3.40}$$

The temperature dependence arises primarily from the Langevin or Curie part.

10.4.2 3.2 Metals

We now turn to the response of a nonmagnetic metal to an applied field. The static response to a uniform field was first investigated by Pauli, who considered the spin paramagnetism, and by Landau, who considered the orbital diamagnetism.

3.2.1 Landau Diamagnetism

Let us consider the orbital moment induced in a free-electron metal by a static applied field. For the time being we shall neglect the spin associated with the electrons in a metal. As mentioned in Sect. 1.2, the classical approach to this problem leads to the conclusion that the metal shows no response to the applied field. Landau was the first to show that a quantum-mechanical approach indicates that a free-electron system does in fact respond to the applied field, and that this response is diamagnetic [40]. What he did was to solve the Schrödinger equation for a single electron in a magnetic field; with the resulting eigenvalues he was able to compute the free energy, from which he obtained the magnetization. The corresponding susceptibility was found to be $-\frac{1}{3}$ that obtained by Pauli for the spin contribution. This relationship is valid for all temperatures in degenerate as well as nondegenerate free-electron systems.

Peierls extended Landau's calculation to the case of electrons that are nearly bound by a periodic lattice potential [41], and Wilson considered nearly free electrons [42]. It was found that the effect of the lattice was, essentially, to modify Landau's result by a factor m/m^*, where m^* is the electron effective mass at the Fermi surface.

It is interesting to consider why Miss van Leeuwen's theorem, which we mentioned in Sect. 1.2, does not hold in the quantum-mechanical case. The reason is that those electrons whose orbits are interrupted by the boundary have, in general, high quantum numbers. Therefore,

their energy is also high which makes them thermo-dynamically less important than the "bulk" electrons. Consequently, the diamagnetic moments now dominate the paramagnetic moments arising from the "boundary" electrons.

The problem of boundary effects arises when one calculates the susceptibility by first solving the Schrödinger equation for the energy levels explicitly and then using them to calculate the partition function. Sondheimer and Wilson
have pointed out that the whole question of boundary effects may be avoided by calculating the partition function without explicit knowledge of the energy levels [43].

We shall now calculate the Landau susceptibility by another method [44] that is more in keeping with our linear response philosophy. Let us assume that we have a free-electron system characterized by a spherical Fermi surface. To this system we apply a static external magnetic field $\boldsymbol{H}_0(\boldsymbol{r})$, which arises from an external current density $\boldsymbol{j}_0(\boldsymbol{r})$ according to the Maxwell equation

$$\nabla \times \boldsymbol{H}_0(\boldsymbol{r}) = \frac{4\pi}{c} \langle \boldsymbol{j}_0(\boldsymbol{r}) \rangle . \tag{3.41}$$

This field induces a current density $\boldsymbol{j}(\boldsymbol{r})$ in the electron system, which in turn acts as the source for an additional magnetic field. The resulting total local field, denoted by $\boldsymbol{H}(\boldsymbol{r})$, is also related to its current sources by the Maxwell equation

$$\nabla \times \boldsymbol{H}(\boldsymbol{r}) = \frac{4\pi}{c} \langle \boldsymbol{j}_0(\boldsymbol{r}) + \boldsymbol{j}(\boldsymbol{r}) \rangle . \tag{3.42}$$

The local field defined in this manner has zero divergence. Therefore it may be expressed in terms of a vector potential as

$$\boldsymbol{H}(\boldsymbol{r}) = \nabla \times \boldsymbol{A}(\boldsymbol{r}). \tag{3.43}$$

In the presence of this potential the electron velocity becomes

$$\boldsymbol{v}_i = \frac{1}{m} \left[\boldsymbol{p}_i - \frac{e}{c} \boldsymbol{A}(\boldsymbol{r}_i) \right] . \tag{3.44}$$

The induced-current-density operator is related to this velocity by

$$\boldsymbol{j}(\boldsymbol{r}) = \frac{e}{2} \sum_i \left[\boldsymbol{v}_i \delta(\boldsymbol{r} - \boldsymbol{r}_i) + \delta(\boldsymbol{r} - \boldsymbol{r}_i) \boldsymbol{v}_i \right] . \tag{3.45}$$

According to (1.8), the average value of this current density determines the magnetization,

$$\langle \boldsymbol{j}(\boldsymbol{r}) \rangle = c \nabla \times \boldsymbol{M}(\boldsymbol{r}). \tag{3.46}$$

Using the Fourier transforms of (3.42), (3.43), (3.46) and the definition of the susceptibility, $\boldsymbol{M}(\boldsymbol{q}) = \boldsymbol{\chi}(\boldsymbol{q}) \cdot \boldsymbol{H}_0 \boldsymbol{q}$), we obtain

$$\chi(\boldsymbol{q}) = \frac{1}{4\pi} \frac{\langle \boldsymbol{j}(\boldsymbol{q}) \rangle}{\langle \boldsymbol{j}(\boldsymbol{q}) \rangle - (cq^2/4\pi) \boldsymbol{A}(\boldsymbol{q})}. \tag{3.47}$$

Therefore, if we can compute the average current density to lowest order in the vector potential, we shall have an expression for the susceptibility.

Introducing the Fourier momentum operator $\boldsymbol{p}_{\boldsymbol{q}}$, defined by

$$\boldsymbol{p}_{\boldsymbol{q}} \equiv \frac{1}{2} \sum_i \left[\boldsymbol{p}_i \exp(-i\boldsymbol{q} \cdot \boldsymbol{r}_i) + \exp(-i\boldsymbol{q} \cdot \boldsymbol{r}_i) \boldsymbol{p}_i \right] \tag{3.48}$$

we find from (3.45) that

$$\langle \boldsymbol{j}(\boldsymbol{q})\rangle = \frac{e}{m}\left\langle \boldsymbol{p}_{\boldsymbol{q}}\right\rangle - \frac{Ne^2}{mcV}\boldsymbol{A}(\boldsymbol{q}). \tag{3.49}$$

To obtain the diagonal components of χ we must find the average value of that component of $\boldsymbol{p}_{\boldsymbol{q}}$ parallel to $\boldsymbol{A}(\boldsymbol{q})$ to the lowest order in the vector potential. If $\hat{\boldsymbol{n}}$ is a unit vector parallel to $\boldsymbol{A}(\boldsymbol{q})$, then we must compute $\left\langle \boldsymbol{p}_{\boldsymbol{q}}\cdot\hat{\boldsymbol{n}}\right\rangle$. This we shall do by perturbation theory.
In the presence of the field the Hamiltonian is

$$\mathcal{H} = \sum_i \frac{1}{2m}\left[\boldsymbol{p}_i - \frac{e}{c}\boldsymbol{A}\left(\boldsymbol{r}_i\right)\right]^2. \tag{3.50}$$

Since the kinetic-energy part gives rise to our unperturbed Fermi sphere, the perturbation, to lowest order, is

$$\mathcal{H}_1 = -\frac{e}{mc}\sum_i \boldsymbol{A}\left(\boldsymbol{r}_i\right)\cdot\boldsymbol{p}_i. \tag{3.51}$$

Expanding $\boldsymbol{A}\left(\boldsymbol{r}_i\right)$ in a Fourier series and making use of (3.48), we obtain

$$\mathcal{H}_1 = -\frac{e}{mcV}\sum_k \boldsymbol{A}(\boldsymbol{k})\cdot\boldsymbol{p}_{\boldsymbol{k}}. \tag{3.52}$$

Writing the eigenstate of the perturbed system as $|n\rangle$ and that of the unperturbed system as $|\ n)$, we have

$$|0\rangle = |\ 0) + \sum_{n\neq 0}\left(\frac{(n\,|\mathcal{H}_1|\,0)}{E_0 - E_n}\right| n\Big). \tag{3.53}$$

Therefore

$$\left\langle \boldsymbol{p}_{\boldsymbol{q}}\cdot\hat{\boldsymbol{n}}\right\rangle\hat{\boldsymbol{n}} = \frac{2e}{mcV}\sum_{n\neq 0}\frac{\left|\left(n\left|\boldsymbol{p}_{\boldsymbol{q}}\cdot\hat{\boldsymbol{n}}\right|0\right)\right|^2}{E_n - E_0}\boldsymbol{A}(\boldsymbol{q}). \tag{3.54}$$

Since the states $|\ n)$ are many-electron states and the operator $\boldsymbol{p}_{\boldsymbol{q}}$ is a manyelectron operator, it is convenient to second quantize $\boldsymbol{p}_{\boldsymbol{q}}$:

$$\boldsymbol{p}_q\cdot\hat{\boldsymbol{n}} = \sum_{\boldsymbol{k}}\hbar\boldsymbol{k}\cdot\hat{\boldsymbol{n}}c^{\dagger}_{\boldsymbol{k}+\boldsymbol{q}}c_{\boldsymbol{k}} \tag{3.55}$$

Therefore the state $|\ n)$ corresponds to the one in which we have taken an electron in state $\boldsymbol{k}$ inside the Fermi surface and moved it to the state $\boldsymbol{k}+\boldsymbol{q}$ outside the Fermi surface. Equation (3.54) becomes

$$\left\langle \boldsymbol{p}_{\boldsymbol{q}}\cdot\hat{\boldsymbol{n}}\right\rangle\hat{\boldsymbol{n}} = \frac{2e}{cV}\sum_{\substack{k<k_F\\ |\boldsymbol{k}+\boldsymbol{q}|>k_F}}\frac{(\boldsymbol{k}\cdot\hat{\boldsymbol{n}})^2}{\boldsymbol{k}\cdot\boldsymbol{q}+q^2/2}\boldsymbol{A}(\boldsymbol{q}). \tag{3.56}$$

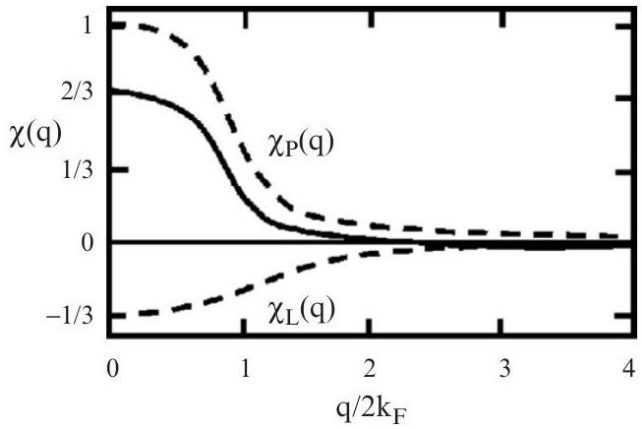


Fig. 3.3. Wave-vector dependence of the Landau and Pauli susceptibilities. The sum is shown by the solid curve

Converting the sum to an integral in (3.56) leads to the susceptibility

$$\chi_L(\boldsymbol{q}) = \chi_L \frac{3k_F^2}{2q^2}\left[1 + \frac{q^2}{4k_F^2} - \frac{k_F}{q}\left(1 - \frac{q^2}{4k_F^2}\right)^2 \ln\left|\frac{2k_F + q}{2k_F - q}\right|\right], \tag{3.57}$$

where

$$\chi_L = -\frac{(N/V)e^2}{4mc^2k_F^2} = -\frac{e^2k_F}{12\pi^2mc^2}. \tag{3.58}$$

The second expression was arrived at by using $N/V = k_F^3/3\pi^2$, which assumes that each point in k space has a twofold spin degeneracy. This result for $q = 0$ was first obtained by Landau, and as we shall see in the next section, it is $-\frac{1}{3}$ the Pauli susceptibility. The q dependence of $\chi_L(\boldsymbol{q})$ is shown in Fig. 3.3.

As we mentioned above, Peierls considered the effect of a periodic potential upon the orbital susceptibility. His result, which neglects interband effects, takes the form [41].

$$\chi_{LP} = -\frac{e^2}{48\pi^3\hbar^2c^2}\int\left[\frac{\partial^2\epsilon}{\partial k_x^2}\frac{\partial^2\epsilon}{\partial k_y^2} - \left(\frac{\partial^2\epsilon}{\partial k_x\partial k_y}\right)^2\right]\frac{dS}{|\nabla_k\epsilon|}. \tag{3.59}$$

If $\epsilon(\boldsymbol{k}) = \hbar^2k^2/2m^*$ we obtain the factor m/m^* mentioned above.
The general form of the orbital susceptibility of noninteracting Bloch electrons was established in the early 1960s. However, the formulas were quite complicated. In 1970 Fukayama [45] reformulated these results in a way which allows calculations to be carried out for realistic energy bands. The susceptibility is found to be a sensitive function of the position of the Fermi level relative to critical points in the Brillouin zone. This is illustrated by Fig. 3.4 which shows the susceptibility, in units of the Landau susceptibility, for a simple model which is periodic in the x direction and free-electron-like in the other directions. The large diamagnetism of bismuth is due to such an interband effect and has the same physical interpretation of large interatomic circulating currents that give rise to large g values (see p. 79).

In dealing with electrons in bands other than s bands the possibility arises of a paramagnetic orbital contribution. Such a contribution was first pointed

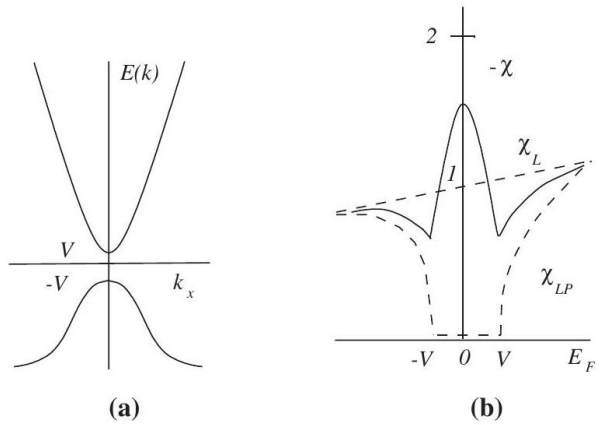


Fig. 3.4. Orbital susceptibility (b) as a function of the Fermi level for the band structure shown in (a)

out by Kubo and Obata [46] and has the same origin as the Van Vleck susceptibility in the case of localized electrons. That is, it is a second-order effect involving the orbital part of the Zeeman interaction. In a metal this takes the form

$$\chi_{\text{orb}} = \frac{2\mu_B^2}{(2\pi)^3} \int d\boldsymbol{k} \sum_{m' \neq m} \frac{f(\epsilon_{mk}) - f(\epsilon_{m'k})}{\epsilon_{m'k} - \epsilon_{mk}} \left| \langle m'\boldsymbol{k}| \boldsymbol{L} |m\boldsymbol{k} \right\rangle |^2 \tag{3.60}$$

where ϵ_{mk} is the energy of an electron in the band m with wave vector $\boldsymbol{k}$. The factor 2 in the numerator arises from the two spin states.

Finally, it is worth mentioning that in a superconductor the presence of a gap in the energy spectrum leads to the result that $\langle \boldsymbol{p_q} \cdot \hat{\boldsymbol{n}} \rangle = 0$. Therefore

$$\chi(\boldsymbol{q})_{\text{super}} = -\frac{1}{4\pi} \frac{\omega_p^2/c^2}{q^2 + \omega_p^2/c^2}, \tag{3.61}$$

where $\omega_p^2 = 4\pi ne^2/m$ is the plasma frequency. The corresponding screening length is the London penetration length $(mc^2/4\pi ne^2)^{1/2}$. As $q \to 0$ this reduces to the susceptibility of a perfect diamagnet, $-1/4\pi$.

3.2.2 The de Haas-van Alphen Effect

Just as in the ionic case, when the temperature becomes very low or the field very large, such that $\mu_B H > k_B T$, nonlinear effects begin to appear. In the case of metallic diamagnetism this nonlinearity manifests itself as a periodic variation of the induced moment as a function of $1/H$. Such behavior was
suspected by Landau [40], worked out in detail by Peierls [41], and observed by de Haas and van Alphen [47]. This de Haas-van Alphen effect, as it is now known, has become important in determining the nature of Fermi surfaces.

The perturbation approach used to compute the Landau susceptibility is not sufficient to describe the de Haas-van Alphen effect. Therefore, we shall compute the susceptibility from the free energy which is, in turn, obtained from an explicit knowledge of the energy levels. It should be mentioned that Sondheimer and Wilson [43] have shown that for a degenerate electron system the free energy is given by the inverse Laplace transform of the classical partition function $Z(\beta)$ where $\beta = 1/k_B T$ is regarded as a complex variable. The interesting feature of this approach is that $Z(\beta)$ contains a branch point at the origin which is responsible for the

Landau diamagnetism and a row of poles on the imaginary axis that lead to the de Haas-van Alphen oscillations.

The Schrödinger equation for an electron in a uniform magnetic field is

$$\frac{1}{2m}\left(\boldsymbol{p}-\frac{e}{c}\boldsymbol{A}\right)^2\psi = E\psi. \tag{3.62}$$

Landau noticed that it is particularly convenient to choose the gauge $\boldsymbol{A} = \boldsymbol{H}\times\hat{\boldsymbol{y}}$. Equation (3.62) then becomes

$$\frac{\partial^2\psi}{\partial x^2}+\left(\frac{\partial}{\partial y}-i\frac{eH}{\hbar c}x\right)^2\psi+\frac{\partial^2\psi}{\partial z^2}+\frac{2mE}{\hbar^2}\psi = 0. \tag{3.63}$$

If we assume a solution of the form

$$\psi = \exp\left[i\left(k_y y + k_z z\right)\right]\varphi(x), \tag{3.64}$$

then $\varphi(x)$ must satisfy

$$\frac{\partial^2\varphi}{\partial x^2}+\left[\frac{2mE}{\hbar^2}-k_z^2-\left(k_y-\frac{eH}{\hbar c}x\right)^2\right]\varphi = 0. \tag{3.65}$$

But this is just the Schrödinger equation for a one-dimensional harmonc oscillator centered at $x_0 = \hbar c k_y/eH$. Therefore we may immediately write the total eigenfunction as

$$\psi_{n,k_y,k_z} = \frac{N_n}{\sqrt{L_y L_z}}\exp\left[i\left(k_y y+k_z z\right)\right]\exp\left[-\left(\alpha^2/2\right)\left(x-x_0\right)^2\right]H_n\left[\alpha\left(x-x_0\right)\right], \tag{3.66}$$

where N_n is a normalization factor, H_n is a Hermite polynomial, and $\alpha = \left(m\omega_c/\hbar\right)^{1/2}, \omega_c$ being the cyclotron frequency $|e|H/mc$. This has the eigenvalue

$$E_{n,k_y,k_z} = \left(n+\frac{1}{2}\right)\hbar\omega_c+\frac{\hbar^2 k_z^2}{2m}. \tag{3.67}$$

We could now proceed to compute the moment by means of the partition function (3.7). However, in computing the trace we must be careful to preserve
the total number of electrons. This was not a problem with insulators, since each of the ions or electrons was distinguishable. In a metal, however, we have no ionic cores on which to "trap" electrons and thereby identify them. Thus we are left with a system of N truly identical particles. In such a situation it is convenient to sum over all possible states, irrespective of the number of particles involved, but to include in the distribution function a factor that ensures that the total number of particles is conserved. In mathematics such a factor is called a Lagrange multiplier. In our particular application it is referred to as the chemical potential. Thus we shall take as the equilibrium density matrix [36]

$$\rho = \exp[\beta(\Omega-\mathcal{H}+\lambda\mathcal{N})], \tag{3.68}$$

where Ω is the thermodynamic potential

$$\Omega = -k_B T\ln\mathcal{Q} \tag{3.69}$$

and $\mathcal{Q}$ is the grand partition function

$$\mathcal{Q} = \mathrm{Tr}\{\exp[-\beta\mathcal{H}-\lambda\mathcal{N}]\}. \tag{3.70}$$

Here $\mathcal{N}$ is the total number operator and λ is the chemical potential, which is to be determined by the condition

$$\langle \mathcal{N} \rangle = N. \tag{3.71}$$

The trace in (3.70) is taken without any restriction on the total number of particles. Equation (3.68) is referred to as the grand canonical distribution.

Since the total number of particles is not dependent on the magnetic field, we generalize (1.22) to

$$M = \frac{k_B T}{V} \frac{\partial}{\partial H} \ln \Omega. \tag{3.72}$$

In computing the trace indicated in (3.70) we shall take as our basis states those of the form (1.106). Since there is no restriction on the total number of particles, there will be a state with no particles, $|0_1, 0_2, \ldots, 0_i, \ldots, 0_\infty\rangle$; infinitely many with one particle, $|1_1, 0_2, \ldots, 0_i \ldots\rangle$; etc. Here i stands for the set of quantum numbers $\{n, k_y, k_z\}$. Thus $\mathcal{Q}$ takes the form

$$\mathcal{Q} = 1 + \sum_i \exp\left[-\beta\left(E_i - \lambda\right)\right] + \sum_{i \neq j} \exp\left[-\beta\left(E_i - \lambda\right)\right] \exp\left[-\beta\left(E_j - \lambda\right)\right] + \ldots \tag{3.73}$$

This may be rewritten as

$$\mathcal{Q} = \prod_i \left(1 + \exp\left[-\beta\left(E_i - \lambda\right)\right]\right). \tag{3.74}$$

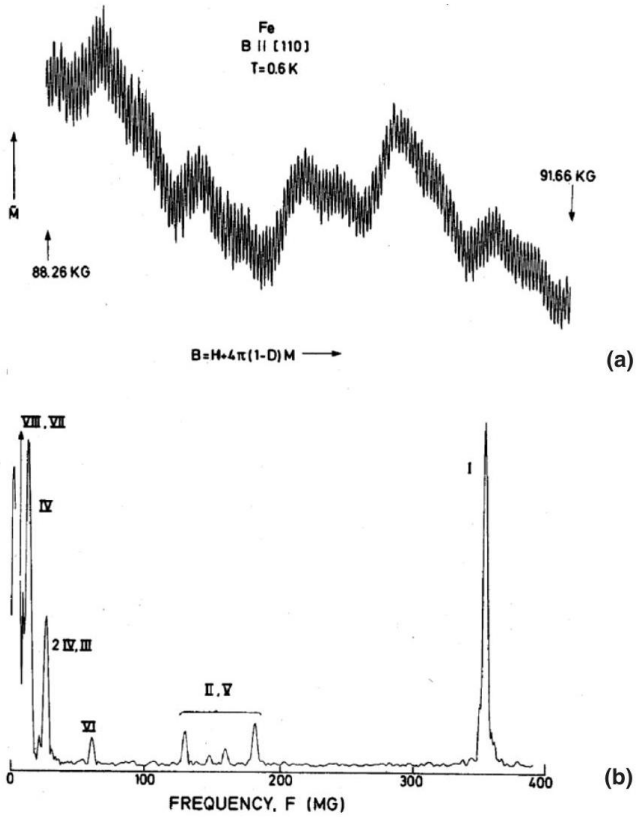


Fig. 3.5. Oscillatory component of the diamagnetic moment in iron (a) with its frequency spectrum (b) (courtesy of G.G. Lonzarich)

Therefore the thermodynamic potential becomes

$$\Omega = -k_B T \sum_i \ln\left(1 + \exp\left[-\beta\left(E_i - \lambda\right)\right]\right). \tag{3.75}$$

In the absence of a magnetic field the states would be characterized by their wave vector $\boldsymbol{k}$, and the sum in (3.75) would be converted to an integral over $\boldsymbol{k}$ space. In the presence of the field, however, we find that the states are characterized by n, k_y, and k_z. What has happened is that the original distribution of points in the $k_x k_y$ plane has condensed onto "tubes" parallel to the k_z axis labeled by the quantum number n. Since this total number of states must remain the same, we find that the number of states along a length dk_z of such a tube is $(eHV/2\pi^2\hbar c)\, dk_z$. Therefore the sum over i in (3.75) may be replaced by a sum over n plus an integral along k_z [48],

$$\Omega = k_B T \left(\frac{eHV}{2\pi^2\hbar c}\right) \int_{-\infty}^{\infty} dk_z \sum_{n=0}^{\infty} \ln\left(1 + \exp\left\{-\beta\left[E\left(n, k_z\right) - \lambda\right]\right\}\right). \tag{3.76}$$

The resulting thermodynamic potential leads to two contributions to the susceptibility. The first is just the Landau result we have already obtained. The second contribution to the thermodynamic potential has the oscillatory form

$$\Omega^{(2)} \simeq k_B T \sum_{s=1}^{\infty} (-1)^s \left(\frac{eH}{2\pi s c\hbar}\right)^{3/2} \frac{\left|A_0''\right|^{-1/2}}{\sinh\left(2\pi^2 s k_B T/\hbar\omega_c\right)} \cos\left(\frac{sc\hbar}{eH} A_0 \pm \frac{\pi}{4}\right). \tag{3.77}$$

In this expression A_0 is the extremal cross-sectional area of the Fermi surface at some height $k_z = k_0$. The expansion of the cross-sectional area about this point has the form

$$A = A_0 \pm \frac{1}{2}\left(k_z - k_0\right)^2 A_0'' + \ldots, \tag{3.78}$$

which defines A_0''. Equation (3.77) implies that the magnetic susceptibility has a contribution which oscillates as a function of $1/H$. Furthermore, the period of this oscillation directly measures the extremal area of the Fermi surface perpendicular to the direction of the magnetic field.

In Fig. 3.5a we show the oscillatory component of the magnetization in iron. Since both the "up" and "down" spins have their own Fermi surfaces, and since both have a complex topology, there are numerous areas contributing to the de Haas-van Alphen effect. These are identified by frequency analyzing the magnetization spectrum as shown in Fig. 3.5b. The cross-sectional area in atomic units is related to the frequency in megagauss by $A = 2.673 \cdot 10^{-3}$ F. Combining such data with band calculations enables one to construct the Fermi surface.

3.2.3 Quantized Hall Conductance

With the development of thin film growth techniques it is possible to fabricate devices capable of confining electrons to a region in which one of the dimensions is small compared with the de Broglie wavelength. Under these conditions the electrons behave like a two-dimensional electron gas. If a magnetic field is applied perpendicular to this two-dimensional gas the energy spectrum analogous to (3.67) consists of a ladder of discrete Landau levels with energy gaps between them. This has dramatic consequences that are not found in 3-D, or even 1-D, systems. The equation of motion for an electron in electric and magnetic fields is

$$m^* \frac{d\boldsymbol{v}}{dt} = e\left(\boldsymbol{\mathcal{E}} + \frac{1}{c}\boldsymbol{v} \times \boldsymbol{H}\right) - m^* \frac{\boldsymbol{v}}{\tau} \tag{3.79}$$

In steady state ($d\boldsymbol{v}/dt = 0$) and in high magnetic fields ($\omega_c\tau \gg 1$), this equation reduces to

$$0 = e\mathcal{E}_x + \frac{eH}{c}v_y. \tag{3.80}$$

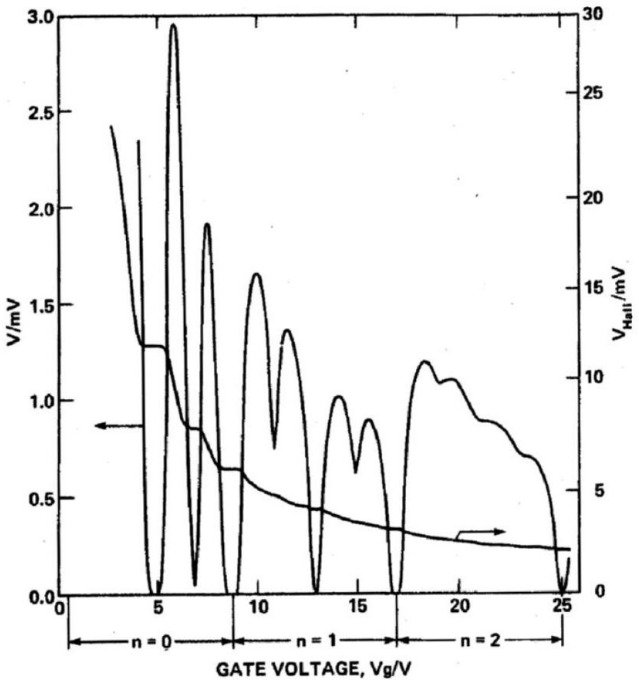


Fig. 3.6. Longitudinal and Hall voltages of a silicon MOSFET with a source drain current of \$1 \mu \mathrm{˜A
\$ as a function of gate voltage in a magnetic field of 18 T . The index n refers to the Landau level occupied. Structure within a given level is due to spin and band structure degeneracies [49]}

The current density in the y direction is Nev_y where N is the density of electrons. Consequently, the Hall conductivity is

$$\sigma_{xy} = \frac{j_y}{\mathcal{E}_x} = -\frac{Nec}{H}. \tag{3.81}$$

If we assume that all the electrons are describable by the conducting Landau states and that the Fermi level is somehow pinned between the n and $n+1$ Landau levels then the density N is simply the product of the number of filled Landau levels, n, times the degeneracy of each level, eH/hc, per unit area. Thus the Hall conductivity becomes

$$-\sigma_{xy} = n\frac{e^2}{h}, \tag{3.82}$$

that is, it is quantized in multiples of e^2/h. Furthermore, as long as the Fermi level lies between the Landau levels, the density of states is zero and $\sigma_{xx} = 0$. The fact that $\sigma_{xx} = 0$ while $\sigma_{xy} \neq 0$ means that the diagonal resistivity

$$\rho_{xx} = \frac{\sigma_{xx}}{\sigma_{xx}^2 + \sigma_{yy}^2} \tag{3.83}$$

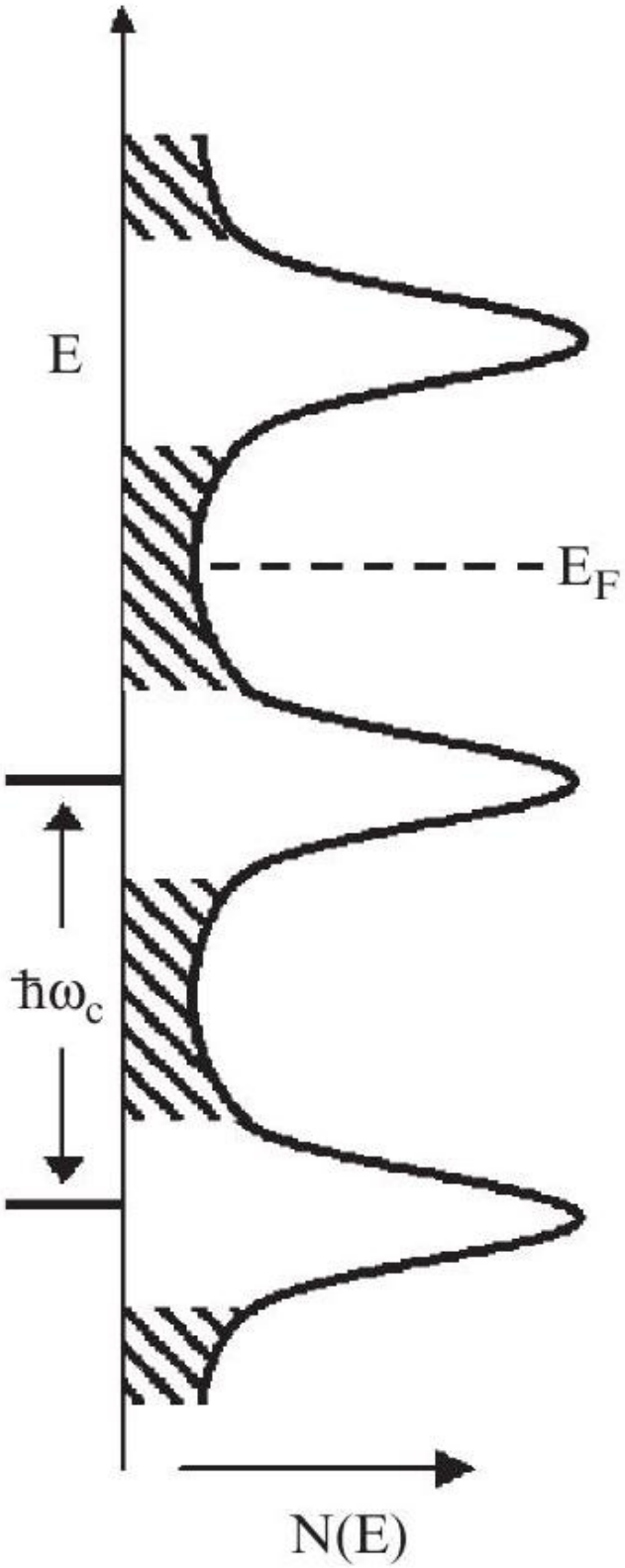


Fig. 3.7. Density of states showing broadened Landau levels and localized states (shaded)

also vanishes and the two-dimensional system is in a zero-resistance state!
This two-dimensional behavior has, in fact, been observed by von Klitzing et al. [49] in the inversion layer of a silicon MOSFET (metal-oxidesemiconductor field-effect transistor). As one changes the gate voltage on such a device, the Fermi level is swept through the Landau levels. Figure 3.6 shows the longitudinal voltage V associated with a constant current as well as the

Hall voltage V_{Hall} as functions of the gate voltage. The plateaus in the Hall voltage correspond to a Hall resistance $R_H = h/e^2n$ to an accuracy of better than one part in 10^5. Tsui and Gossard [50] have also observed this effect in gallium arsenide heterojunctions as a function of an applied magnetic field.

What is puzzling about these observations is the extraordinary accuracy with which the plateaus in σ_{xy} give e^2/h. The fact that the plateaus exist suggests that there are localized states between the Landau levels through which the Fermi level can be moved, as illustrated in Fig. 3.7. However, since such states do not carry current, we might have expected σ_{xy} to be reduced from its ideal value. Prange [51] as well as Tsui and Allen [52], have shown that even if localized states exist, the remaining nonlocalized states carry an extra Hall current which exactly compensates for that not carried by the localized states. Then there are also interaction effects and edge effects to worry about. Laughlin [53] pointed out that the quantization is more fundamental than

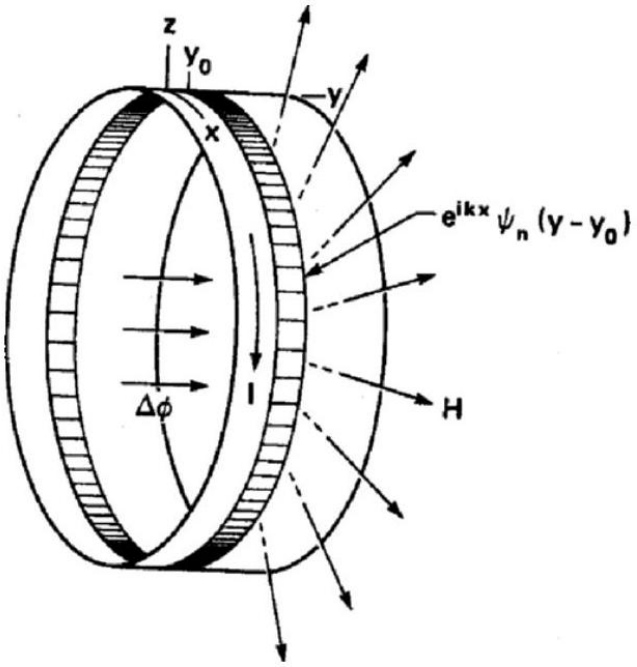


Fig. 3.8. Electron wavefunction associated with a metal loop in a radial magnetic field

that suggested above, being the consequence of gauge invariance. Consider a conducting strip connected in a ring of circumference L, as shown in Fig. 3.8. Let us assume there is a magnetic field normal to the strip. The electron wave functions are then plane waves in the x direction, with a harmonic oscillator envelope in the y direction centered at y_0. If we now change the flux $\Delta\phi$ which threads the ring, this corresponds to a change in the vector potential in the direction around the ring; i.e., $\boldsymbol{A} \to \boldsymbol{A} + \Delta\boldsymbol{A} = \boldsymbol{A} + (\Delta\phi/L)\hat{\boldsymbol{x}}$. This gauge change shifts the center of the wave function by an amount $\Delta A/H_0$. In particular, if the flux change is a quantum of flux, hc/e, then the shift moves each state precisely to an adjacent center. The net effect is to move one state per Landau level from one edge to the other. Assuming there is a voltage V across the strip, this corresponds to a total increase in energy of neV, where n is the number of filled Landau levels. This change in energy ΔU is related to the current flowing around the strip by the general relation

$$I = \frac{c\Delta U}{\Delta\phi} = \frac{\text{nec}\,V}{(hc/e)} = \frac{ne^2V}{h}. \tag{3.84}$$

Thus the Hall conductance I/V has the quantized form found above. Any localized states present will have their phase shifted by the gauge transformation but their energy will be unchanged. Thus they will not contribute to the Hall currents. Furthermore, since the voltage V is by definition the difference in the Fermi levels at the edge, any edge effects are automatically incorporated.

If one plots the Hall resistance as a function of magnetic field the 2-D resistance follows the linear relation, $R_H = B/Ne$, while the quantum Hall resistance consists of steps along this line with the values $(h/e^2)\,(1/n)$. At a
field of approximately 10 T in a GaAs/GaAl As heterostructure the level with $n = 1$ is reached. This corresponds to all the electrons being in the lowest Landau level. One might think that the Hall resistance would simply remain at the value h/e^2 for higher fields. Surprisingly, however, the Hall resistance continues to exhibit plateaus given by

$$R_H = \left(h/e^2\right)(p/q) \tag{3.85}$$

where p and q are integers [54]. While the "integer" quantum Hall effect, where p/q is an integer, depends upon the localized states between the Landau levels as described above, the "fractional" quantum Hall effect (FQHE) is quite different. To understand the fractional quantum Hall effect we recall that a magnetic field consists of discrete flux quanta. The magnitude of the flux quantum is $\phi_0 = h/e$. The plateaus of the FQHE correspond to the system being in a correlated, many-body state in which flux quanta are bound to electrons. In the state $p/q = 1/3$, for example, three flux quanta are bound to each electron. For more details we refer the reader to [55].

3.2.4 Pauli Paramagnetism

Let us now turn our attention to the spin response of a free-electron system. As mentioned at the beginning of this chapter, we can construct the response to an arbitrary static field if we know the response to a single Fourier component. Therefore, let us apply an external field of the form $\boldsymbol{H}\cos(\boldsymbol{q}\cdot\boldsymbol{r})$. The interaction with the electron spins is

$$\mathcal{H} = \sum_i g\mu_B \boldsymbol{s}_i \cdot \boldsymbol{H}\cos\left(\boldsymbol{q}\cdot\boldsymbol{r}_i\right). \tag{3.86}$$

According to (1.22), the moment induced by the field may be obtained by computing the change in the energy of the system that is proportional to the field. This we shall do by perturbation theory. Notice that if the quantization axis of our system is chosen to lie along the direction of $\boldsymbol{H}$, then when $\boldsymbol{q} = 0$, (3.86) has first-order diagonal matrix elements. Therefore a perturbation calculation would have to distinguish between $\boldsymbol{q} = 0$ and $\boldsymbol{q} \neq 0$. In order to avoid this difficulty we shall assume that the field lies in the x direction, $\boldsymbol{H} = H\hat{\boldsymbol{x}}$. Then

$$\mathcal{H} = \frac{g\mu_B H}{4}\sum_i \left(s_i^+ + s_i^-\right)\left[\exp\left(i\boldsymbol{q}\cdot\boldsymbol{r}_i\right) + \exp\left(-i\boldsymbol{q}\cdot\boldsymbol{r}_i\right)\right]. \tag{3.87}$$

Second quantizing this interaction by the prescription given in Chap. 1 gives

$$\mathcal{H} = \frac{g\mu_B H}{4}\sum_{\boldsymbol{k}}\left(a^\dagger_{\boldsymbol{k}+\boldsymbol{q},\uparrow}a_{\boldsymbol{k}\downarrow} + a^\dagger_{\boldsymbol{k}+\boldsymbol{q},\downarrow}a_{\boldsymbol{k}\uparrow} + a^\dagger_{\boldsymbol{k}-\boldsymbol{q},\uparrow}a_{\boldsymbol{k},\downarrow} + a^\dagger_{\boldsymbol{k}-\boldsymbol{q},\downarrow}a_{\boldsymbol{k}\uparrow}\right). \tag{3.88}$$

We see immediately that there are no first-order diagonal matrix elements. The second-order correction to the energy is

$$\Delta E = \left(\frac{g\mu_B H}{4}\right)^2 \sum_{\boldsymbol{k},\sigma} \frac{\left|\left\langle a^\dagger_{\boldsymbol{k}+\boldsymbol{q},-\sigma} a_{\boldsymbol{k}\sigma}\right\rangle\right|^2}{\epsilon_{\boldsymbol{k}\sigma} - \epsilon_{\boldsymbol{k}+\boldsymbol{q},-\sigma}} + (\text{ terms involving } -\boldsymbol{q}). \tag{3.89}$$

The sum over $\boldsymbol{k}$ is restricted to those values of $\boldsymbol{k}$ inside the Fermi sphere such that $\boldsymbol{k}+\boldsymbol{q}$ lies outside the Fermi sphere. Since the unperturbed energies are independent of spin, (3.89) reduces to

$$\Delta E = -\frac{g^2\mu_B^2 H^2}{8} \sum_{\boldsymbol{k}} \left[\frac{f_{\boldsymbol{k}}\left(1 - f_{\boldsymbol{k}+\boldsymbol{q}}\right)}{\epsilon_{\boldsymbol{k}+\boldsymbol{q}} + \epsilon_{\boldsymbol{k}}} + \frac{f_{\boldsymbol{k}}\left(1 - f_{\boldsymbol{k}-\boldsymbol{q}}\right)}{\epsilon_{\boldsymbol{k}-\boldsymbol{q}} - \epsilon_{\boldsymbol{k}}}\right], \tag{3.90}$$

when the sum now runs over all $\boldsymbol{k}$ and $f_{\boldsymbol{k}}$ is the Fermi function

$$f_{\boldsymbol{k}} = \frac{1}{\exp\left[\left(\epsilon_{\boldsymbol{k}} - \epsilon_F\right)/k_B T\right] + 1}. \tag{3.91}$$

Since the sum does run over all $\boldsymbol{k}$, we may write $\boldsymbol{k} \to \boldsymbol{k}+\boldsymbol{q}$ in the second term. The two terms then combine to give

$$\sum_k \frac{f_k - f_{k+q}}{\epsilon_{k+q} - \epsilon_k}.$$

Since

$$\Delta E = -\frac{1}{2}\int d\boldsymbol{r}\boldsymbol{M}(\boldsymbol{r})\cdot\boldsymbol{H}(\boldsymbol{r}) = -\frac{H}{2}\int d\boldsymbol{r}\boldsymbol{M}(\boldsymbol{r})\cos(\boldsymbol{q}\cdot\boldsymbol{r}) \tag{3.92}$$

the derivative of $-\Delta E$ with respect to H gives $\boldsymbol{M}(\boldsymbol{q})$ directly. Therefore, with $H(\boldsymbol{q}) = HV/2$, the susceptibility becomes

$$\chi(\boldsymbol{q}) = \frac{g^2\mu_B^2}{V} \sum_{\boldsymbol{k}} \frac{f_{\boldsymbol{k}} - f_{\boldsymbol{k}+\boldsymbol{q}}}{\epsilon_{\boldsymbol{k}+\boldsymbol{q}} - \epsilon_{\boldsymbol{k}}}. \tag{3.93}$$

The sum in (3.93) may be evaluated by converting it into an integral. At $T = 0$ K we obtain

$$\chi(\boldsymbol{q}) = \frac{3g^2\mu_B^2(N/V)}{8\epsilon_F} F\left(\frac{q}{2k_F}\right), \tag{3.94}$$

where

$$F\left(\frac{q}{2k_F}\right) = \frac{1}{2} + \frac{k_F}{2q}\left(1 - \frac{q^2}{4k_F^2}\right)\ln\left|\frac{2k_F + q}{2k_F - q}\right| \tag{3.95}$$

and N is the total number of electrons in the system. Notice that

$$\lim_{q\to 0} F\left(\frac{q}{2k_F}\right) = 1. \tag{3.96}$$

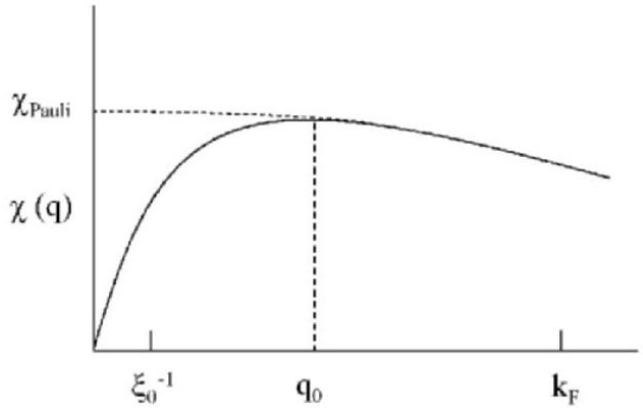


Fig. 3.9. Wave-vector dependence of the spin susceptibility in a superconductor

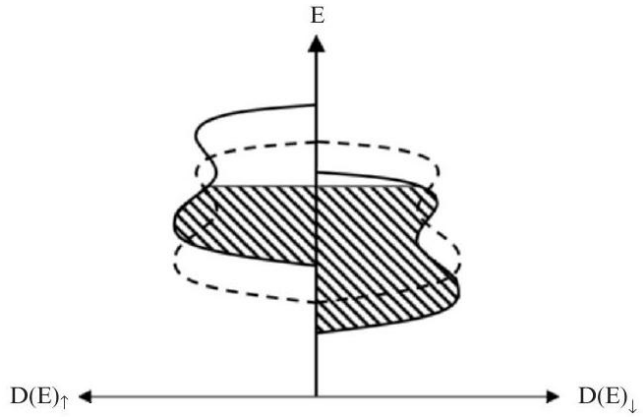


Fig. 3.10. Relative splitting of a d band in the presence of an external field

With $g = 2$ and $\mu_B = e\hbar/2mc$, (3.85) may be rewritten as

$$\chi(q) = \chi_{\text{Pauli}}\, F\left(\frac{q}{2k_F}\right), \tag{3.97}$$

where

$$\chi_{\text{Pauli}} = \frac{3(N/V)e^2}{2k_F}. \tag{3.98}$$

The Pauli susceptibility, given by (3.98), describes the spin response of an electron system to a uniform field. To see this, note that for a uniform field $\boldsymbol{H}(\boldsymbol{q}) = \boldsymbol{H}\Delta(\boldsymbol{q})$. Therefore

$$\boldsymbol{M}(\boldsymbol{r}) = \sum_{\boldsymbol{q}} e^{i\boldsymbol{q}\cdot\boldsymbol{r}}\chi(\boldsymbol{q})\boldsymbol{H}(\boldsymbol{q}) = \chi_{\text{Pauli}}\boldsymbol{H} \tag{3.99}$$

The wave-vector dependence of $\chi(\boldsymbol{q})$ is sketched in Fig. 3.3. Notice that this function has an infinite derivative when $q = 2k_F$. The susceptibility (3.97) was derived under the assumption that the temperature was small in comparison with the Fermi temperature. Since Fermi temperatures are typically of the

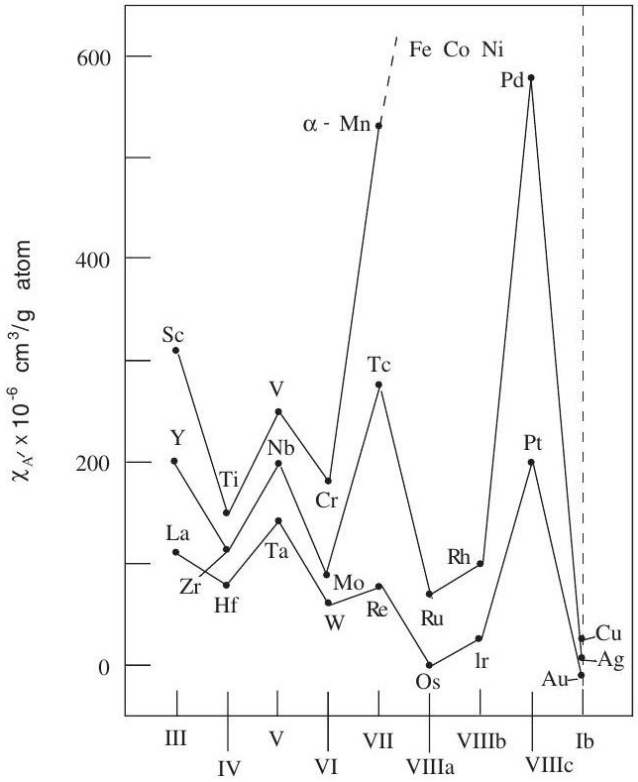


Fig. 3.11. The susceptibility of the transition metals at room temperature [56]

order of 10,000 K, this is generally a good approximation. However, there are situations, as in metals like palladium and in nondegenerate semiconductors, where the effective Fermi temperature is much lower. It is then possible to reach temperatures where the electrons behave like Boltzman particles and exhibit a Curie-like susceptibility.

It is also interesting to consider how the spin susceptibility is changed if the metal becomes superconducting. In a superconductor electrons with opposite spin are correlated over a distance ξ_0 called the coherence length. This suppresses the magnetic response of such pairs at wave vectors below ξ_0^{-1}. When this effect is combined with the decreasing response beyond $2k_F$ we obtain a maximum in the susceptibility at $q_0 = (2\pi k_F^2/\xi_0)^{1/3}$ as shown in Fig. 3.9.

As with the orbital susceptibility, the periodicity of the lattice also modifies the spin susceptibility. However, in the limit of weak spin-orbit coupling one can talk about spin-up and spin-down bands and the Pauli susceptibility takes a particularly simple and useful form. Let us consider a metal whose band structure gives rise to a density of states, i.e., number of states per unit energy per unit volume, as illustrated in Fig. 3.10. To lowest order, the effect of the applied field is to shift the relative energy distributions of the up and down spins. Since the electrons do, in fact, interact with one another, the up- and down-spin distributions will reach equilibrium in such a manner that their Fermi levels are the same. Therefore, as we see from Fig. 3.10 there will be more spins pointing anti-parallel to the field than parallel to it. Since the

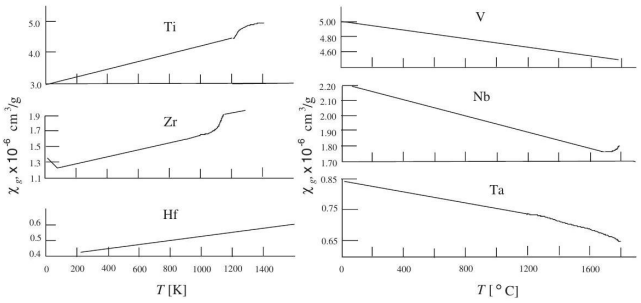


Fig. 3.12. The temperature dependence of the specific susceptibility of selected transition metals [57]. The structure in the data for Ti and Zr near 1200 K is associated with structural phase transitions

gyromagnetic ratio of an electron is negative, this leads to a net moment in the direction of the field. To find this moment we need only compute the difference in the number of up- and down-spin electrons,

$$M = \frac{1}{2} g\mu_B \int_{-\infty}^{\infty} \left[\mathcal{D}(\epsilon) f\left(\epsilon - \frac{1}{2} g\mu_B H\right) - \mathcal{D}(\epsilon) f\left(\epsilon + \frac{1}{2} g\mu_B H\right) \right] d\epsilon \tag{3.100}$$

We expand the Fermi functions about $H = 0$ to yield

$$M \simeq \frac{1}{2} g^2 \mu_B^2 H \int_0^{\infty} \left(-\frac{\partial f}{\partial \epsilon}\right) \mathcal{D}(\epsilon) d\epsilon \tag{3.101}$$

At low temperatures the Fermi function is essentially a step function at the Fermi level. Therefore its derivative is a delta function. Thus (3.101) leads to the susceptibility

$$\chi_{\text{Pauli}} = \frac{1}{2} g^2 \mu_B^2 \mathcal{D}\left(\epsilon_F\right). \tag{3.102}$$

For a free-electron system

$$\mathcal{D}\left(\epsilon_F\right) = \frac{m k_F}{2\pi^2 \hbar^2}, \tag{3.103}$$

and (3.102) reduces to our previous result (3.98).
From (3.102) we would expect the Pauli susceptibility to be roughly ten times larger when the Fermi surface lies within a d band than when it lies

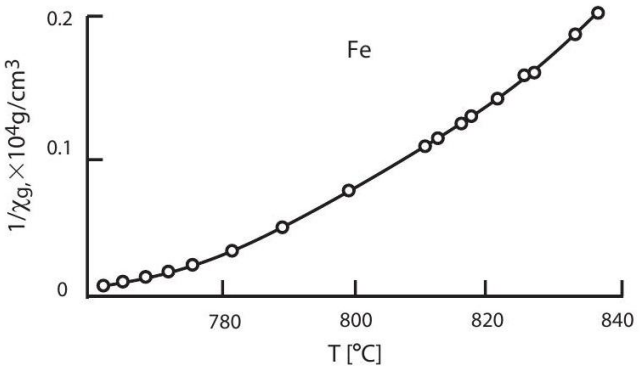


Fig. 3.13. Temperature dependence of the reciprocal susceptibility of Fe [56]

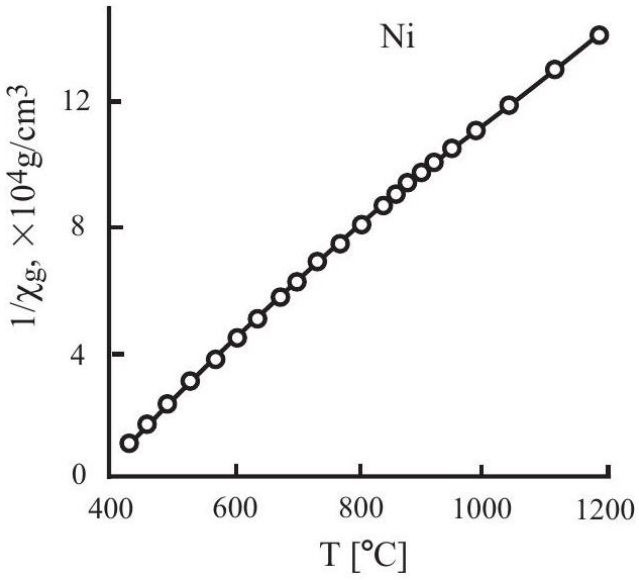


Fig. 3.14. Temperature dependence of the reciprocal susceptibility of Ni [56]

within an s band since d bands are generally narrower with a higher density of states. We might also expect this d-band contribution to exhibit a temperature dependence. The reason for this is that at a temperature T the susceptibility essentially involves an average of $\mathcal{D}(\epsilon_F)$ over a region of the order of k_BT. Therefore, if the width of the structure in the d-band density states is of this order, it will be reflected in the susceptibility. Interactions between the electrons can also lead to a temperature dependence.

Susceptibilities for the various transition metals are shown in Fig. 3.11. First, note that the susceptibility of those metals with unfilled d bands is much larger than for the s-band metals Cu, Ag, and Au . Second, note the variation in the susceptibility across the periodic table. This might be associated with variations in d-band density of states. This rigid-band explanation is consistent with the temperature dependence of these susceptibilities. As shown in Fig. 3.12, the susceptibility of metals such as Ti, Zr, and Hf , which
have a relative minimum in Fig. 3.11, increases with temperature, while for those with a maximum it decreases. Those metals with d^1 and d^9 configurations are further complicated by interaction effects which we shall discuss in the next chapter.

In Mn, Fe, Co, and Ni interactions among the electrons are sufficient to produce an ordered magnetic state. Above their ordering temperature these metals exhibit a relatively strong temperature dependence as shown in Figs. 3.13 and 3.14. This behavior is indicative of some sort of local-moment formation. It is not simply Curie-Weiss, however, for $d(\chi^{-1})/dT$ is not temperature independent.
If band effects may be represented by an effective mass, then

$$\chi_{\text{Pauli}} = g^2\mu_B^2 \frac{m^* k_F}{4\pi^2\hbar^2}. \tag{3.104}$$

Notice that the mass entering the Bohr magneton is unaffected. However, the Landau-Peierls susceptibility (3.59) becomes

$$\chi_{\text{LP}} = -\frac{e^2 k_F}{12\pi^2 m^* c^2} = -\frac{1}{3}\left(\frac{m}{m^*}\right)^2 \frac{e^2\hbar^2 m^* k_F}{m^2 c^2 4\pi^2\hbar^2}. \tag{3.105}$$

Therefore

$$\chi_{\text{LP}} = -\frac{1}{3}\left(\frac{m}{m^*}\right)^2 \chi_{\text{Pauli}}\,. \tag{3.106}$$

Collecting all our results, we may express the total susceptibility of a nonmagnetic metal as

$$\chi(T) = \chi_{\text{core}} + \chi_{\text{LP}} + \chi_{\text{orb}} + \chi(T)_{\text{Pauli}} . \tag{3.107}$$

The temperature dependence arises primarily from the d-band contribution to the Pauli susceptibility.

10.4.3 3.3 Measurement of the Susceptibility

At this point it is appropriate to say a little about measuring the susceptibility. A broad class of techniques use the fact that when a sample with a susceptibility χ_g per gram is placed in a magnetic field gradient it experiences a force. In particular, if the field is in the x direction with a gradient in the z direction, this force is

$$\Delta F_z = \left(H_x \frac{dH_z}{dz} \right) \chi_g \cdot m(gr). \tag{3.108}$$

If the sample is mounted at the end of a pendulum and suspended at a right angle to the field gradient, there will be a torque tending to displace the pendulum. This is the pendulum galvanometer [58]. The torque due to the sample

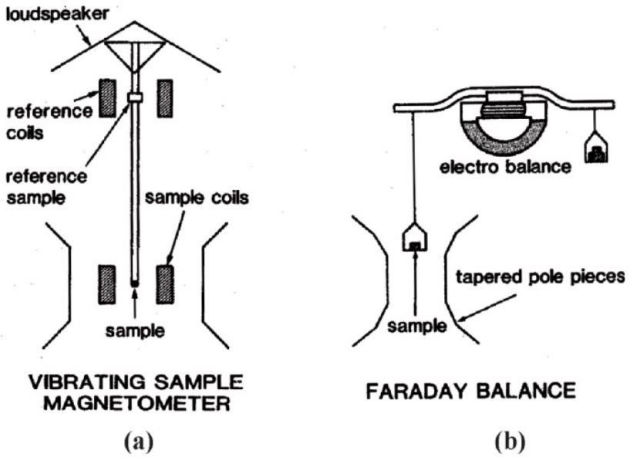


Fig. 3.15. Simplified illustrations of (a) a vibrating-sample magnetometer and (b) a Faraday susceptibility balance

can be balanced by an opposing torque due to current flowing in a solenoid surrounding the sample. The measurement of the susceptibility is thus reduced to measuring the current flow at zero displacement and consequently has a large dynamic range well suited for ferromagnets and strong paramagnets.

For strongly magnetic materials Foner's vibrating-sample magnetometer is particularly convenient (Fig. 3.15a) [59]. In this arrangement the sample is placed at the end of a rod which vibrates up and down inside a pickup coil. If the whole assembly is placed between the pole faces of a magnet, a moment is induced in the sample which, by virtue of its vibration, induces a signal in the pickup coil that is proportional to the magnetization of the sample.

For weakly magnetic materials it is convenient to measure the force directly with an electrobalance. The most popular balance technique today is the Faraday balance (Fig. 3.15b). For this method the pole pieces of the magnet are carefully machined to produce a fairly large region over which $\boldsymbol{H}_x d\boldsymbol{H}_x/dz$ is constant. Samples small with respect to this region are suspended in the region and the force measured. The quantity $\boldsymbol{H}_x d\boldsymbol{H}_x/dz$ is determined by

calibrating the system with a standard such as platinum. A typical microbalance can resolve 5 micrograms. In a related technique, known as the Gouy method, one employes a long uniform sample, frequently a liquid, which extends over the entire field region. Thus the force is the integral of (3.108).

Another way of measuring the susceptibility of small or magnetically weak samples is by a mutual-inductance technique [60]. The sample is mounted inside a pair of secondary coils on top of which is wound a primary coil. Measurement of the mutual inductance of this coil system with and without the sample yields an absolute determination of the susceptibility of the sample.

A very different method for obtaining the static susceptibility makes use of the KramersKronig relation

$$\chi(0) = \frac{2}{\pi}\int_0^\infty \frac{\chi''(\omega)d\omega}{\omega} \tag{3.109}$$

where $\chi(0)$ is measured in relation to $\chi(\infty)$. If the imaginary part of the susceptibility is strongly peaked around a particular frequency ω_0, then

$$\chi(0) \simeq \frac{2}{\pi\omega_0}\int_0^\infty \chi''(\omega)d\omega \tag{3.110}$$

If, as in most cases, the experiment entails varying the magnetic field at a fixed frequency, this becomes

$$\chi(0) = \frac{g\mu_B}{\pi\hbar\omega}\int_0^\infty \chi''(H)dH \tag{3.111}$$

Therefore the static susceptibility is simply given by the area under the magnetic resonance. Since this entails an absolute intensity measurement, it is rather difficult. However, Schumacher and Slichter [61], in determining the spin susceptibility of lithium and sodium, neatly solved this problem by comparing spin resonances with the nuclear resonance in the same sample. The nuclear susceptibility χ_n may be computed from the Langevin susceptibility (3.27). The spin susceptibility is determined by measuring the area under the conduction-electron spin resonance A_e and the area under the nuclear resonance A_n and using the relation

$$\chi_e = \chi_n \frac{\mu_B A_e}{\mu_n A_n} \tag{3.112}$$

There is another technique involving nuclear magnetic resonance which is very helpful in determining the various contributions to the total susceptibility in a metal. As we saw in Sect.2.2, s-like conduction electrons produce a contact hyperfine field at the nuclei of a metal. Thus, if the conduction electrons are polarized by an external field, they will, in turn, shift the nuclear magnetic resonance frequency from what it would be in a diamagnetic environment. This is known as the Knight shift. Since the ability of the conduction electrons to be polarized is determined by their Pauli susceptibility, the resulting Knight shift K will be proportional to this susceptibility. In particular,

$$K = \left(\frac{8\pi}{3}\right)\chi_P\Omega\langle|\psi_S(0)\Big|^2\Big\rangle_F , \tag{3.113}$$

where χ_P is the Pauli susceptibility per unit volume, Ω is the volume of the unit cell, and $\langle|\psi_S(0)^2|\rangle_F$ is the average value of the s-electron probability density at the nucleus for electrons near the Fermi surface.

As was also pointed out in Sect.2.2, d-like electrons do not contribute to the contact hyperfine field. However, their presence can lead to a slight polarization of the otherwise paired core

s electrons. This core polarization depends on the d-band polarization and will therefore contribute a term to

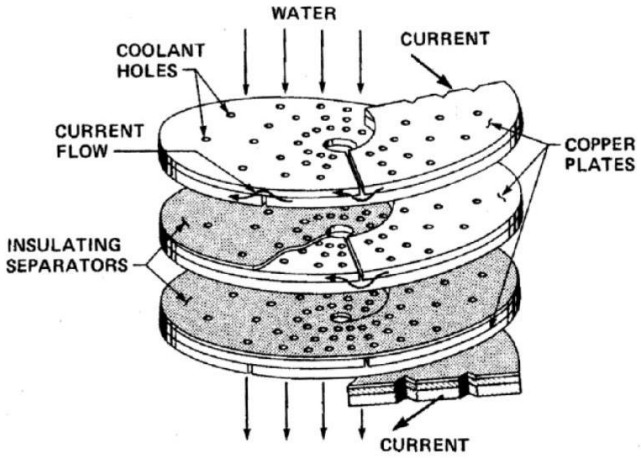


Fig. 3.16. Windings of a Bitter magnet

the Knight shift that is proportional to the d-band susceptibility. It turns out that the core polarization generally leads to a negative Knight shift. Furthermore, since χ_P is temperature dependent, the resulting Knight shift will also be temperature dependent. Thus the Knight shift, together with direct susceptibility measurements, enables us to unravel the various contributions to the total susceptibility given by (3.98). A beautiful example of this is the work of Clogston et al. on metallic platinum [62].

In some cases, we are interested in the magnetic response of a system at very high fields, such as in de Haas-van Alphen studies or studies of magnetic impurity levels in semiconductors. The production of such high fields has become a specialty practiced only within the world's few high-field laboratories, such as the National High Magnetic Field Laboratory in Tallahassee or the joint German-French high-field laboratory in Grenoble.

The classical approach to generating large continuous magnetic fields is to use water-cooled coils. In the late 1930s, Francis Bitter pioneered the design of such high-field magnets. A typical Bitter magnet is illustrated in Fig. 3.16. In order to produce 15 T in a $5-$cm bore, five megawatts of power are required. If the coil is not stress limited, the central field is proportional to the square root of the power dissipated. As a result, the practical limit for such resistive magnets is about 20 T . If one tries to increase the field by reducing the inner diameter, the Lorentz force on the conductor may exceed its tensile strength.

10.4.4 3.4 Local Moments in Metals

When a potentially magnetic atom such as iron is dissolved in a nonmetallic host, the electronic states are localized at the impurity site. The magnetic character of this state is then determined by the intraatomic Coulomb interaction (Hund's rules). If the crystalline fields become large, the magnetic state
may be affected. Fe^{2+}, for example, in an octahedral environment can have either a high-spin ${}^5F_2\left(t_{2g}{}^4e_g^2\right)$ state or a low-spin ${}^1A_a\left(t_{2g}^6\right)$ state. But the lowtemperature susceptibility is readily calculated by the formalism described in the earlier sections of this chapter.

In a metallic host, the situation is much more complex because the electronic states associated with the impurity are often not localized. Even when the impurity appears to contain an integral number of electrons, the fact that they are coupled to a Fermi sea, no matter how weak the coupling, can lead to a correlated nonmagnetic ground state.

3.4.1 Virtual Bound States

If we consider the conduction electrons as plane waves, then a convenient formalism for describing the effect of a single impurity is scattering theory. If the impurity is represented by a real spherical potential (i.e., we neglect inelastic scattering), then the problem has symmetry about the axis defined by the incident wave vector $\boldsymbol{k}$, so the wave function, call it $u_k(r,\theta)$, may be expanded as a sum of products of radial functions and Legendre polynomials,

$$u_k(r,\theta) = \sum_{l=0}^{\infty}(2l+1)i^l R_l(k,r)P_l(\cos\theta) \tag{3.114}$$

The Legendre polynomials are related to the spherical harmonics of Table 2.1 by

$$P_l(\cos\theta) = \sqrt{\frac{4\pi}{2l+1}}Y_l^0(\theta,\varphi)$$

The coefficient $R_l(k,r)$ is called the partial wave amplitude. This is an important quantity for, when the energy variable k is regarded as a complex number, the poles of $R_l(k,r)$ along the positive imaginary axis correspond to bound states. Far from the impurity, this wave function has the form

$$\lim_{r\to\infty} u_k(r,\theta) = e^{ik\cdot r} + f(k,\theta)\left(e^{ikr}/r\right), \tag{3.115}$$

where $f(k,\theta)$ is called the scattering amplitude and is a function of the scattering phase shift δ_l,

$$f(k,\theta) = (1/k)\sum_l [4\pi(2l+1)]^{1/2}\exp(i\delta_l)(\sin\delta_l)\,Y_l^0(\theta). \tag{3.116}$$

Substituting (3.116) into (3.115) shows that $R_l(kr)$ has the asymptotic form

$$R_l(kr) \underset{r\to\infty}{\longrightarrow} \frac{1}{kr}\exp(i\delta_l)\sin\left(kr - \frac{1}{2}l\pi + \delta_l\right) \tag{3.117}$$

The phase shift, δ_l depends upon the energy of the scattering particle and on the details of the impurity potential. If the potential is a delta function of strength V_0, then there is only s-wave scattering and

$$\delta_0(E) = \tan^{-1}\left\{\pi V_0 N(E)/\left[1 - V_0 I(E)\right]\right\} \tag{3.118}$$

where $N(E)$ is the host density of states and

$$I(E) = \mathcal{P}\int dE' N(E')/(E'-E) \tag{3.119}$$

where $\mathcal{P}$ means that we take the principal part of the integral. If we define E_0 as the value of E where $V_0 I(E_0) = 1$ and expand $V_0 I(E)$ about $E = E_0$, then

$$\delta_0(E) = \tan^{-1}\frac{\Delta}{E - E_0}, \tag{3.120}$$

where Δ is proportional to $N(E_0)$. Notice that when the phase shift goes through $\pi/2$, the scattering amplitude goes through a maximum. This is referred to as a resonance or virtual bound state.

Suppose we surround the impurity by a sphere of radius R. If we require that the wavefunction go to zero on the boundary, then, from (3.117)

$$kR - \frac{1}{2}l\pi + \delta_l = n\pi \tag{3.121}$$

where n is an integer. In the absence of scattering $\delta_l = 0$ As the scattering potential is "turned on" δ_l becomes nonzero. If the potential is attractive δ_l is positive and δ_l/π is the number of extra nodes introduced within the sphere. Each of these nodes corresponds to a state that can accomodate $2(2l+1)$ electrons. Therefore, if z is the charge differential between the impurity and the host

$$\sum_l 2(2l+1)\frac{\delta_l}{\pi} = z \tag{3.122}$$

This is known as the Friedel sum rule. The density of states associated with these extra states is

$$\eta(E) = \frac{dz}{dE} = \sum_l \frac{2(2l+1)}{\pi}\frac{d\delta_l}{dE} \tag{3.123}$$

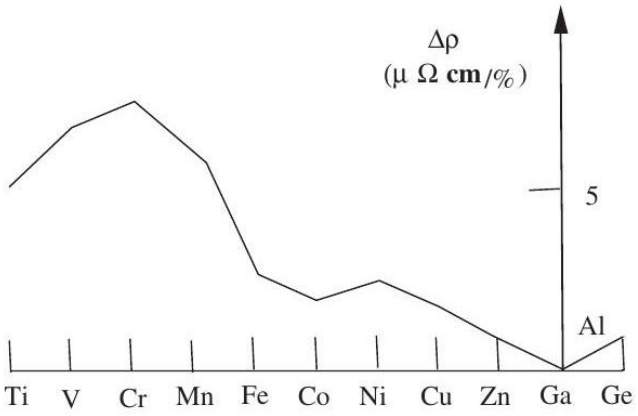


Fig. 3.17. Residual resistivities of transitional impurities in aluminum

For the phase shift given by (3.122) this virtual bound state is characterized by the Lorentzian form

$$\eta(E) = \frac{1}{\pi}\frac{\Delta}{(E-E_0)^2 + \Delta^2}. \tag{3.124}$$

The differential scattering cross section $\sigma(\theta)$ is given by the square of the scattering amplitude, i.e., $\sigma(\theta) = |f(\theta)|^2$. The effective average cross section σ (not to be confused with conductivity) that governs the resistivity is the angular average of $\sigma(\theta)$ weighted by $(1-\cos\theta)$, which measures the relative change in the component of the electron's velocity along its initial direction of motion:

$$\sigma = 2\pi \int_{-\pi}^{\pi} d\theta \sin\theta \sigma(\theta)(1-\cos\theta) \tag{3.125}$$

The mean free path is $\Lambda = 1/x\sigma$, where x is the atomic concentration of impurities. The integral over the $\cos\theta$ term in (3.125) is facilitated by the relation

$$Y_l^0 \cos\theta = \frac{l+1}{\sqrt{(2l+1)(2l+3)}}Y_{l+1}^0 + \frac{l}{\sqrt{(2l+1)(2l-1)}}Y_{l-1}^0$$

and then using the orthogonality of the spherical harmonics. The resistivity associated with the impurities is then

$$\Delta\rho = \frac{4\pi x}{ne^2 k_F} \sum_{l=0}^{\infty} (l+1) \sin^2 \left[\delta_l \left(E_F\right) - \delta_{l+1} \left(E_F\right)\right] \tag{3.126}$$

where n is the density of conduction electrons. Thus, again if we only consider s-wave scattering ($l = 0$), the resistivity will show a peak when the Fermi level and the center of the virtual bound state coincide. There is evidence for such behavior in aluminum as shown in Fig. 3.17. This phase shift description has been successfully applied to many alloys, particularly by Friedel [63] and his collaborators. The problem arises when one begins to consider magnetic properties. Figure 3.18, for example, shows that iron may or may not exhibit a

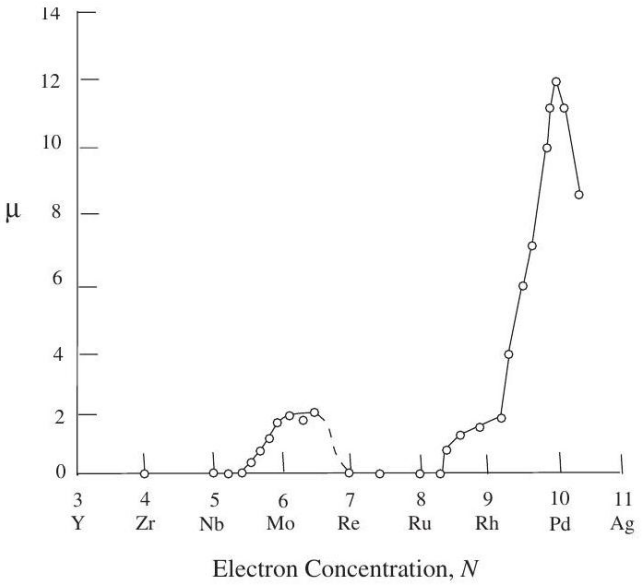


Fig. 3.18. Magnetic moment in Bohr magnetons of an iron atom dissolved in various second-row transition metals (except Re which is a 5 d element) and alloys as a function of electron concentration [64]

magnetic moment when it is dissolved in 4d transition metals and alloys. This behavior attracted the attention of Anderson, who considered how a virtual bound state develops a moment.

3.4.2 Anderson's Theory of Moment Formation

The presence of an impurity in a metallic host leads to additional terms in the Hamiltonian of the system. Since the size or the charge of the impurity may differ from that of the host, there will be a change $V(\boldsymbol{r})$ in the crystal potential. Since this potential is screened by the conduction electrons, it is essentially confined to the impurity cell. In addition, the Coulomb interaction between electrons within the impurity cell, v, will be different from that between electrons within host cells. Most of the hosts we shall consider do not exhibit a significant exchange enhancement, and so we shall neglect the Coulomb interactions within the host cells, retaining it only for the impurity cell. Assuming that these are the major effects of the impurity, we write the total Hamiltonian as

$$\mathcal{H} = \mathcal{H}_0 + \sum_i eV(\boldsymbol{r}_i) + \sum_{\substack{i,j \\ \text{(within} \\ \text{impurity cell)}}} v(\boldsymbol{r}_i - \boldsymbol{r}_j). \tag{3.127}$$

Here $\mathcal{H}_0$ is the Hamiltonian for the pure host.

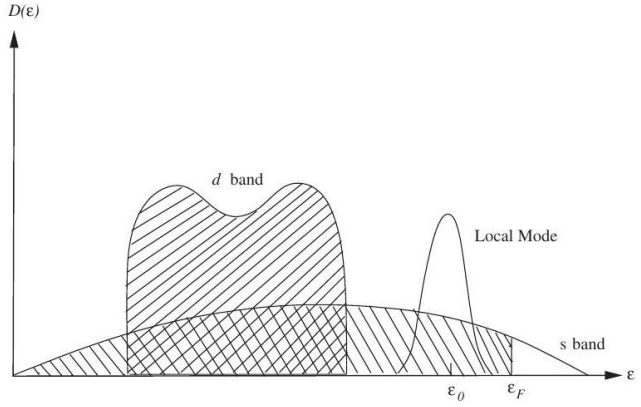


Fig. 3.19. Schematic density of states of copper containing an impurity mode

The next step is to write this Hamiltonian in its second-quantized form. The actual form will depend on the states we use to expand the field operator. For example, suppose we dissolve a transition metal such as manganese into copper. Since Mn^{2+} has four fewer electrons than Cu^{2+}, it constitutes a strong perturbation. Therefore we might expect the localized state to be pushed out of the d band of the copper host, as shown in Fig. 3.19. Since the d band is filled, it will not contribute to the properties of the dilute alloy. Therefore we expand our field operator in terms of the s-band states and a localized nondegenerate d-like state. The one-electron terms in (3.127) then becomes

$$\sum_{k,\sigma} \epsilon_k c^\dagger_{k\sigma} c_{k\sigma} + \sum_\sigma \epsilon_0 c^\dagger_{0\sigma} c_{0\sigma} + \sum_{k,\sigma} \left(V_{0k} c^\dagger_{k\sigma} c_{0\sigma} + V^*_{0k} c^\dagger_{0\sigma} c_{k\sigma} \right). \tag{3.128}$$

The first term is the s-band energy, the second term is the contribution from the impurity state, and the last term is the so-called sd mixing term. Notice that this mixing is a one-electron effect. It corresponds to the hopping of an electron from the localized d orbital into the conduction band, or vice versa. The interaction part of (3.127) becomes

$$U n_{0\uparrow} n_{0\downarrow}, \tag{3.129}$$

where U is the intraatomic Coulomb repulsion between opposite spins in the localized orbital. Equations (3.128), (3.129) constitute what is known as the Anderson Hamiltonian [65].

To study the implications of this model, Anderson employed the so-called Hartree-Fock approximation which is a mean field approximation. Namely, the interaction term $U n_{0\uparrow} n_{0\downarrow}$ is replaced by $U\left(\langle n_{0\uparrow}\rangle n_{0\downarrow} + \langle n_{0\downarrow}\rangle n_{0\uparrow} - \langle n_{0\uparrow}\rangle \langle n_{0\downarrow}\rangle\right)$. Thus, in the Hartree-Fock approximation the Anderson Hamiltonian for a spin σ becomes

$$\mathcal{H}^\sigma_{\mathrm{HF}} = \sum_{\boldsymbol{k}} \epsilon_{\boldsymbol{k}} c^\dagger_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma} + \sum_\sigma \epsilon_{0\sigma} c^\dagger_{0\sigma} c_{0\sigma} + \sum_{\boldsymbol{k}} V_{0\boldsymbol{k}} \left(c^\dagger_{\boldsymbol{k}\sigma} c_{0\sigma} + c^\dagger_{0\sigma} c_{k\sigma} \right), \tag{3.130}$$

where

$$\epsilon_{0\sigma} = \epsilon_0 + U \langle n_{0,-\sigma} \rangle \tag{3.131}$$

Solving the Schrodinger equation is equivalent to finding the so-called Green's function, defined by

$$G^{\sigma}(\mathcal{E}) = \lim_{s \to 0} \frac{1}{\mathcal{E} - \mathcal{H}^{\sigma} + is} \tag{3.132}$$

The usefulness of this function lies in its relation to the density of states. In particular, since the density of states may be represented by a sum of delta functions at energies given by the exact eigenvalues, i.e.,

$$\eta(E) = \sum_{n} \delta\left(E - E_n\right) \tag{3.133}$$

then

$$\eta(E) = -\frac{1}{\pi} \operatorname{Im}[\operatorname{Tr} G(E)] \tag{3.134}$$

Multiplying both sides (3.132) by $(\epsilon - \mathcal{H}^{\sigma} + \mathrm{i}s)$ and computing the matrix elements of the resulting equation between the states $|\boldsymbol{k}\rangle$ and $|0\rangle$ gives the following four relations:

$$\left(\epsilon - \epsilon_{0\sigma} + \mathrm{i}s\right) G^{\sigma}_{00}(\epsilon) - \sum_{\boldsymbol{k}} V_{0\boldsymbol{k}} G^{\sigma}_{\boldsymbol{k}0}(\epsilon) = 1 \tag{3.135}$$

$$\left(\epsilon - \epsilon_{\boldsymbol{k}} + \mathrm{i}s\right) G^{\sigma}_{\boldsymbol{k}\boldsymbol{k}'}(\epsilon) - V_{\boldsymbol{k}0} G^{\sigma}_{0\boldsymbol{k}'}(\epsilon) = \delta_{\boldsymbol{k}\boldsymbol{k}'} \tag{3.136}$$

$$\left(\epsilon - \epsilon_{\boldsymbol{k}} + \mathrm{i}s\right) G^{\sigma}_{\boldsymbol{k}0}(\epsilon) - V_{\boldsymbol{k}0} G^{\sigma}_{00}(\epsilon) = 0 \tag{3.137}$$

$$\left(\epsilon - \epsilon_{0\sigma} + \mathrm{i}s\right) G^{\sigma}_{0\boldsymbol{k}}(\epsilon) - \sum_{\boldsymbol{k}'} V_{0\boldsymbol{k}'} G^{\sigma}_{\boldsymbol{k}'\boldsymbol{k}}(\epsilon) = 0 \tag{3.138}$$

By solving these equations we obtain the various matrix elements of the Green's function. We then use (3.134) to compute the local density of states $\eta_{0\sigma}(\epsilon)$. This has the Lorentzian form

$$\eta_{0\sigma}(\epsilon) = \frac{1}{\pi} \frac{\Delta}{\left(\epsilon - \epsilon_0 - U\left\langle n_{0-\sigma}\right\rangle - \delta\epsilon\right)^2 + \Delta^2} \tag{3.139}$$

where

$$\Delta = \pi \sum_{\boldsymbol{k}} \left|V_{0\boldsymbol{k}}\right|^2 \delta\left(\epsilon - \epsilon_{\boldsymbol{k}}\right) \tag{3.140}$$

and

$$\delta\epsilon = \mathcal{P} \sum_{k} \frac{\left|V_{0k}\right|^2}{\epsilon - \epsilon_k} \tag{3.141}$$

From this density of states we can compute the occupation number of the localized state:

$$\left\langle n_{0\sigma}\right\rangle = \int_{-\infty}^{\epsilon_F} \eta_{0\sigma}(\epsilon) d\epsilon \tag{3.142}$$

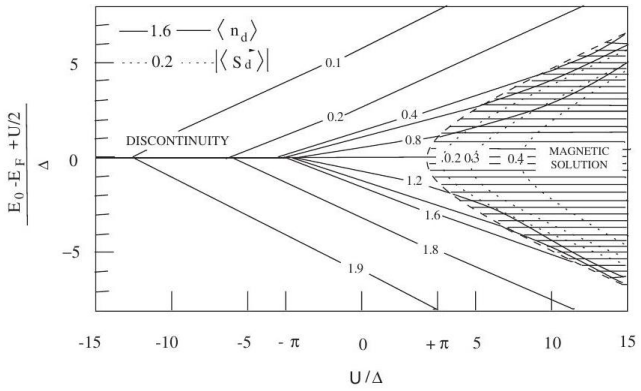


Fig. 3.20. "Phase diagram" corresponding to the Hartree-Fock solution of the Anderson Hamiltonian

Since $\eta_{0\sigma}$ is a function of $\langle n_{0-\sigma}\rangle$, we have a set of coupled equations. These may have the single nonmagnetic solution $\langle n_{0\sigma}\rangle = \langle n_{0-\sigma}\rangle$, or the two symmetrical magnetic solutions $\langle n_{0\sigma}\rangle \neq \langle n_{0-\sigma}\rangle$, depending upon the relative values of the parameters ϵ_0, U, and Δ. These regimes are indicated in Fig. 3.20. Notice that integrating (3.123) and comparing it with (3.142) gives $\delta_\sigma(\epsilon_F) = \pi\langle n_{0\sigma}\rangle$. If Z is the total number of localized electrons, then $\delta_\sigma + \delta_{-\sigma} = \pi Z$, which is a special case of the Friedel sum rule, (3.122).

To determine the boundary between the magnetic and nonmagnetic regions let us introduce the number of localized electrons n_0 and the localized moment m_0 by

$$n_0 = \frac{1}{2}\left(\langle n_{0\uparrow}\rangle + \langle n_{0\downarrow}\rangle\right) \tag{3.143}$$

$$m_0 = \frac{1}{2}\left(\langle n_{0\uparrow}\rangle - \langle n_{0\downarrow}\rangle\right) \tag{3.144}$$

Then

$$m_0 = \frac{1}{2}f\left(\epsilon_F; \epsilon_0 + Un_0 - Um_0\right) - \frac{1}{2}f\left(\epsilon_F; \epsilon_0 + Un_0 + Um_0\right), \tag{3.145}$$

where f is the $\cot^{-1}$ function obtained from integrating the Lorentzian density of states.

In the limit of large m_0 the right-hand side goes to 0 . Therefore, if there is a nonzero solution, the right-hand side must have a slope greater than 1 at $m_0 = 0$. Since

$$\left.\frac{\partial f\left(\epsilon_F; \epsilon_0 + Un_0 \mp Um_0\right)}{\partial m_0}\right|_{m_0=0} = \mp U\frac{\partial f\left(\epsilon_F; \epsilon_0 + Un_0\right)}{\partial\left(\epsilon_0 + Un_0\right)} \tag{3.146}$$

the condition for the appearance of the localized moment is

$$-U\frac{\partial f\left(\epsilon_F; \epsilon_0 + Un_0\right)}{\partial\left(\epsilon_0 + Un_0\right)} > 1. \tag{3.147}$$

If the localized state lies close to the Fermi surface, then Δ and $\delta\epsilon$ may be approximated by their values at the Fermi surface. Then

$$\frac{\partial f\left(\epsilon_F; \epsilon_0 + Un_0\right)}{\partial\left(\epsilon_0 + Un_0\right)} = \int_{-\infty}^{\epsilon_F} d\epsilon \frac{\partial \eta_{0\sigma}(\epsilon)}{\partial\left(\epsilon_0 + Un_0\right)} = -\eta_{0\sigma}\left(\epsilon_F\right). \tag{3.148}$$

The condition (3.147) then becomes

$$U\eta_{0\sigma}\left(\epsilon_F\right) > 1. \tag{3.149}$$

From (3.139) we see that a localized spin moment appears when U is sufficiently large in comparison with the width Δ of the resonant level. Since $\Delta = \pi V^2 \eta(\epsilon_F)$ a small density of states for the conduction band and a small covalent admixture favor the formation of a moment. Thus we expect to find moments when $3d$ atoms are dissolved in noble metals but not in aluminum, as is the case.

One of the consequences of Anderson's Hartree-Fock solution is that in the magnetic regime the density of states is characterized by two Lorentzians separated by an energy $U(\langle n_{0\uparrow}\rangle - \langle n_{0\downarrow}\rangle)$. If we again look to the residual resistivity for evidence of such behavior, we do indeed find two peaks for transition-metal impurities dissolved in Cu as shown in Fig. 3.21. More microscopic probes, however, indicate that these solutes behave more atomically than the itinerant Anderson model would suggest. An extensive study of the moments of Fe in various hosts have been carried out by Riegel et al. [66] using timedifferential perturbed angular γ-ray distribution. In this technique a scandium (^{45}Sc) target is bombarded by carbon ions which produce ^{54}Fe ions

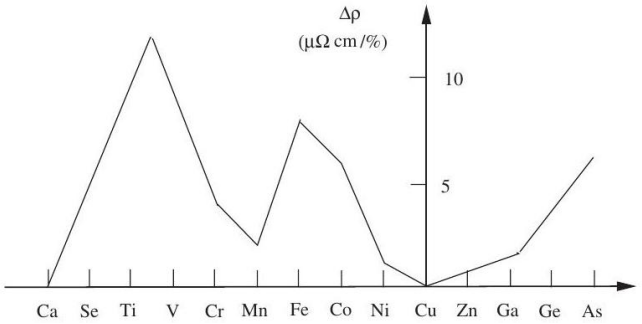


Fig. 3.21. Residual resistivities of transitional impurities in copper

(together with a proton and two neutrons) which recoil and imbed themselves in an adjacent film of the host to be studied. The ^{54}Fe then decays with the emission of two γ-rays. The second γ-ray emitted has an asymmetric distribution with respect to the first. If a field H_0 is applied perpendicular to the plane of the two γ-rays the asymmetric distribution rotates in space with a Larmour frequency $\omega_L(T) = \hbar^{-1}\mu_N g_N (H_0 + H_i)$ where H_i is the hyperfine field (2.47). According to the Wigner-Eckart theorem the electronic matrix element of (2.47) is proportional to $\langle J_z\rangle$. In the high-temperature limit $\langle J_z\rangle = g_J \mu_B J(J+1)H_0/3k_BT$. Therefore $H_0 + H_i$ may be written as $\beta(T)H_0$ where $\beta - 1$ may be thought of as a "local" susceptibility. These studies show that the behavior of Fe in the third-row transition metals has the same form as that shown in Fig. 3.18, i.e., Fe is nonmagnetic in Ta, Re, and Os, while it is magnetic in Mo, Ir, Pt, and Au with non-half-integral moments. However, in non-d-state hosts such as K, Rb, and Cs iron shows an ionic moment appropriate to a $3d^6$ configuration. This suggests that moment formation in the noble- and transition-metal hosts involves the d-states of the host. This is particularly true in Pd and Pt where iron polarizes its host to produce a "giant" moment. Wolff [67] has developed a theory of local moments similar to Anderson's but obtains the local moment as a virtual bound state thereby explicitly involving the host.

3.4.3 The Kondo Effect

Much more dramatic is the behavior associated with the breakdown of the Hartree-Fock approximation itself, as manifest in the Kondo effect. Figure 3.22 shows the general behavior observed in "local-moment" systems such as those illustrated in Fig. 3.21 with decreasing temperature. The resistivity increases logarithmically and eventually saturates. When this increasing impurity contribution is combined with the decreasing phonon contribution, one obtains a minimum

in the resistivity, a phenomenon that was known but not understood for many years. At the same time, the magnetic moment, as measured by the susceptibility, decreases, and the specific heat shows a peak corresponding to an entropy change of approximately $k_B \ln 2$. Although this transition is gradual, it is possible to assign a temperature, called the Kondo temperature, below which these anomalous properties appear. Figure 3.23 shows how this Kondo temperature varies for $3d$ transition metals dissolved in copper.

It is now known that these properties, which are collectively referred to as the "Kondo effect", are associated with the behavior of a single localized electron spin interacting with a degenerate electron system. This seemingly simple problem has required an enormous effort to bring us to the point where we think that we understand what is going on. We shall not cover all this development (see for example [68a] or the reviews [68b]), but merely indicate a few aspects that will give the reader a "feel" for the Kondo effect.

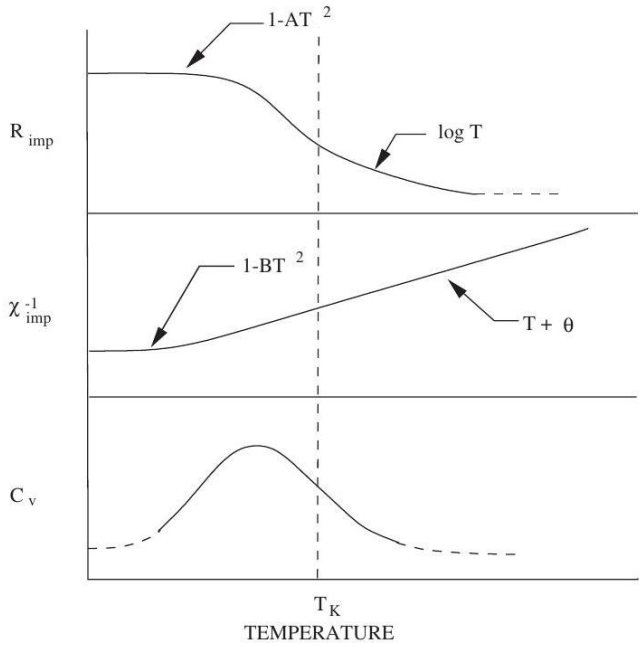


Fig. 3.22. Schematic behavior characteristic of a typical Kondo alloy

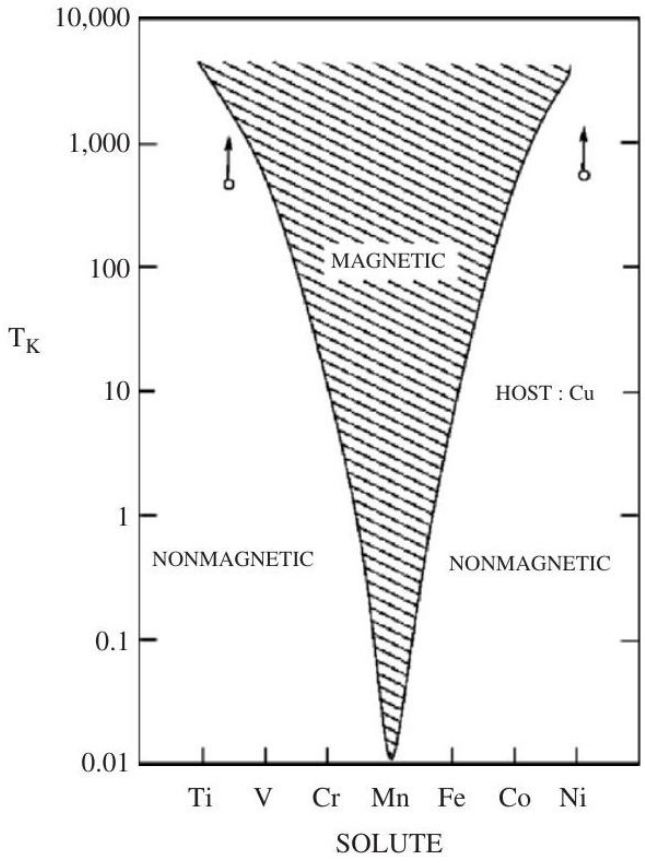


Fig. 3.23. Variation of the Kondo temperature for the $3d$ transition metals in copper

Let us begin by assuming we are in the magnetic regime of the Anderson model, that is $U \gg \Delta$. In this case Schrieffer and Wolff showed that one could perform a canonical transformation which eliminates V_{0k} to first order [69].

They found that the resulting Hamiltonian contains a contribution of the form

$$\mathcal{H}_{sd} = \sum_{\boldsymbol{k},\boldsymbol{k}'} J_{\boldsymbol{k}'\boldsymbol{k}} \psi^{\dagger}_{\boldsymbol{k}'} s \psi_{\boldsymbol{k}} \cdot \psi^{\dagger}_{0} \boldsymbol{S} \psi_{0} \tag{3.150}$$

where

$$\psi_{\boldsymbol{k}} = \begin{bmatrix} c_{\boldsymbol{k}\uparrow} \\ c_{\boldsymbol{k}\downarrow} \end{bmatrix} \quad \text{and} \quad \psi_{0} = \begin{bmatrix} c_{0\uparrow} \\ c_{0\downarrow} \end{bmatrix}. \tag{3.151}$$

This has the form of an sd exchange interaction. For wave vectors near the Fermi surface the exchange parameter becomes

$$J_{\boldsymbol{k}_F\boldsymbol{k}_F} = 2\,|V_{0\boldsymbol{k}_F}|^2 \,\frac{U}{\epsilon_0\,(\epsilon_0 + U)}. \tag{3.152}$$

In this expression ϵ_0 is measured in relation to the Fermi surface, which makes $J_{\boldsymbol{k}_F\boldsymbol{k}_F}$ negative, corresponding to antiferromagnetic coupling.

As we mentioned above, those alloys possessing local moments are also the ones which show a minimum in their resistivity as a function of temperature. This motivated Kondo [70] to calculate the spin-flip scattering amplitude of a conduction electron using the $s-d$ interaction (3.150). Let us outline this calculation. We shall characterize the impurity spin by its spin-projection quantum numbers M_S and add an extra electron in the state ($\boldsymbol{k}, \sigma$) to the Fermi

sea. The state of the system at time $t = -\infty$ may then be written $c^\dagger_{k\sigma}\,|FS; M'_S\rangle$. In the absence of an external field the unperturbed Hamiltonian is just

$$\mathcal{H}_0 = \sum \epsilon_{\boldsymbol{k}} c^\dagger_{\boldsymbol{k}\sigma} c_{\boldsymbol{k}\sigma}.$$

We now adiabetically turn on the interaction $\mathcal{H}_1$ and examine the amplitude for scattering into the state $c^\dagger_{\boldsymbol{k}'\sigma'}\,|FS; \boldsymbol{M}_S\rangle$. This amplitude is

$$\begin{aligned}\langle \boldsymbol{k}'\sigma'; \boldsymbol{M}'_S \mid \phi(\infty)\rangle =& \delta\left(\boldsymbol{k}-\boldsymbol{k}'\right)\delta_{\sigma,\sigma'}\delta_{M_S\cdot M'_S} \\ & - 2\pi i \left\langle \boldsymbol{k}'\sigma'; \boldsymbol{M}'_S\right| \mathcal{H}_1 + \mathcal{H}_1 \frac{1}{\epsilon_k - \mathcal{H}_0}\mathcal{H}_1 + \ldots \left|\boldsymbol{k}, \sigma; \boldsymbol{M}_S\right\rangle \\ & \times \delta\left(\epsilon_{\boldsymbol{k}} - \epsilon_{\boldsymbol{k}'}\right) \end{aligned} \tag{3.153}$$

Suppose $\mathcal{H}_1$ corresponded to ordinary potential scattering; then it would have the form $\sum c^\dagger_{l'\sigma} V_{l'l} c_{l\sigma}$. The interesting term in (3.153) is the one involving $\mathcal{H}_1$ twice. There are two distinct ways in which the scattering $\boldsymbol{k} - \boldsymbol{k}'$ can occur through this term. These are illustrated in Fig. 3.24. The matrix element for process (a) is

$$\langle FS| c_{\boldsymbol{k}'} c^\dagger_{\boldsymbol{k}'} V_{\boldsymbol{k}'l} c_l \frac{1}{\epsilon_{\boldsymbol{k}} - \mathcal{H}_0} c^\dagger_l V_{l\boldsymbol{k}} c_{\boldsymbol{k}} c^\dagger_{\boldsymbol{k}} |FS\rangle, \tag{3.154}$$

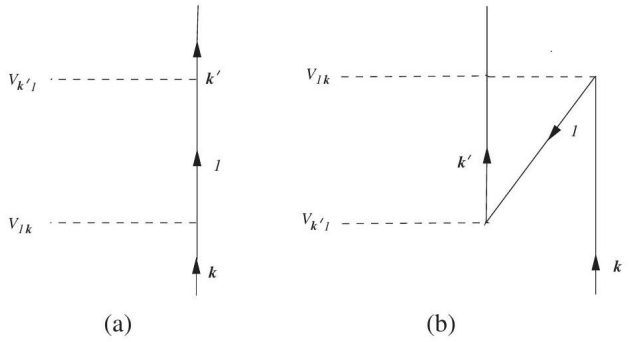


Fig. 3.24. Contributions to the second-order scattering amplitude associated with a spin-independent potential

which reduces to

$$V_{\boldsymbol{k}'l} \frac{1 - f\left(\epsilon_l\right)}{\epsilon_{\boldsymbol{k}} - \epsilon_l} V_{l\boldsymbol{k}} \tag{3.155}$$

Process (b) becomes

$$\langle FS| c_{\boldsymbol{k}'} V_{l\boldsymbol{k}} c_{\boldsymbol{k}} \frac{1}{\epsilon_{\boldsymbol{k}} - \mathcal{H}_0} c^\dagger_{\boldsymbol{k}'} V_{\boldsymbol{k}'l} c_l c^\dagger_{\boldsymbol{k}} |FS\rangle, \tag{3.156}$$

which contracts to

$$-V_{l\boldsymbol{k}} \frac{f\left(\epsilon_l\right)}{\epsilon_{\boldsymbol{k}} - \left(\epsilon_{\boldsymbol{k}} - \epsilon_l + \epsilon_{\boldsymbol{k}'}\right)} V_{\boldsymbol{k}'l} = V_{l\boldsymbol{k}} \frac{f\left(\epsilon_l\right)}{\epsilon_{\boldsymbol{k}'} - \epsilon_l} V_{\boldsymbol{k}'l} \tag{3.157}$$

When these two processes are added together, the Fermi function cancels out. Thus, the exclusion principle does not enter into ordinary potential scattering. This is not the case, however, for a spin-dependent potential. The Hamiltonian (3.150) becomes

$$\mathcal{H}_1 = -J\boldsymbol{S}\cdot\sum_{l,l'}\sum_{\sigma,\sigma'} c^{\dagger}_{l\sigma}s_{\sigma\sigma'}c_{l'\sigma'} \tag{3.158}$$

The two second-order processes arising from this interaction are illustrated in Fig. 3.25. Process (a) gives

$$J^2\left\langle FS;\boldsymbol{M}'_S\right|S_\nu S_\mu c_{\boldsymbol{k}'\sigma'}c^{\dagger}_{\boldsymbol{k}'\sigma'}s^{\nu}_{\sigma'\sigma''}c_{l\sigma''}\frac{1}{\epsilon_{\boldsymbol{k}}-\mathcal{H}_0}c^{\dagger}_{l\sigma''}s^{\mu}_{\sigma''\sigma}c_{\boldsymbol{k}\sigma}c^{\dagger}_{\boldsymbol{k}\sigma}\left|FS;\boldsymbol{M}_S\right\rangle, \tag{3.159}$$

which reduces to

$$J^2\frac{1-f\left(\epsilon_l\right)}{\epsilon_{\boldsymbol{k}}-\epsilon_l}\left\langle\boldsymbol{M}'_S\right|S_\nu S_\mu s^{\nu}_{\sigma'\sigma''}s^{\mu}_{\sigma''\sigma}c_{\boldsymbol{k}\sigma}c^{\dagger}_{\boldsymbol{k}\sigma}\left|\boldsymbol{M}_S\right\rangle. \tag{3.160}$$

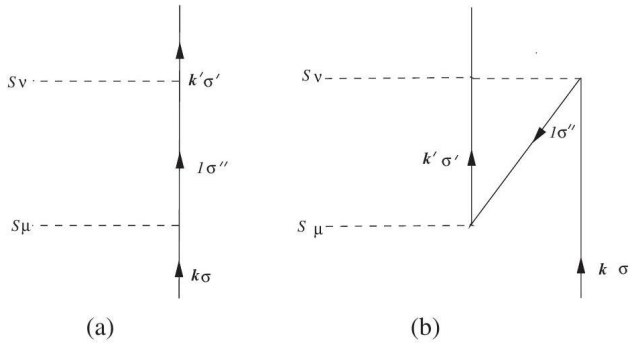


Fig. 3.25. Contributions to the second-order scattering amplitude associated with the $s-d$ interaction

Process (b) gives

$$J^2\frac{f\left(\epsilon_l\right)}{\epsilon_{\boldsymbol{k}}-\epsilon_l}\left\langle\boldsymbol{M}'_S\right|S_\nu S_\mu s^{\nu}_{\sigma''\sigma}s^{\mu}_{\sigma'\sigma''}\left|M_S\right\rangle. \tag{3.161}$$

When these are added together, we obtain

$$\frac{J^2}{\epsilon_{\boldsymbol{k}}-\epsilon_l}\left[\frac{1}{4}S(S+1)-\left\langle M'_S,\sigma'\right|\boldsymbol{S}\cdot\boldsymbol{s}\left|M_S,\sigma\right\rangle\right]+\frac{J^2f\left(\epsilon_l\right)}{\epsilon_{\boldsymbol{k}}-\epsilon_l}\left\langle M'_S,\sigma'\right|\boldsymbol{S}\cdot\boldsymbol{s}\left|M_S,\sigma\right\rangle. \tag{3.162}$$

The second term in (3.162) arises from the fact that the spin operators may not be interchanged. Therefore we obtain a contribution to the scattering which does depend on the exclusion principle. Summing over the intermediate states introduces the factor

$$\sum_l\frac{f\left(\epsilon_l\right)}{\epsilon_k-\epsilon_l}.$$

Introducing the conduction-electron density states of $N(\epsilon)$, this sum may be written

$$\int N\left(\epsilon'\right)\frac{f\left(\epsilon'\right)}{\epsilon-\epsilon'}d\epsilon' \tag{3.163}$$

Assuming a simple rectangular band,

$$N(\epsilon)=\begin{cases}N(0) & -D<\epsilon<D\\ 0 & \text{otherwise}\end{cases} \tag{3.164}$$

the integral for $\epsilon = 0$ gives

$$-N(0) \ln \frac{k_B T}{D} + \text{ const.} \tag{3.165}$$

Since the resistivity involves the square of the scattering amplitude, the cross term between the first-order term and this second-order term gives a contribution proportional to $J^3 \ln T$.

For antiferromagnetic exchange ($J < 0$) the resistivity increases with decreasing temperature as shown in Fig. 3.22. The divergence as $T \to 0$, however, indicates that the perturbation theory is not valid in this regime. This remarkable result initiated intense activity. Experimental studies established the systematics of this phenomenon while theoretical work explored the properties of the $s - d$ Hamiltonian as well as the Anderson Hamiltonian.

The first attempts to improve on Kondo's calculation led to a scattering amplitude that diverged at a finite temperature

$$T_K = D \exp \left[1/2JN\left(E_F\right)\right] \tag{3.166}$$

which provides a quantitative expression for the Kondo temperature. Although the divergence indicated that the problem was still not being treated properly, it did indicate the formation of a resonance in the scattering amplitude at the Fermi level. That is, the Kondo effect is basically an impurity-induced Fermi surface instability. A great deal of effort has gone into identifying the nature of the ground state. From the very beginning it was suspected that it must be a singlet. However, it was only the exact diagonalization of the $s-d$ Hamiltonian by Andrei [71] and Wiegmann [72] that this singlet nature was confirmed. Yosida and collaborators [73] have shown that if one assumes a singlet wave function of the form

$$\psi_{\text{Kondo}} = \frac{1}{\sqrt{2}} \left(\chi_\uparrow \psi_\uparrow - \chi_\downarrow \psi_\downarrow\right) \tag{3.167}$$

where χ_σ is the wave function of the impurity spin, then the conductionelectron component $\psi_\uparrow$ consists of half an electron with down spin and half a hole with up spin bound, to the impurity. This is illustrated schematically in Fig. 3.26.

Anderson [74] also considered the implications of a singlet ground state on the conduction-electron phase shifts. In the Hartree-Fock approximation discussed above, in the case of a well-localized moment $\delta_\uparrow = \pi$ and $\delta_\downarrow = 0$. However, if we do have a singlet ground state there is rotational invariance and the phase shift must be independent of σ. Anderson postulated that in the vicinity of the Fermi level the phase shifts readjust as indicated in Fig. 3.26. The appearance of $\pi/2$ phase shifts at the Fermi level implies a resonance at the Fermi level. This "Abrikosov-Suhl" resonance, as it is sometimes called, is also shown in Fig. 3.26. Such a peak in the density of states has also been derived [75] directly from the Anderson Hamiltonian. This result shows that the peak has a width $k_B T_K$ and disappears at the Kondo temperature. If only electrons within a characteristic energy $k_B T_K$ of the Fermi level are involved, this implies a spatial extent of

$$\xi_K = \frac{v_F}{k_B T_K} \tag{3.168}$$

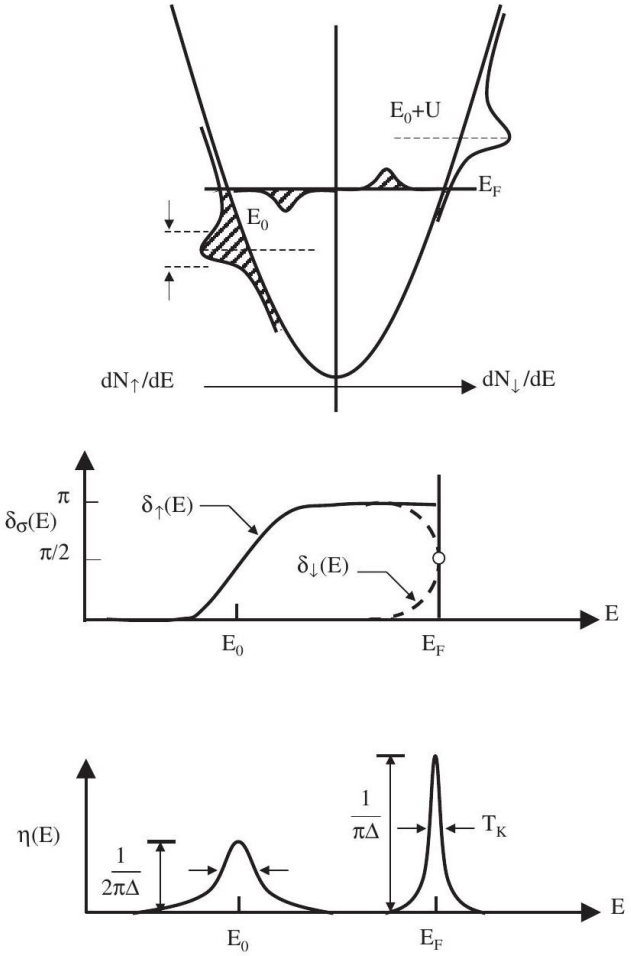


Fig. 3.26. (a) Schematic illustration of how the Hartree-Fock solution is augmented by the appearance of a resonance at the Fermi level consisting of half a spin-down electron and half a spin-up hole. (b) Variation of the phase shift with energy. (c) Variation in the local density of states

Problems

3.1. Show that the sum of χ_{diamag} defined by (3.23) and χ_{VV} defined by (3.28) is invariant under a change in origin.
Hint: Use the relation

$$\langle\alpha|p_\mu|\beta\rangle = \frac{iE_{\alpha\beta}}{\hbar}\langle\alpha|\mu|\beta\rangle,$$

where $\mu = x, y$, or z.

3.2. (a) Compute the partition function, Z, for a collection of N spin- J particles in a field H.
(b) Using the thermodynamic relation for the entropy,

$$S = -\frac{\partial\,(k_B T \ln Z)}{\partial T}$$

compute the entropy for the collection of spin- J particles.
(c) Let us assume that the sample can be thermally insulated so that the entropy remains constant. Then if we reduce the field from some initial value, H_i, to some final value H_f, the temperature will change from its initial value, T_i, to a new final value, T_f. Calculate T_f. This

is the principle of magnetic cooling.
3.3. Assume that the saturation magnetization (magnetic moment per unit volume) for iron arises only from the ionic moments associated with the electronic configuration Fe^{2+} $\left(3\ d^6\right)$. Calculate the saturation magnetization for a body-centered cubic structure with a lattice parameter $a = 2.87\text{Å}$. How does this compare with the experimental value for iron?
3.4. Calculate the de Haas-van Alphen period $\Delta(1/H)$ expected for copper assuming a free electron model with a Fermi wave vector of $k_F = 1.36 \times 10^8\ \text{cm}^{-1}$.
3.5. Consider a two-dimensional electron gas (2 DEG) of density n (electrons/ cm^2) in which the electrons have an effective g-factor g^* and an effective mass m^*.
(a) Show that the spin susceptibility is proportional to g^*m^*.
(b) Calculate the field, B_p, at which the system would exhibit total spin polarization.
3.6. Consider a dimer consisting of two Cu^{2+} $(3d^9)$ ions. Assume that the 3d holes interact through an antiferromagnetic exchange, $\mathcal{H} = Js_1.s_2$. Using (3.19) derive the susceptibility of this isolated dimer.

10.5 The Static Susceptibility of Interacting Systems: Local Moments

In this chapter we shall turn our attention to the static response of systems in which interactions among localized moments may lead to long-range order. A thorough treatment of such interactions and the resulting phase transition is an extremely complicated subject. Although very sophisticated techniques have been developed to treat these problems, there are still gaps in our understanding. However, except for the region very close to the critical point, an effective field theory works quite well. Therefore we shall employ this approximation throughout most of this chapter.

In Sect. 2.2 we found that the Coulomb interaction between the valence electrons on different ions could be expressed as an effective interaction between the individual electron spins. We also found that under certain conditions this interaction could be expressed in terms of the total ionic spin or, in some cases, the effective spin of the ground-state multiplet. Let us consider a system whose interactions may, in fact, be described by the Heisenberg exchange Hamiltonian

$$\mathcal{H}_{\text{ex}} = -\sum_i \sum_{j \neq i} J_{ij} S_i \cdot S_j, \tag{4.1}$$

where J_{ij} is a function of the relative separation $\boldsymbol{R}_i - \boldsymbol{R}_j$.
Although many systems do not satisfy our conditions for writing the interaction in this form, it is often found that such systems may, nevertheless, be described surprisingly well by it. However, as we shall see, the treatment of this interaction entails certain approximations. Hence our success may result more from our methods than from the interaction with which we start. Since the exchange interaction is often large in comparison with other interactions, such as the magnetic-dipole-dipole interaction, the gross features of such a system may be described by (4.1) alone.

Let us now see how such a system responds to a static applied field. Since (4.1) is rotationally invariant, the direction of this applied field may be taken to define the z-axis. Just as in the preceding chapter, we apply a field of the form

$$\boldsymbol{H}(\boldsymbol{r}) = \boldsymbol{H}\hat{\boldsymbol{z}} \cos(\boldsymbol{q} \cdot \boldsymbol{r}), \tag{4.2}$$

where $\boldsymbol{q}$ corresponds to some point within the first Brillouin zone of the lattice. If the g value associated with the effective spin is g, then the total Hamiltonian becomes

$$\mathcal{H} = -\sum_i \sum_{j \neq i} J_{ij} S_i \cdot S_j + g\mu_B H \sum_i S_i^z \cos\left(\boldsymbol{q} \cdot \boldsymbol{R}_i\right). \tag{4.3}$$

As in most cases involving coupling between lattice sites, it is convenient to express the Hamiltonian in Fourier components. Since

$$S\left(\boldsymbol{q}'\right) = \int d\boldsymbol{r} e^{i\boldsymbol{q}'\cdot\boldsymbol{r}} S(\boldsymbol{r}) = \sum_i \int d\boldsymbol{r} e^{i\boldsymbol{q}'\cdot\boldsymbol{r}} S_i \delta\left(\boldsymbol{r} - \boldsymbol{R}_i\right) = \sum_i e^{i\boldsymbol{q}'\cdot\boldsymbol{r}_i} S_i \tag{4.4}$$

we may write (4.3) as

$$\mathcal{H} = -\sum_{\boldsymbol{q}'} J\left(-\boldsymbol{q}'\right) S\left(\boldsymbol{q}'\right) \cdot S\left(-\boldsymbol{q}'\right) + \frac{1}{2} g\mu_B H \left[S_z(\boldsymbol{q}) + S_z(-\boldsymbol{q})\right] \tag{4.5}$$

where

$$J\left(-\boldsymbol{q}'\right) \equiv \frac{1}{N} \sum_{i \neq j} J\left(\boldsymbol{R}_i - \boldsymbol{R}_j\right) e^{i\boldsymbol{q}'\cdot\left(\boldsymbol{r}_i - \boldsymbol{r}_j\right)} \tag{4.6}$$

Notice that if the crystal has inversion symmetry, $J(-\boldsymbol{q}) = J(\boldsymbol{q})$. The analysis of the Hamiltonian (4.5) is very difficult and has been the subject of much work. The simplest approach is to consider one of the spins in (4.1) and replace its interaction with the other spins by an effective field [76]. This is referred to as the mean-field approximation. The concept appears to have first been applied by van der Waals in 1873 to develop the liquid-gas equation of state. Its application to ferromagnetism was made by P. Weiss in 1907. The same approximation was employed by Bragg and Williams in 1934 to describe orderdisorder transitions such as occur in β-brass when the copper and zinc atoms, through diffusion, order themselves with the copper at the center and the zinc at the corners of the body-centered-cubic lattice. The physics of such orderdisorder transitions is identical to that of an Ising magnet. In 1937 Landau developed a generalized mean-field theory applicable to all second-order phase transitions. We shall return to Landau's theory later in this chapter.

Since we have been emphasizing the usefullness of working with Fourier components it is appropriate to apply a variation of the mean-field approximation in which each of the Fourier components in (4.5) is assumed to be independent. This if referred to as the random-phase approximation. In this approximation the exchange interaction becomes

$$\mathcal{H}_{\text{ex}} = -\sum_{\boldsymbol{q}'} J\left(-\boldsymbol{q}'\right) S\left(\boldsymbol{q}'\right) \cdot \left\langle S\left(-\boldsymbol{q}'\right)\right\rangle. \tag{4.7}$$

10.5.1 4.1 High Temperatures

If the temperature is higher than that at which long-range order sets in, we may assume that the only components with nonzero average values are those driven by the applied field. Thus $\left\langle S\left(-\boldsymbol{q}'\right)\right\rangle$ is replaced by $\left\langle S_z\left(-\boldsymbol{q}'\right)\right\rangle \hat{\boldsymbol{z}}\Delta\left(\boldsymbol{q} - \boldsymbol{q}'\right)$. The total Hamiltonian in the random-phase approximations becomes

$$\begin{aligned}\mathcal{H}_{\text{RPA}} = &-\left[J(\boldsymbol{q})\left\langle S_z(-\boldsymbol{q})\right\rangle S_z(\boldsymbol{q}) + J(\boldsymbol{q})\left\langle S_z(\boldsymbol{q})\right\rangle S_z(-\boldsymbol{q})\right] \\ &+ \frac{1}{2} g\mu_B H \left[S_z(\boldsymbol{q}) + S_z(-\boldsymbol{q})\right]\end{aligned} \tag{4.8}$$

which defines the total effective field seen by the $\boldsymbol{q}$ th component as

$$\boldsymbol{H}(-\boldsymbol{q})_{\text{eff}} = \frac{J(-\boldsymbol{q})V}{g\mu_B} \langle S_z(-\boldsymbol{q})\rangle \, \hat{\boldsymbol{z}} + \frac{1}{2} HV\hat{\boldsymbol{z}} \tag{4.9}$$

In discussing noninteracting systems we found that the response of the $\boldsymbol{q}$ th component of the spin at high temperatures was determined by

$$\chi_0(\boldsymbol{q}) = \frac{C}{T} \equiv \chi_0 \tag{4.10}$$

where

$$C = \frac{N}{V} g^2 \mu_B^2 \frac{S(S+1)}{3k_B}$$

Therefore, employing our definition of the susceptibility, we have

$$\begin{aligned} M_z(\boldsymbol{q}) &= -g\mu_B \langle S_z(\boldsymbol{q})\rangle = \chi_0(\boldsymbol{q}) H_z(\boldsymbol{q})_{\text{eff}} \\ &= -\frac{J(\boldsymbol{q})\chi_0(\boldsymbol{q})V}{g\mu_B} \langle S_z(\boldsymbol{q})\rangle + \frac{1}{2} HV\chi_0(\boldsymbol{q}) \end{aligned} \tag{4.11}$$

Solving for $\langle S_z(\boldsymbol{q})\rangle$ and using (4.10) gives us the susceptibility of the interacting system at high temperatures,

$$\chi(\boldsymbol{q}) = \frac{\chi_0}{1 - \chi_0 \frac{J(\boldsymbol{q})V}{g^2\mu_B^2}} = \frac{C/T_c}{(T - T_c)/T_c + [1 - J(\boldsymbol{q})/J(\boldsymbol{q})]} \tag{4.12}$$

We notice immediately that as the temperature decreases a divergence appears at a critical temperature

$$T_c = \frac{J(\boldsymbol{q})CV}{g^2\mu_B^2} \tag{4.13}$$

where $\boldsymbol{q}$ is the wave vector for which $J(\boldsymbol{q})$ is a maximum. The fact that the susceptibility diverges for this particular wave vector means that the component of the magnetization with that wave vector would remain nonzero even if the probing field were to go to 0 . Thus we have a spontaneous magnetization. The wave vector $\boldsymbol{q}$ for which this transition occurs depends on the nature of $J(\boldsymbol{q})$. Let us consider four possible cases.

Ferromagnetism. Suppose that the exchange J_{ij} is positive and extends only to nearest neighbors. Then

$$J(\boldsymbol{q}) = \frac{J}{N} \sum_{\delta} \exp(-i\boldsymbol{q} \cdot \boldsymbol{\delta}), \tag{4.14}$$

where $\boldsymbol{\delta}$ is a vector to the nearest neighbor. In particular, if we have a simple cubic lattice with a lattice parameter a, then

$$J(\boldsymbol{q}) = \frac{2J}{N} (\cos q_x a + \cos q_y a + \cos q_z a) \tag{4.15}$$

Figure 4.1 shows $1/\chi(\boldsymbol{q})$ plotted as a function of temperature for various values of $\boldsymbol{q}$. We see immediately that as the temperature decreases the first divergence appears at $\boldsymbol{q} = 0$.

This instability results in a uniformly magnetized system, a ferromagnet. The temperature associated with this transition is the Curie temperature, which, for our simple cubic example, is

$$T_c = \frac{J(0)CV}{g^2\mu_B^2} = \frac{6JS(S+1)}{3k_B} \tag{4.16}$$

The uniform susceptibility is

$$\chi(0)_{\text{ferro}} = \frac{C}{T - T_c} \tag{4.17}$$

In fact, ferromagnetism is not compatible with cubic symmetry. This has to do with time reversal. In the ferromagnetically ordered state time reversal by itself is not a symmetry operation, but only occurs in conjunction with some other symmetry operator. Consider, for example, a body-centered-cubic (bcc) system. In the paramagnetic state the symmetry operations which transform this structure into itself constitute the octahedral group denoted O_h or m 3 m . If, in the ordered state, the spins at each site all point in the z-direction then those operations of O_h which involve rotations about the x - or y-axes must be combined with time reversal to preserve the direction of the spins. The resulting allowed operations constitute the so-called magnetic point group, which, in this case, is denoted 4/mmm. The important feature of this group is that it is tetragonal. When a material develops a spontaneous magnetization, it also generally develops a strain proportional to the magnetization. This magnetostrictive effect, as it is known, reflects the symmetry of the system. Thus, iron, which is bcc above its Curie temperature, develops a tetragonal distortion in the ferromagnetic state. However, since this distortion is very small one often finds ferromagnetic iron characterized as bcc.

Another general result which is obscured by the above mean field approximation has to do with the dimensionality of the system. Landau and

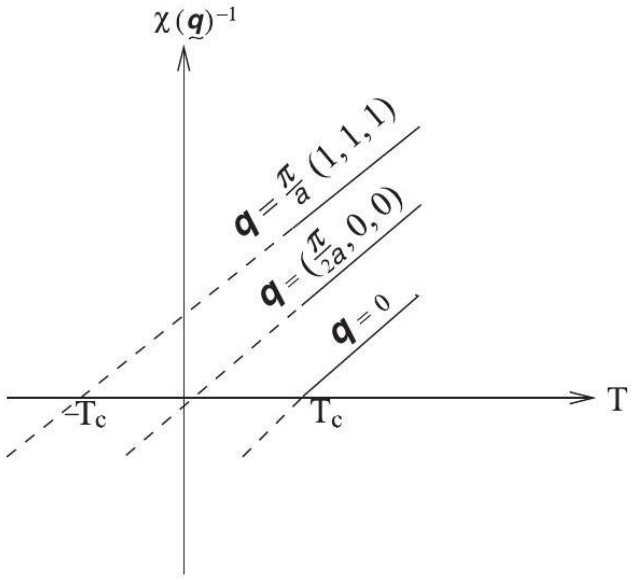


Fig. 4.1. Temperature dependence of the inverse susceptibility \$1 / \chi(\boldsymbol{q)\$ of a ferromagnet for various values of $\boldsymbol{q}$}

Lifshitz [77a] pointed out that one cannot have long-range order in one dimension at a finite temperature. Consider, for example, a chain of spins. The long-range order is destroyed by simply "flipping" one spin. Although this costs exchange energy, the entropy gained is proportional to the logarithm of the number of ways this can be done. For an infinitely long chain this entropy term will always dominate the free energy, resulting in disorder. Lieb and Mattis have shown [77b] that at $T = 0$ the ground state of a system of N electrons in 1-dimension subject to an arbitrary symmetric potential is unmagnetized. While it is not possible

to realize a truely 1-dimensional system, it is possible to fabricate GaAs-AlGaAs heterojunctions that behave as quantum wires. The quantization of conductance in integer multiples of $2e^2/\hbar$ is the signature of 1-D transport. Recent results suggest that under such conditions the transport electrons are spin polarized [77c]. This is a subject of current interest.

In two dimensions the situation depends upon the details of the interaction - a Heisenberg system, for example, does not exhibit ferromagnetism, while an Ising system does.
Antiferromagnetism. From Fig. 4.1 it is obvious that if the exchange constant between nearest neighbors is negative, then the first divergence appears for $\boldsymbol{q} = (\pi/a)(\hat{\boldsymbol{x}} + \hat{\boldsymbol{y}} + \hat{\boldsymbol{z}})$ or one of the other similar wave vectors in a bodydiagonal direction. This means that the phase of the magnetization changes by π from one plane of atoms to the next as we proceed along the direction of $\boldsymbol{q}$. Thus we have an antiferromagnetic system. The temperature at which $\chi(\boldsymbol{q})$ diverges is the Néel temperature,

$$T_N = \frac{6|J|S(S+1)}{3k_B}. \tag{4.18}$$

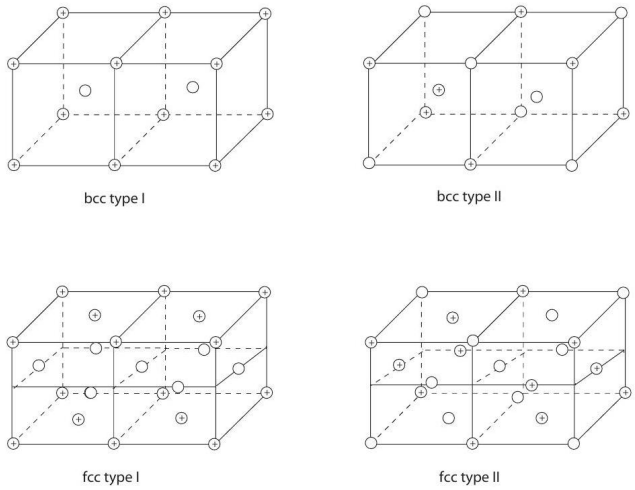


Fig. 4.2. Possible types of antiferromagnetic order in body-centered and facecentered cubic lattices. The symbols $\oplus$ and $\ominus$ refer to "up" and "down" magnetic moments, respectively

The uniform susceptibility is given by

$$\chi(0)_{\text{antiferro}} = \frac{C}{T + T_N}, \tag{4.19}$$

which is well behaved at the Néel temperature.
In general the exchange interaction will extend beyond just the nearest neighbors. This can lead to a variety of antiferromagnetic configurations. Antiferromagnetism is compatible with cubic symmetry, and several "types" of antiferromagnetic orderings in cubic lattices are illustrated in Fig. 4.2. The oxides MnO , FeO , and NiO are examples of the fcc type II order.
Helimagnetism. In Chap. 9 we shall discuss the RKKY exchange interaction, an indirect exchange between localized moments which arises through the presence of conduction electrons. For our present purposes it is sufficient to note that this interaction is long range and oscillatory. To see what effect such an exchange may have on our spin configuration, let us consider our simple cubic lattice and take

$$J_{ij} = \begin{cases} J_1 & \text{for the six nearest neighbors.} \\ -J_2 & \text{for the twelve next-nearest neighbors.} \end{cases}$$

			$RE^{3+}\left(RE^{2+}\right)$								
Ce	$4f^2$	$^2F_{5/2}$	AFM	PARA							
Pr	$4f^3$	$^3H_4$	PARA (SINGLET GROUND STATE)								
Nd	4f 4	$^4I_{9/12}$	AFM	PARA							
Pm	4f ^{5}	$^5I_4$	? INTENSELY RADIOACTIVE ?								
Sm	$4f^6$	$^6H_{5/2}$	AFM	AFM (hexagonal sites)					PARA		
Eu	4f ^{7}	$\left(^8S_{7/2}\right)$	AFM				1st		PARA		
Gd	$4f^8$	$^8S_{7/2}$	PLANAR FM PARA								
Tb	4f 9	$^7F_6$	PLANAR FM $H_{E_{L_{I_X}}}$ PARA								
Dy	4f ^{10}	$^6H_{15/2}$	PLANAR FM				HELIX		PARA		
Ho	4f ^{11}	$^5I_8$	CONE	HELIX							
Er	4f ^{12}	$^4I_{15/2}$	CONE	LW+ HELIX LW			PARA		⊗	Φ	
Tm	$4f^{13}$	$^3H_6$	ANTIPHASE	LW			PARA		8	(1) 0 (1) 0	
Yb	4f 14	$^2F_{7/2}\,(^1S_0)$	PARA (PAULI)						∇ Q CONE		
Lu	$4f^{14}5$ d	$^1S_0$	SUPERCONDUCTING $T_c = 0.1$ K							LW	
			0								
				100							
				TEMPERATURE (K)							

Fig. 4.3. Magnetic properties of the rare-earth metals. Data assembled from [78]

Then

$$J(\boldsymbol{q}) = \frac{2J_1}{N}\left(\cos q_x a + \cos q_y a + \cos q_z a\right) - \frac{4J_2}{N}\left[\left(\cos q_x a\right)\left(\cos q_y a\right) + \left(\cos q_y a\right)\left(\cos q_z a\right) + \left(\cos q_z a\right)\left(\cos q_x a\right)\right] \tag{4.20}$$

This exchange has an extremum when $a\boldsymbol{q} = \cos^{-1}\left(J_1/4J_2\right)\left(\hat{\boldsymbol{x}} + \hat{\boldsymbol{y}} + \hat{\boldsymbol{z}}\right)$. If $J_1 < 4J_2$, it will lead to a spin configuration whose wavelength is incommensurate with the lattice spacing. An example of such a system is the rareearth metal thulium (Tm) at temperatures between 40 and 56 K . The spin configuration of Tm in this region is illustrated in Fig. 4.3. This particular configuration is sometimes referred to as a longitudinal wave.

The general situation is actually more complicated. So far our probing field has had a sinusoidally varying amplitude but has been fixed in direction. In principle, we could also imagine a field whose direction also varies in space. Should the response to such a peculiar field diverge, a correspondingly peculiar spin configuration would occur. Such spin configurations are, in fact, found in the heavy rare-earth metals as indicated in Fig. 4.3. The insert illustrates various spin configurations. Such systems are referred to as helimagnets. Dealing with the

generalized susceptibility is not a very practical way of determining the existence of such spin configurations. A better method is to minimize the exchange energy. This is, in fact, the way these configurations were first obtained [79]. For our longitudinal-wave helimagnet, $J(\boldsymbol{q}) = 3J_1^2/4NJ_2$ and

$$T_c = \frac{S(S+1)J_1^2}{4k_B J_2} \tag{4.21}$$

Therefore the uniform susceptibility is

$$\chi(0)_{\text{heli}} = \frac{C}{T - \left[8J_2\left(J_1 - 2J_2\right)/J_1^2\right]T_c}. \tag{4.22}$$

Since $J_1 < 4J_2$, the coefficient of T_c in the denominator is always less than 1 . Thus the uniform susceptibility will be finite at the transition temperature. However, its high-temperature behavior will resemble that of a ferromagnet if $J_1 > 2J_2$ and that of an antiferromagnet if $J_1 < 2J_2$.

Ferrimagnetism. In the preceding three cases we have assumed that the spins and their environments were all identical. If some spins or some environments are different, then we have an impurity or alloy problem. However, if there are two different sites, each of which is translationally equivalent, i.e., two sublattices, then we can have what Néel called a ferrimagnet. As an example of such a system, let us consider an NaCl -like structure with spins S_1, say, at the Na-like sites and S_2 at the Cl-like sites. Thus we have a basis of two spins, just as NaCl has a basis of two ions. The Bravais lattice is face-centered cubic.

If we denote the spin in the i th cell with the basis vector $\boldsymbol{b}$ as S_{ib} and assume a nearest-neighbor exchange interaction, then the exchange Hamiltonian is

$$\mathcal{H}_{\text{ex}} = -J\sum_{i,j}\sum_{\boldsymbol{b},\boldsymbol{b}'} S_{i\boldsymbol{b}}\cdot S_{j\boldsymbol{b}'} \tag{4.23}$$

where the sums involve only nearest-neighbor spins. The reciprocal lattice associated with our system is that appropriate to the face-centered-cubic Bravais lattice. Therefore the Fourier expansion of the spins will involve wave vectors whose components are multiples of $2\pi/L$ up to $\pi/2a$, where L is the crystal dimension and a is the ionic spacing. Applying an external field with a spatial variation characterized by one of these wave vectors gives a Zeeman Hamiltonian of the form

$$\mathcal{H}_Z = g\mu_B H\sum_{i,\boldsymbol{b}} S_{i\boldsymbol{b}}^z\cos\left(\boldsymbol{q}\cdot\boldsymbol{R}_i + \boldsymbol{q}\cdot\boldsymbol{b}\right) \tag{4.24}$$

In Fourier components, the total Hamiltonian is

$$\begin{aligned}\mathcal{H} = &-\sum_{\boldsymbol{q}'}\sum_{\boldsymbol{b}\neq\boldsymbol{b}'} J\left(-\boldsymbol{q}'\right)S_{\boldsymbol{b}}\left(\boldsymbol{q}'\right)\cdot S_{\boldsymbol{b}'}\left(-\boldsymbol{q}'\right) + \frac{1}{2}g\mu_B H\\ &\times\sum_{\boldsymbol{b}}\left[S_{\boldsymbol{b}}^z(\boldsymbol{q})e^{-i\boldsymbol{q}\cdot\boldsymbol{b}} + S_{\boldsymbol{b}}^z(-\boldsymbol{q})e^{i\boldsymbol{q}\cdot\boldsymbol{b}}\right]\end{aligned} \tag{4.25}$$

where

$$J(-\boldsymbol{q}) = \frac{J}{N}\sum_{\delta}{}' e^{-i\boldsymbol{q}\cdot\boldsymbol{\delta}} \tag{4.26}$$

The prime indicates that $\boldsymbol{\delta}$ is restricted to 0 and five vectors to nearest neighbors, which depend on the definition of the basis. If we take $\boldsymbol{b} = a\hat{\boldsymbol{x}}$, then

$$J(-\boldsymbol{q}) = \frac{2J}{N}\left(\cos \boldsymbol{q}_x a + \cos \boldsymbol{q}_y a + \cos \boldsymbol{q}_z a\right) \exp\left(-i\boldsymbol{q}_x a\right). \tag{4.27}$$

The effective field seen by the spin S_b is

$$\boldsymbol{H}_{\boldsymbol{b}}(-\boldsymbol{q})_{\text{eff}} = -\frac{J(-\boldsymbol{q})V}{g\mu_B}\left\langle S^z_{\boldsymbol{b}'}(-\boldsymbol{q})\right\rangle \hat{\boldsymbol{z}} + \frac{1}{2}HVe^{-i\boldsymbol{q}\cdot\boldsymbol{b}}\hat{\boldsymbol{z}}. \tag{4.28}$$

Making use of (4.10), we have

$$\left\langle S^z_{\boldsymbol{b}}(\boldsymbol{q})\right\rangle = -\frac{\chi_0(\boldsymbol{q})}{g\mu_B}H_{\boldsymbol{b}}(\boldsymbol{q})_{\text{eff}} = \frac{C_{\boldsymbol{b}}J(\boldsymbol{q})V}{g^2\mu_B^2 T}\left\langle S^z_{\boldsymbol{q}'}(\boldsymbol{q})\right\rangle - \frac{C_{\boldsymbol{b}}HV}{2g\mu_B T}e^{-i\boldsymbol{q}\cdot\boldsymbol{b}} \tag{4.29}$$

with a similar equation for $\langle S^z_{b'}(\boldsymbol{q})\rangle$. Solving these two equations gives

$$\left\langle S^z_{\boldsymbol{b}}(\boldsymbol{q})\right\rangle = -\frac{HV}{2g\mu_B T}\frac{C_{\boldsymbol{b}}e^{i\boldsymbol{q}\cdot\boldsymbol{b}} + \left[C_{\boldsymbol{b}}C_{\boldsymbol{b}'}J(\boldsymbol{q})V/g^2\mu_B^2 T\right]e^{i\boldsymbol{q}\cdot\boldsymbol{b}'}}{1 - C_{\boldsymbol{b}}C_{\boldsymbol{b}'}J(\boldsymbol{q})^2V^2/g^4\mu_B^4T^2}. \tag{4.30}$$

Since

$$\begin{aligned} M_z(\boldsymbol{q}) &= -g\mu_B\int d\boldsymbol{r}e^{i\boldsymbol{q}\cdot\boldsymbol{r}}\sum_{i,\boldsymbol{b}}\left\langle S^z_{i\boldsymbol{b}}\right\rangle\delta\left(\boldsymbol{r} - \boldsymbol{R}_i - \boldsymbol{b}\right) \\ &= -g\mu_B\sum_{\boldsymbol{b}}\left\langle S_{\boldsymbol{b}}(\boldsymbol{q})\right\rangle e^{-i\boldsymbol{q}\cdot\boldsymbol{b}} \end{aligned} \tag{4.31}$$

we obtain for the susceptibility of our ferrimagnetic system,

$$\chi(\boldsymbol{q}) = \frac{\left(C_1 + C_2\right)T + 2C_1C_2J(\boldsymbol{q})\cos\left(q_x a\right)V/g^2\mu_B^2}{T^2 - C_1C_2J(\boldsymbol{q})^2V^2/g^4\mu_B^4}. \tag{4.32}$$

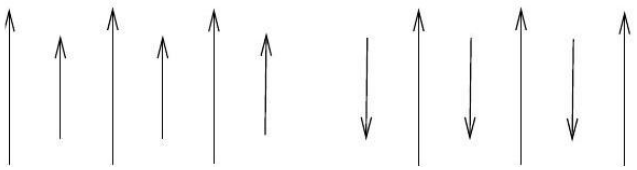


Fig. 4.4. Two possible basis configurations for a ferrimagnet

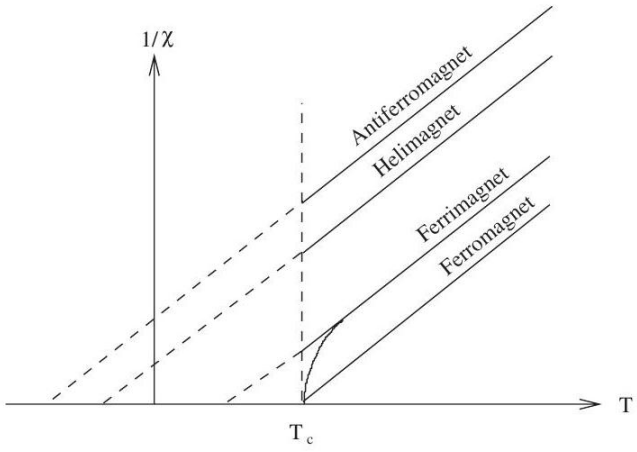


Fig. 4.5. General behavior of the high-temperature inverse susceptibility associated with various systems

From the nature of $J(\boldsymbol{q})$, this susceptibility diverges for $\boldsymbol{q} = 0$. Furthermore, this divergence is independent of whether J is positive or negative. If it is positive, the two spins within the basis are parallel; if it is negative, they are antiparallel. These two possible configurations are shown in Fig. 4.4. The uniform susceptibility of this ferrimagnetic system is

$$\chi(0)_{\text{ferri}} = \frac{(C_1 + C_2)\,T + 2\sqrt{C_1 C_2}T_c}{T^2 - T_c^2}, \tag{4.33}$$

where the transition temperature is

$$T_c = \frac{6J\sqrt{S_1 S_2\,(S_1+1)\,(S_2+1)}}{3k_B}. \tag{4.34}$$

The high-temperature susceptibilities associated with the various types of magnetic order are shown in Fig. 4.5. The rare-earth iron garnets, $RE_3Fe_5O_{12}$, are typical examples of ferrimagnets.

Although our preceeding discussion has focused on lattices, the existence of such atomic long-range order is not required for magnetic long-range order. There are a large number of materials that can be prepared in the amorphous state which show magnetic order. In particular, amorphous films of rare-earth and transition metals, such as Gd-Co and Tb-Fe, are ferrimagnetic.

Below the ferromagnetic transistion temperature the system is characterized by a spontaneous uniform magnetization. In most ferromagnets the magnetization develops smoothly from zero, identifying such transitions as second order. A typical magnetization curve is shown in Fig. 4.6.

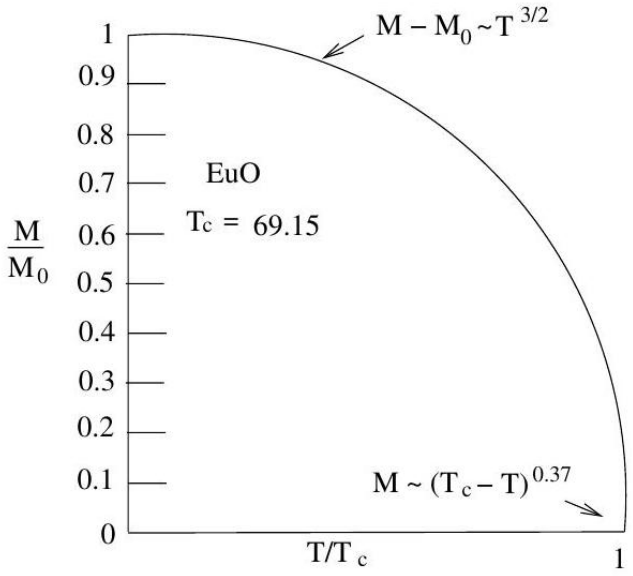


Fig. 4.6. Magnetization of europium oxide as a function of temperature [80]

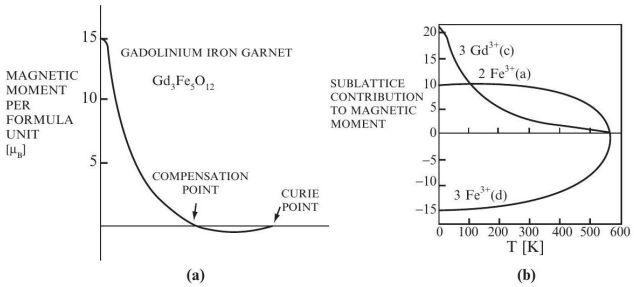


Fig. 4.7. Magnetization of gadolinium iron garnet as a function of temperature (a). The garnets have three inequivalent magnetic sublattices. The " a " ions are arranged on a bcc lattice in octahedral sites; the " c " and " d " ions sit in dodecahedral and tetrahedral sites, respectively. The dominant exchange interaction is an antiferromagnetic coupling between the a and d sites. The calculated contributions to the magnetization from these sublattices is shown in (b) [81]

The magnetization curve for a ferrimagnet is more complex. Figure 4.7 shows the net magnetization associated with a two-component, antiferromagnetically coupled ferrimagnet. Because the temperature dependences of the individual components are different, the net magnetization may go to zero at a point called the compensation temperature. This temperature is very important in certain applications involving the switching of the magnetization.
Spin Glasses. In a metal the exchange between local moments can be positive (ferromagnetic) or negative (antiferromagnetic) (see Sect. 9.2.1). This means

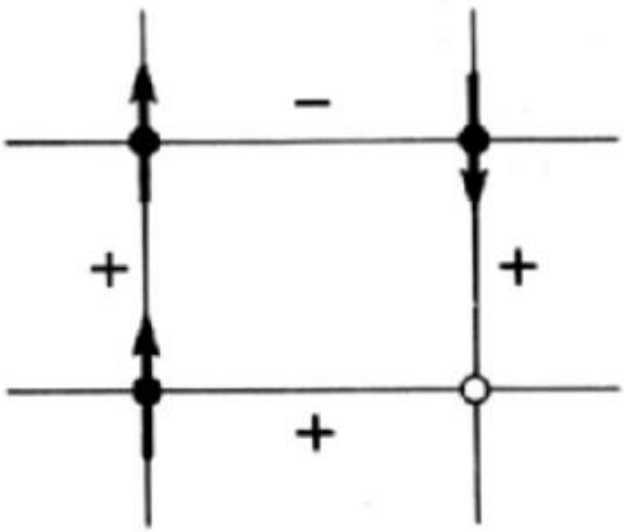


Fig. 4.8. Simple example of frustration. The signs refer to the exchange. Thus, if the three spins are aligned as indicated, the fourth will be frustrated

that some spins may find themselves in configurations where it is not possible to satisfy all the exchange interactions. Such a situation is illustrated in Fig. 4.8 and is referred to as "frustration". The first model of a spin system incorporating frustration was suggested by Edwards and Anderson [82]. Their model consists of spins on every site of a regular lattice interacting with z neighbors through an exchange J_{ij}, which is random and governed by a probability distribution $P(J_{ij})$. To see the qualitative implications of this model, consider the mean-field equation (4.35) for the thermal average of a spin $\frac{1}{2}$,

$$\langle S_i \rangle = \tanh\left(\sum_j J_{ij} \langle S_j \rangle / k_B T\right). \tag{4.35}$$

Linearizing gives

$$\langle S_i \rangle = \sum_j J_{ij} \langle S_j \rangle / k_B T. \tag{4.36}$$

Averaging over $J_{ij} = 0$ leads to the result that $\overline{\langle S_i \rangle} = 0$. However, if we square (4.36) and again average over the exchange values, we obtain

$$\overline{\langle S_i \rangle^2} = \frac{z}{(k_B T)^2} \overline{J_{ij}^2} \overline{\langle S_j \rangle^2}. \tag{4.37}$$

This means that $\overline{\langle S_i \rangle^2}$ becomes nonzero at a critical temperature $T_f = z\overline{J_{ij}^2}/k_B$ and is identified as the order parameter q. The fact that $\overline{\langle S_i \rangle^2} \neq 0$, while $\overline{\langle S_i \rangle} = 0$, implies a spin configuration in which each spin is frozen in some specific direction, but this direction varies randomly throughout the alloy. This is generally what is meant by a "spin glass". The mean-field solution to the Edwards-Anderson model gives a cusp in the low-field ac susceptibility as observed experimentally (Fig. 4.9). However, the theory also predicts a cusp in the specific heat at T_f which is not observed.

A difficulty in comparing experimental results with random exchange models is that the data seem to suggest that the actual systems may be more complex. The ac susceptibility, for example, is slightly frequency dependent,

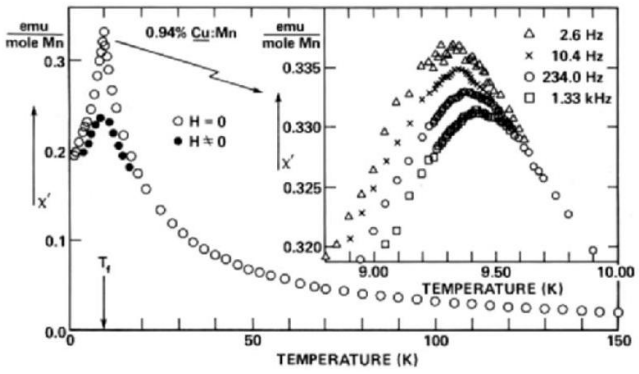


Fig. 4.9. ac susceptibility as a function of temperature for $\mathbf{C u : \mathrm{Mn}$ [83]}

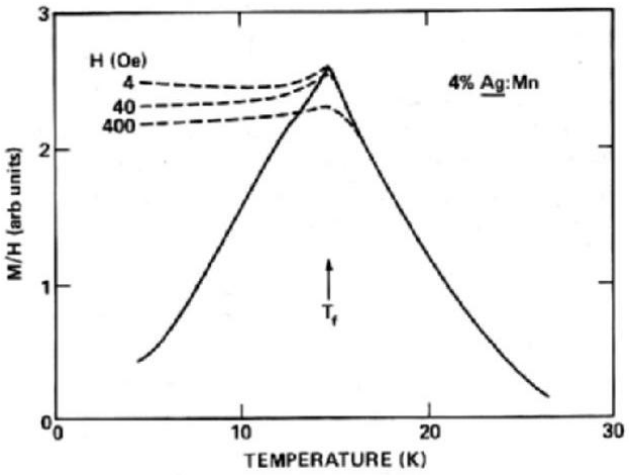


Fig. 4.10. dc susceptibility of Ag:Mn [84]

as shown in the insert in Fig. 4.9. The dc susceptibility is also marked by timedependent and irreversible behavior. Figure 4.10 shows the dc susceptibility of Mn in Ag . If the sample is cooled in zero magnetic field (which involves cancellation of the earth's field) and then the susceptibility measured with increasing temperature, one obtains the solid curve. On the other hand, if one cools the sample in the presence of the field which is used to measure susceptibility, then one obtains the larger values indicated by the dashed lines. These field-cooled susceptibilities are reversible whereas the zero-field-cooled curve is not. Furthermore, if one sits at a given temperature $T < T_f$ on the zero-field-cooled curve, the magnetization increases logarithmically with time, eventually approaching the field-cooled value. This behavior is believed to be associated with the existence of a large number of low-lying states connected by energy barriers. These barriers restrict the phase space available to the system and the low temperature magnetic properties are governed by the shape of these wells.

10.5.2 4.2 Low Temperatures

Let us now turn briefly to the question of how the systems we have just discussed respond to a static field when they are below their ordering temperature. In the treatment above we considered only an isotropic exchange interaction. If this were the only contribution to the Hamiltonian, then the Hamiltonian, as well as its eigenstates, would be invariant under rotations. This would mean that the isotropic exchange interaction itself could not lead to a macroscopically ordered state. There are, in fact, additional interactions which favor certain crystallographic directions. Nickel, for example, prefers to magnetize along its $\left[\begin{array}{ccc} 1 & 1 & 1 \end{array}\right]$ direction while iron prefers the $\left[\begin{array}{ccc} 1 & 0 & 0 \end{array}\right]$ direction. Such magnetic anisotropy has its origin in the spin-orbit interaction as we discussed in Sect. 2.3.1. In metals this interaction leads to shifts in the energy bands as the magnetization direction is varied. Such shifts are accompanied by an electron redistribution among the bands with a corresponding change in the total electronic energy.

Macroscopically, the magnetic anisotropy is characterized by an energy which is a function of the direction cosines of the magnetization relative to the crystallographic axes. A uniaxial anistropy, for example, would be represented by $K_1 \sin^2 \theta$. Microscopically, at least for an ionic solid, the anisotropy appears as terms in the Hamiltonian involving products of the components of the spin. In particular, for a uniaxial system,

$$\mathcal{H}_{\text{aniso}} = -D \sum_i (S_i^z)^2 \,, \tag{4.38}$$

where D is the anisotropy constant. The total effective field acting on the i th spin in the molecular field approximation is therefore

$$\boldsymbol{H}_{i,\text{eff}} = -\frac{1}{g\mu_B}\sum_{j\neq i} J_{ij}\left\langle S_j\right\rangle - \frac{D}{g\mu_B}\left\langle S_i^z\right\rangle \hat{\boldsymbol{z}} + \boldsymbol{H}\cos\left(\boldsymbol{q}\cdot\boldsymbol{R}_i\right). \tag{4.39}$$

Because there is now an inherent direction in the system, the response to an applied field will depend upon the direction of such a field. Let us first consider the case in which the applied field is parallel to the anisotropy axis. Then

$$\left\langle S_j\right\rangle = \left\langle S_i^z\right\rangle_0 \cos\left[\boldsymbol{Q}\cdot\left(\boldsymbol{R}_j-\boldsymbol{R}_i\right)\right]\hat{\boldsymbol{z}} + \delta S_j^z\hat{\boldsymbol{z}}, \tag{4.40}$$

where $\left\langle S_i^z\right\rangle_0$ is the average value of S_i^z in the absence of the applied field and Q characterizes the order. Since the exchange energy exceeds the thermal energy in the ordered state, this is given by the Brillouin function as

$$\left\langle S_i^z\right\rangle_0 = -SB_S\left(-\frac{\left[NSJ(\boldsymbol{Q})+DS\right]\left\langle S_i^z\right\rangle_0}{k_BT}\right). \tag{4.41}$$

The term δS_j^z in (4.40) is that part of $\left\langle S_j^z\right\rangle$ which is induced by the applied field. This response is what defines the susceptibility. The average value of S_i^z
in the presence of the field is given by an expression similar to (4.41), with the applied field added to the argument of the Brillouin function as well as terms involving δS_i^z. However, if all these additional terms are small, we may expand the Brillouin function in powers of these terms. To lowest order, we obtain

$$\begin{aligned}\left\langle S_i^z\right\rangle = \left\langle S_i^z\right\rangle_0 &+ \frac{NS^2B_S'}{k_BT}\frac{1}{N}\sum_{j\neq i} J_{ij}\delta S_j^z \\ &+ \frac{DS^2B_S'}{k_BT}\delta S_i^z - \frac{g\mu_BS^2HB_S'}{k_BT}\cos\left(\boldsymbol{Q}\cdot\boldsymbol{R}_i\right),\end{aligned} \tag{4.42}$$

where B_S' is the derivative of the Brillouin function with respect to its argument, evaluated at zero field. If we take the Fourier transform of (4.39), the susceptibility is found to be

$$\chi_{\parallel}(\boldsymbol{q}) = \frac{3SCB_S'/(S+1)}{T-\left[3SB_S'/(S+1)\right]\left[NJ(\boldsymbol{q})+D\right]S(S+1)/3k_B}. \tag{4.43}$$

By similar arguments, the response to a field applied perpendicular to the z-axis is characterized by

$$\chi_{\perp}(\boldsymbol{q}) = \frac{g^2\mu_B^2N/V}{N[J(\boldsymbol{Q})-J(\boldsymbol{q})]+D}. \tag{4.44}$$

From these expressions we find that the perpendicular susceptibility is independent of temperature, while the parallel susceptibility decreases to 0 as the temperature does. This leads to a rather interesting phenomenon in antiferromagnets. At zero temperature the application of a parallel field produces no polarization. Thus there is no Zeeman energy associated with this state. A perpendicular field, however, produces a small canting of the spins which does lead to a polarization. The Zeeman energy associated with this state is essentially $-\boldsymbol{M}\cdot\boldsymbol{H}$. Thus, if there were no anisotropy energy, the spins in a parallel field would all "flop" over into the lower-energy perpendicular configuration. Because of the anisotropy energy there is a minimum field at which this "spin-flop" transition occurs. By equating the energies in the two states we find, for nearest-neighbor interactions,

$$H_{\text{crit}} = \frac{S}{g\mu_B}\sqrt{4zD|J|}, \tag{4.45}$$

which is the geometrical mean of the exchange field and the anistropy field. This will not happen in a ferromagnet, even though the parallel response function does go to 0 , because of the large dipolar energy associated with the aligned state. Figure 4.11 shows the phase diagram for an antiferromagnet.

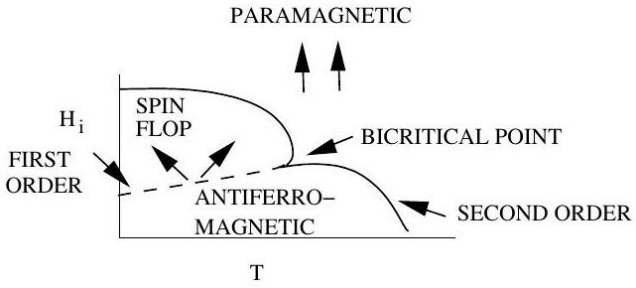


Fig. 4.11. Phase diagram for an antiferromagnet ($H_{i $ is the internal field)}

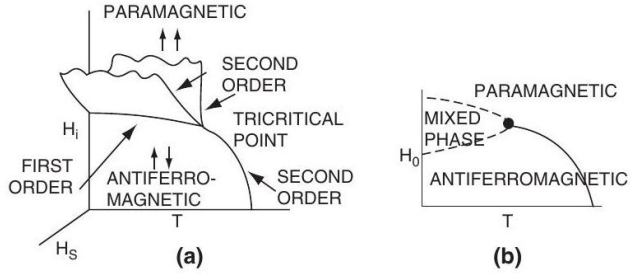


Fig. 4.12. Phase diagram for a metamagnet in terms of the internal field, $H_{i $ (a) and the applied field, H_0 (b)}

Metamagnetism. In a material with strong anistropy the spin-flop state is suppressed and below a certain temperature the system goes directly into the paramagnetic state via a first-order transition. Such systems are called metamagnets. Classic examples are $FeCl_2$ and $DyPO_4$. A schematic phase diagram is shown in Fig. 4.12a. The point at which the line of second-order transitions changes into a line of first-order transitions is called the tricritical point. In an antiferromagnet the field conjugate to the staggered magnetization is the staggered field $H(\boldsymbol{q}) \equiv H_s$. Griffiths [85] has pointed out that in the $H_i - H_s - T$ phase diagram for a metamagnet there exist two first-order surfaces extending like wings out into the $H_s \neq 0$ region as shown in Fig. 4.12a. Notice that these wings are bounded by two lines of second-order transitions. Thus the tricritical point is also the meeting of three critical lines. In finite samples the internal field H_i and the applied field H_0 differ due to demagnetizing effects. For an elliptical sample characterized by a demagnetizing factor N, the internal field is given by $H_i = H_0 - 4\pi NM$. Since the magnetization increases discontinuously as we cross the line of first-order transitions, the internal field is discontinuously reduced to a value below that required for the transition. The material resolves this contradiction by breaking into domains of two phases. This mixed phase is analogous to the so-called intermediate state in superconductivity.

10.5.3 4.3 Temperatures Near T_c

Let us now turn to the interesting region of the critical point itself. The object of any critical-point theory is to predict accurately the behavior of various physical quantities as the temperature or some other variable approaches its critical-point value. Generally this approach is represented by a power law such as $(T - T_c)^{\lambda}$. Figure 4.6, for example, shows that the magnetization of EuO vanishes as $(T - T_c)^{0.37}$. A successful theory must correctly predict the value of λ, as well as any relationship, if it exists, between the exponents associated with different physical quantities.

There are two basic steps in this problem. First a model must be selected that adequately represents the actual system but is still mathematically tractable. Our discussions so far have been based on the Heisenberg Hamiltonian, $-\sum J_{ij} S_i \cdot S_j$. We have also mentioned the Ising model, - $\sum J_{ij} S_i^z S_j^z$. Although some materials, such as dysprosium aluminum garnet, may actually be described by this model, its main attraction is that it is much easier to deal with than the Heisenberg model. In fact, the thermodynamic quantities associated with a two-dimensional Ising model may be computed exactly [86].

Since even the relatively simple Ising model has not yet been solved exactly in three dimensions, we are forced to make certain approximations. Therefore the second step in the critical-point problem is the development of mathematical techniques that enable us to calculate physical properties near the phase transition of our model system. In Sect.4.1, we employed the meanfield approximation for low temperatures where each spin was assumed to be statistically independent; that is, $\left\langle S_i^z S_j^z \right\rangle_0 = \langle S_i^z \rangle_0 \left\langle S_j^z \right\rangle_0$, where the subscript indicates that this average is to be taken in zero field. In the hightemperature region we employed the random-phase approximation, in which only the Fourier components were assumed independent. Notice that use of the molecular field approximation in this region would have implied that the energy of the system, which involves $\left\langle S_i^z S_j^z \right\rangle_0$, was 0 above T_c. Thus the specific heat would also have been 0 . In actual fact, however, as T_c is approached from above the spins begin to correlate with one another. That is, $\left\langle S_i^z S_j^z \right\rangle$ is not 0 for spins that are relatively close together. We speak of this as the appearance of short-range order. Owing to the existence of this short-range order, the energy, and consequently the specific heat, are not 0 above the point where the spontaneous magnetization vanishes.

The determinations of the correlation function $\left\langle S_i^z S_j^z \right\rangle_0$ provides us with a nice application of the fluctuation-dissipation theorem. Introducing Fourier components and requiring that this correlation function depend only upon the relative coordinates of the spins gives

$$\left\langle S_i^z S_j^z \right\rangle_0 = \frac{1}{N^2} \sum_{\boldsymbol{q}} \langle S(\boldsymbol{q})_z S(-\boldsymbol{q})_z \rangle_0 \exp\left[i\boldsymbol{q} \cdot (\boldsymbol{R}_i - \boldsymbol{R}_j) \right] \tag{4.46}$$

Since $\mathcal{M}_\eta(\boldsymbol{q}) = -g\mu_B S_\eta(\boldsymbol{q})$, the fluctuation-dissipation theorem (1.88) gives

$$\langle S(\boldsymbol{q}, t)_z S(-\boldsymbol{q})_z \rangle_0 = \int \frac{d\omega}{2\pi} \frac{\hbar V}{g^2 \mu_B^2} \coth(\beta\hbar\omega/2) \chi''_{zz}(\boldsymbol{q}\omega) e^{-i\omega t} \tag{4.47}$$

Setting $t = 0$ and assuming that $k_B T/\hbar$ is larger than any relevant frequencies (we shall discuss the validity of this assumption in Chap. 6), we find

$$\langle S(\boldsymbol{q})_z S(-\boldsymbol{q})_z \rangle_0 = \frac{k_B T V}{\pi g^2 \mu_B^2} \int \frac{\chi''_{zz}(\boldsymbol{q}, \omega)_s}{\omega} d\omega \tag{4.48}$$

By the Kramers-Kronig relation (1.64), the frequency integral is just π times the static susceptibility $\chi_{zz}(\boldsymbol{q}, 0)$. Therefore we obtain

$$\left\langle S_i^z S_j^z \right\rangle_0 = \frac{k_B T V}{N^2 g^2 \mu_B^2} \sum_{\boldsymbol{q}} \chi_{zz}(\boldsymbol{q}) \exp\left[i\boldsymbol{q} \cdot (\boldsymbol{R}_i - \boldsymbol{R}_j)\right] \tag{4.49}$$

Notice that since $\left\langle (S_i^z)^2 \right\rangle = S(S+1)/3$, (4.49) leads to what we refer to as a sum rule

$$\frac{1}{N} \sum_{\boldsymbol{q}} \chi(\boldsymbol{q}) = \chi_0 \tag{4.50}$$

where χ_0 was defined in (4.10).
If we use the random-phase result for $\chi_{zz}(\boldsymbol{q})$ (4.12) and assume a simple cubic lattice, then in the long-wavelength limit (4.49) becomes

$$\left\langle S_i^z S_j^z \right\rangle_0 = \frac{6S(S+1)}{3Na^2} \frac{T}{T_c} \sum_{\boldsymbol{q}} \frac{\exp\left[i\boldsymbol{q} \cdot (\boldsymbol{R}_i - \boldsymbol{R}_j)\right]}{\kappa^2 + q^2}, \tag{4.51}$$

where

$$\kappa^{-1} = a \sqrt{\frac{T_c}{6\,(T - T_c)}} \tag{4.52}$$

is referred to as the correlation length. Notice that as T becomes very large in comparison with T_c, the correlation length becomes very small and

$$\left\langle S_i^z S_j^z \right\rangle_0 \underset{T>T_c}{\longrightarrow} \frac{1}{3} S(S+1) \Delta\left(\boldsymbol{R}_i - \boldsymbol{R}_j\right) \tag{4.53}$$

As T approaches T_c the correlation function takes the so-called Yukawa-like form

$$\left\langle S_i^z S_j^z \right\rangle_0 = \frac{S(S+1)}{16\pi^4 (N/V) a^2} \frac{T}{T_c} \frac{\exp\left(-\kappa \left|\boldsymbol{R}_i - \boldsymbol{R}_j\right|\right)}{\left|\boldsymbol{R}_i - \boldsymbol{R}_j\right|} \tag{4.54}$$

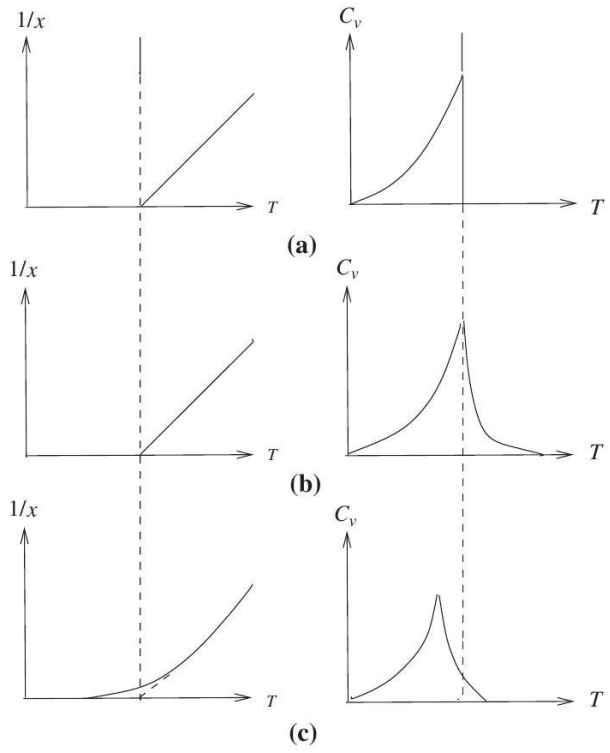


Fig. 4.13. The effect of various approximations on the susceptibility and the specific heat of a Heisenberg ferromagnet: (a) molecular field, (b) random phase, and (c) typical experimental behavior

Thus as T approaches T_c each spin exerts a polarizing influence on those spins within a sphere of radius κ^{-1}. At $T = T_c$ this sphere encompasses the whole sample, and we say that long-range order has set in.

The specific heat associated with this ordering may also be computed from (4.49). The behavior of the susceptibility and the specific heat in the mean-field and random phase approximations are compared schematically with typical experimental data in Fig. 4.13. The most notable discrepancy is the fact that the actual transition occurs below the value predicted by these theories. This is the consequence of an inadequate treatment of short-range order.

We can obtain a better approximation for $\chi(\boldsymbol{q})$ by explicitly taking shortrange order in account. The basic idea is that the spins surrounding any given spin will be correlated to the motion of that spin and will, therefore, not contribute to the mean field seen by that spin. This is incorporated into the theory by subtracting a term from the mean field which is proportional to $\langle S_i^z \rangle$ itself,

$$H_i^{\text{eff}} = -\sum_j \frac{J_{ij}}{g\mu_B} \langle S_j^z \rangle + \sum_j \frac{\lambda_{ij} J_{ij}}{g\mu_B} \langle S_i^z \rangle \,. \tag{4.55}$$

The parameters λ_{ij} are assumed to be temperature dependent and characterize this so-called Onsager reaction field. Proceeding as in the derivation of (4.12) we obtain

$$\chi(\boldsymbol{q}) = \frac{\chi_0}{1 - \frac{\chi_0 V}{g^2 \mu_B^2}[J(\boldsymbol{q}) - \lambda]} \tag{4.56}$$

where $\lambda = (1/N) \sum_j \lambda_{ij} J_{ij}$. This parameter is obtained by requiring that the susceptibility (4.53) satisfies the sum rule (4.50). The result is

$$\lambda_{ij} = \frac{\left\langle S_i^z S_j^z \right\rangle_0}{S(S+1)/3} \tag{4.57}$$

Thus, (4.46) now becomes a self-consistent equation for the correlation function $\left\langle S_i^z S_j^z \right\rangle_0$.

This improved treatment of short-range order also renormalizes the transition temperature relative to the mean-field value T_c. If we sum (4.53) over $\boldsymbol{q}$ and again use the sum rule we obtain the relation

$$G(S) = T_c/T \tag{4.58}$$

where $S = T/T_c + \lambda/J(\boldsymbol{q})$ and the sum

$$G(S) = \frac{1}{N}\sum_{\boldsymbol{q}} \frac{1}{S - J(\boldsymbol{q})/J(\boldsymbol{q})} \tag{4.59}$$

is called the lattice Green's function. From (4.53) we see that $\chi(\boldsymbol{q})$ diverges at a temperature which corresponds to $S = 1$. Therefore,

$$T_c^* = T_c/G(1) \tag{4.60}$$

For simple cubic, body-centered-cubic, and face-centered-cubic lattices the values of $G(1)$ are 1.517, 1.393, and 1.345, respectively. Thus we see that incorporating the Onsager reaction field predicts a lower transition temperature as we observed in Fig. 4.13.

Just below T_c the quantity of thermodynamic interest is the spontaneous magnetization. As T_c is approached from below, $\langle S_i^z \rangle_0$ becomes very small, which enables us to expand the Brillouin function in (4.41). The resulting magnetization in this molecular field approximation is found to vanish as $(T_c - T)^{1/2}$.

At very low temperatures the molecular field expression (4.41) predicts that the magnetization decreases with increasing temperature as $e^{-2T_c/(S+1)T}$. The decrease is found experimentally to be much less rapid. The reason for this has to do with the existence of spin-wave modes, which we shall discuss in Chap. 8.

10.5.4 4.4 Landau Theory of Second-Order Transitions

This idea of treating the magnetization as an expansion parameter is a fundamental aspect of Landau's theory of second-order phase transitions. Landau associated second-order phase transitions with transitions in which there is a "broken symmetry". That is, the new ground state of the sytem does not possess the total symmetry of the Hamiltonian. In the case of ferromagnetism the rotational invariance of the Heisenberg Hamiltonian is broken by the appearance of the spontaneous magnetization M. Landau identified the parameter whose value becomes nonzero in the unsymmetrical state as the order parameter. In this case it is obviously the magnetization. The order parameter represents an additional variable that must now be used to specify the state of the system. For the Heisenberg model the order parameter is a vector. We therefore speak of it having the dimensionality $n = 3$. Since an Ising model involves only the z-component of the spin, the order parameter has the dimensionality $n = 1$.

If the order parameter, call it M, is to grow from zero above T_c to a finite value below T_c then the Landau theory argues that near T_c the free energy has the form

$$F = a\,(T - T_c)\,M^2 + BM^4. \tag{4.61}$$

where a and B are constants. Minimizing this energy with respect to M tells us that $M \sim (T_c - T)^{1/2}$ as we found above. If we add an applied field term, $-MH$, then it can be shown that

the susceptibility diverges as $\chi \sim (T - T_c)^{-1}$. Similarly, if we assume that the magnetization varies slowly in space, then there will be an additional term in the free energy. Since it costs virtually no energy to rotate the magnetization uniformly, we expect that the energy associated with a gradual variation may be expressed in terms of derivatives of $\boldsymbol{M}$. In order to satisfy time reversal, this must be a quadratic expression. There are several such possibilities. In an isotropic medium, for example, there are three: $|\nabla \boldsymbol{M}|^2 = |\nabla M_x|^2 + |\nabla M_y|^2 + |\nabla M_z|^2$, $(\nabla \cdot \boldsymbol{M})^2$, and $|\nabla \times \boldsymbol{M}|^2$. If we assume that all those terms reflecting the spatial symmetry are taken into account through an anisotropy energy, then the exchange energy must be invariant with respect to rotations in spin space. For example, a rotation of π about the z-axis in spin space takes $M_x \to -M_x$ and $M_y \to -M_y$, but leaves x and y unchanged. Thus, the exchange energy in a cubic crystal must have the form

$$\frac{\delta E}{V} = \frac{A}{M_0^2} |\nabla \boldsymbol{M}|^2, \tag{4.62}$$

where A is referred to as the exchange stiffness. This contribution introduces a characteristic length, the correlation length, which varies as $\xi \sim (T_c - T)^{-1/2}$, as we have already found. These critical exponents,

$$\begin{aligned} \beta &= \frac{1}{2} \text{ for the magnetization,} \\ \lambda &= 1 \text{ for the susceptibility, and} \\ \eta &= \frac{1}{2} \text{ for the correlation length,} \end{aligned}$$

do not agree with those observed experimentally. This is not too surprising for the theory itself predicts that as $T \to T_c$ fluctuations in the order parameter become larger than the order parameter itself. Careful measurements of these and other critical exponents in the 1960s stimulated much theoretical activity. It was particularly intriguing that particular critical exponents in different systems seemed to have nearly the same values. Furthermore, certain combinations of exponents seemed to be related.

10.5.5 4.5 Critical Phenomena

The first major assault on this problem was Widom's hypothesis that the critical part of free energy, G, was a homogeneous function of its independent variables. We say that a function $G(\epsilon, H)$ is homogeneous if it can be written in the form

$$G\left(\lambda^x \varepsilon, \lambda^y H\right) = \lambda G(\varepsilon, H) \tag{4.63}$$

where $\epsilon = (T - T_c)/T_c, x$ and y are two critical indices, and λ was subsequently shown by Kadanoff to correspond to a scaling length. With this "scaling" hypothesis all the critical exponents now become combinations of x and y. In 1971 K . Wilson introduced the renormalization group which clarified the nonanalytic nature of the phase transition and provided the basis for scaling. Various methods have been developed for performing renormalization-group calculations for the critical exponents themselves. Such calculations show that the critical exponents are only functions of the dimensionality of space and the dimensionality of the order parameter, a situation referred to as universality (see [87]).

It is difficult to make accurate comparisons with experiment. First of all, impurities and strains tend to smear out the transition, making it difficult to measure close to the transition. And also, weak anisotropies become important near the transition and can lead to "crossover" from one universality class to another. These difficulties are illustrated in Fig. 4.14, where we have plotted the critical exponent β for four representative materials among the constant β

contours in the n, d-plane obtained from a renormalization-group calculation [88]. Dysprosium aluminum garnet (DAG) is considered to be a three-dimensional Ising system, K_2NiF_4 a two-dimensional Heisenberg system, and $GdCl_3$ a three-dimensional Heisenberg system but one in which long-range dipolar interactions are important.

Table 4.1. The critical exponent β for selected materials

Material	β	$\frac{T_c - T}{T_c}$	Refs.
$GdCl_3$	$\beta = 0.43$	$0.002 \leq \epsilon \leq 0.07$	[90]
DAG	$\beta = 0.26$	$0.001 \leq \epsilon \leq -.056$	[91]
K_2NiF_4	$\beta = 0.138$	$0.0003 \leq \epsilon \leq 0.2$	[92]
MnF_2	$\beta = 0.333$	$0.00006 \leq \epsilon \leq 0.08$	[93]

4.5.1 Order in 2D

Notice that the theoretical curves shown in Fig. 4.14 do not extend to $d = 2$. In two dimensions, the nature of the long-range order depends critically on the dimensionality of the order parameter. Mermin and Wagner [89] have shown that two-dimensional systems with a continuous suymmetry, such as a Heisenberg system, will not exhibit conventional long-range order as characterized, for example, by a spontaneous magnetization or a sublattice magnetization.

A good example of 2-dimensional behavior is found in the undoped hightemperature superconductor La_2CuO_4 [94]. This compound has the layered perovskite structure shown in Fig. 4.15. Since La has the oxidation state +3 and oxygen is -2 , this suggests that the copper is in a +2 oxidation state corresponding to a 3 d^9 electronic configuration. With one hole per copper atom one would expect La_2CuO_4 to be a metal. However, it is an antiferromagnetic insulator. This is presumably the result of electron-electron correlations as originally suggested by Mott.

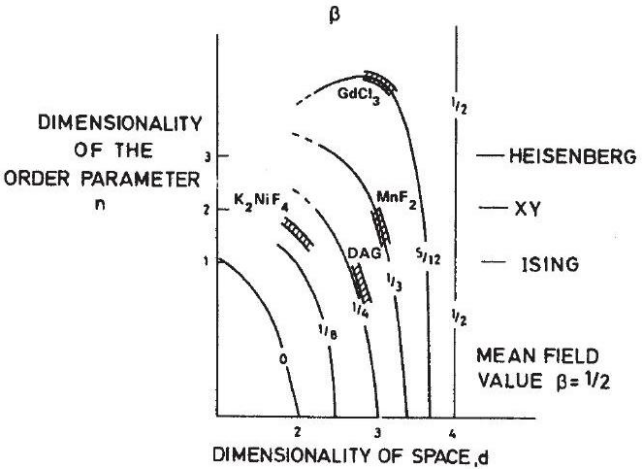


Fig. 4.14. Variation of the critical exponent β (associated with the magnetization) with the dimensionality of space (d) and order parameter (n). The experimental values for several materials are indicated (see Table 4.1 for references)

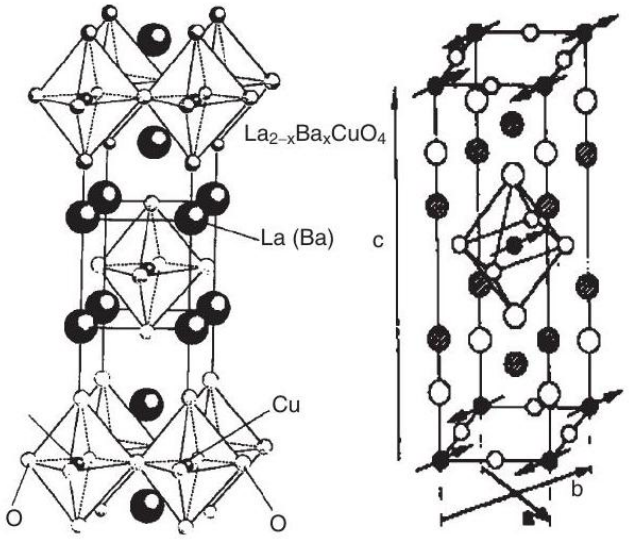


Fig. 4.15. Layered perovskite structure (Figure to right shows spin configuration of $\mathrm{La

{2} \mathrm{CuO}{4}$)}

The spin 1/2 holes in the copper layers describe a spin 1/22D square lattice quantum Heisenberg system. Chakravarty et al. [95] have shown that the correlation length of such a system varies as

$$\frac{\xi}{a} = 0.493 e^{1.15 J/k_B T},$$

which implies that ordering only occurs at $T = 0$ K. The fact that La_2CuO_4 has a relatively high Neél temperature at $T_N = 325$ K is due to anisotropy in the exchange as well as interlayer exchange. The Hamiltonian for La_2CuO_4 has been found [94] to have the form

$$= \mathcal{H} = J \left[\sum_{i,\boldsymbol{\delta}_\parallel} S_i \cdot S_{i+\boldsymbol{\delta}_\parallel} + \alpha_{xy} \sum_{i,\boldsymbol{\delta}_\parallel} S_i^c S_{i+\delta_\parallel}^c \right.$$

$$\left. + \sum_{i,\delta_\perp j} \alpha_{\perp j} S_i \cdot S_{i+\delta_\perp j} + \alpha_{DM} \sum_{i,\boldsymbol{\delta}_\parallel} (-1)^i \hat{\boldsymbol{b}} \cdot S_i \times S_{i+\boldsymbol{\delta}_\parallel} \right]$$

where $\alpha_{xy}, \alpha_{\perp j}$, and α_{DM} represent the $x - y$ anisotropy, the interlayer coupling, and the Dzyaloshinski-Moriya interaction, respectively. The index j refers to the two different out-of-plane neighbors, and the notation " c " indicates the component of the spin in the c-direction. Including the anisotropy term and the interlayer term, the Neél temperature is given by

$$k_B T_N = \left(\frac{\xi}{a} \right)^2 J \left(z_\parallel \alpha_{xy} + z_\perp \alpha_\perp \right),$$

where $z_\parallel$and $z_\perp$ are the in-plane and out-of-plane coordination numbers, respectively. The magnetic order is shown in the insert in Fig. 4.15. The

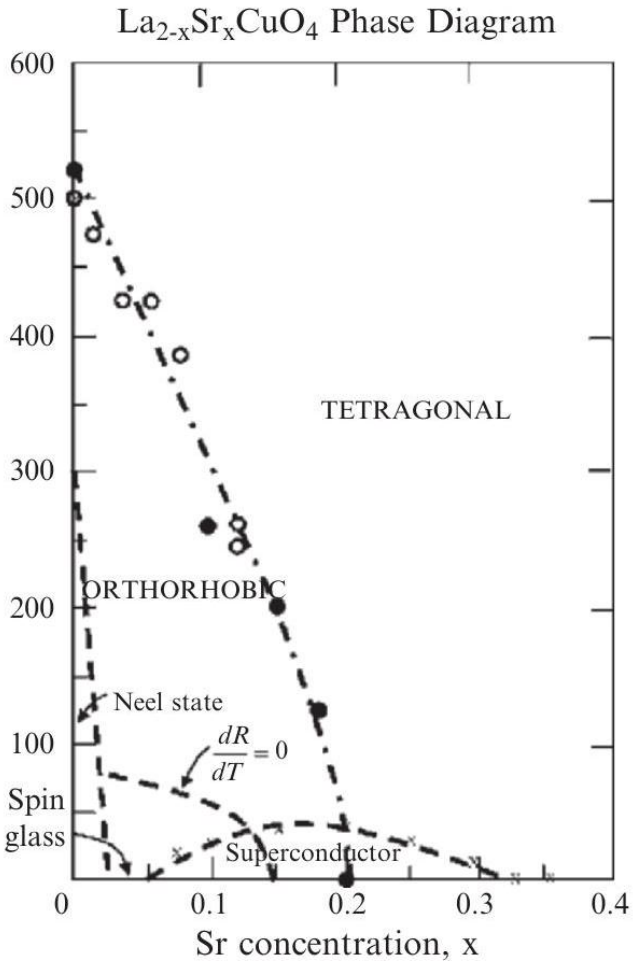


Fig. 4.16. Phase diagram for $\mathrm{La *{2-x}* *mathrm{Sr}*{x} \mathrm{CuO}_{4}$ summarizing structural, magnetic, and transport properties. The narrow dashed line ($dR/dT = 0$) separates the region of metallic linear resistance from that of logarithmically increasing resistance. The conductance in the Néel state is strongly localized. From Keimer, Belk et al. [96]}

Dzyaloshinski-Moriya term causes the spins to cant out of the plane by about 0.17°. The interlayer coupling in turn causes these small moments to orient ferromagnetically from layer to layer.

When La_2CuO_4 is doped with Sr or Ba one obtains many different phases, including the phase exhibiting superconductivity. The phase diagram is shown in Fig. 4.16. Notice the precipitous drop in the Neel temperature with Sr doping. This is due to the fact that the Sr doping introduces holes in the oxygen which reduces the superexchange.

The Mermin-Wagner Theorem mentioned above does not exclude the possibility of a phase transition associated with more complex spin dynamics. The two-dimensional XY model is the classic example. Berezinskii [97] has calculated the spin-spin correlation function $\langle S_i S_j \rangle$ for this model using a low-temperature expansion, and found a power-law behavior

$$\langle S_i S_j \rangle \sim \left(\frac{|R_i - R_j|}{R_0} \right)^{-k_B T/4\pi J} .$$

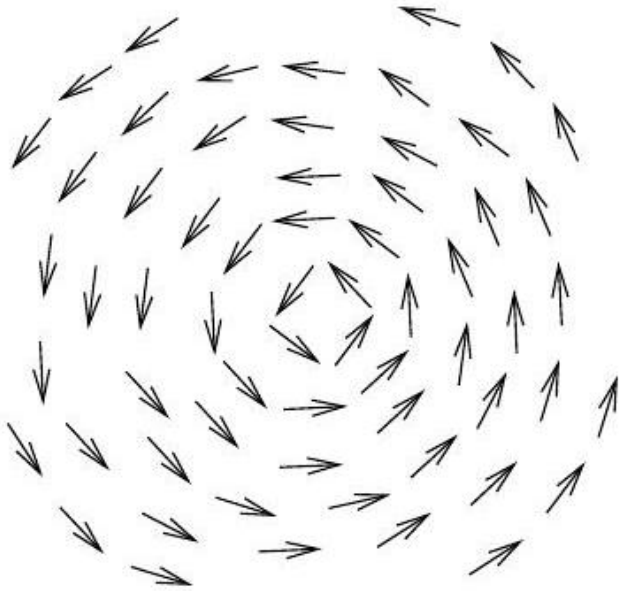


Fig. 4.17. Spin Vortex

This implies that the susceptibility is infinite, i.e., that the system is undergoing large fluctuations which destroy any long-range order. Kosterlitz and Thouless [98], however, have pointed out that, in addition to the longwavelength spin fluctuations responsible for this power-law behavior, there also exist local "topological defects" in the order parameters which, in this case, correspond to vortices as illustrated in Fig. 4.17. Other examples of "defects" in order parameters include dislocations in crystalling solids and vortices in superfluid helium. To calculate the energy of an isolated vortex, we write the XY Hamiltonian as

$$\mathcal{H}_{XY} = -2J \sum_{i>j} \left(S_i^x S_j^x + S_i^y S_j^y \right) = -2J \sum_{i>j} \cos\left(\varphi_i - \varphi_j\right), \tag{4.64}$$

where φ_i is the angle the i th spin makes with some arbitrary axis. Expanding the cosine,

$$E - E_o \approx J \sum_r [\Delta\varphi(r)]^2, \tag{4.65}$$

where Δ denotes the first difference operator. For a vortex as shown in Fig. 4.17, $\Delta\varphi(r) = 2\pi/2\pi r$. Therefore,

$$E - E_0 = 2\pi J \ln(R/a), \tag{4.66}$$

where R is the radius of the system and a the nearest-neighbor distance. Since this vortex could be centered on any one of the $\pi R^2/a^2$ sites, the entropy is

$$S = 2k_B \ln(R/a) \tag{4.67}$$

The free energy is $E - E_0 - TS$. Since both the energy and entropy depend upon the size of the system in the same logarithmic way, the energy term will dominate the free energy at low temperatures, and the probability of a vortex appearing is vanishingly small. However, when the temperature exceeds $\pi J/k_B$, the entropy dominates and free vortices are likely to appear. If one now includes interactions between vortices, the critical temperature is given by the solution of [98]

$$\frac{\pi J}{k_B T_{KT}} - 1 = 2\pi \exp\left(-\pi^2 J/k_B T_{KT}\right) \tag{4.68}$$

Below the Kosterlitz-Thouless temperature, T_{KT}, the vortices are bound in pairs of zero total vorticity. This is referred to as topological long-range order. The spin-spin correlation function still has a power-law behavior indicating an infinite susceptibility. However, above T_{KT}, the vortices give the correlation function an exponential form with a susceptibility given by

$$\chi(T) \sim \exp\left(2.625 t^{-1/2}\right) \tag{4.69}$$

where $t = (T - T_{KT})/T_{KT}$.

10.5.6 4.6 Stoner-Wohlfarth Model

The usefulness of magnetic materials generally depends upon the existence of a net magnetization at room temperature. That is, magnetic materials of practical interest are either ferromagnets or ferrimagnets with Curie temperatures above room temperature. Their electrical conductivity ranges from metallic to insulating.

The susceptibilities given by (4.17) or (4.33) are the "small signal" responses of these systems. If one applies a large enough field opposite to the direction of the magnetization it is possible to reverse the magnetization. As one cycles the magnetic field, positive and negative, the magnetization follows a "hysteresis loop" as shown in Fig. 4.18.

The shapes of hysteresis loops depend upon many extrinsic factors such as crystallographic microstructure and the role of magnetic domains. The magnetic response characterized by a hyperesis loop is refered to as the "technical" magnetization. This subject is beyond the scope of this text. However, there is a very simple model that provides a connection between the hysteresis loop and our microscopic concepts.

This model, due to Stoner and Wohlfarth [99], considers the response of an array of non-interacting magnetic particles. Let us assume these particles are ellipsoidal in shape and are uniformly magnetized as shown in Fig. 4.19. To determine the direction of the magnetization in the presence of an applied field, H_0, we must minimize the energy of the system. Let us consider a single particle. In addition to the Zeeman energy density, $-\boldsymbol{M} \cdot \boldsymbol{H}_0$, there is a magnetostatic energy density associated with the magnetization itself given by $-\frac{1}{2}\boldsymbol{M} \cdot \boldsymbol{H}_D$, where $\boldsymbol{H}_D$ is the "demagnetization" field. To solve for $\boldsymbol{H}_D$ we combine the Maxwell equation $\nabla \cdot \boldsymbol{B} = 0$ with the definition $\boldsymbol{B} = \boldsymbol{H} + 4\pi\boldsymbol{M}$ to obtain

$$\nabla \cdot \boldsymbol{H}_D = -4\pi \nabla \cdot \boldsymbol{M} \tag{4.70}$$

Since there are no "free" currents in this problem, $\nabla \times \boldsymbol{H}_D = 0$. Therefore $\boldsymbol{H}_D$ may be expressed in terms of a magnetic scalar potential, φ, as $\boldsymbol{H}_D = -\nabla\varphi$. Therefore (4.70) becomes

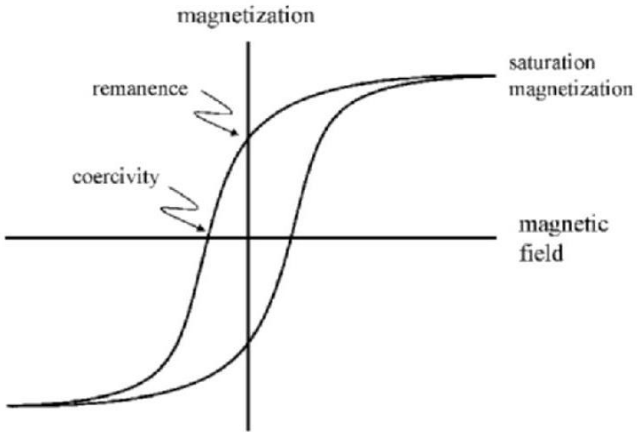


Fig. 4.18. Typical "hysteresis loop" of a ferromagnet

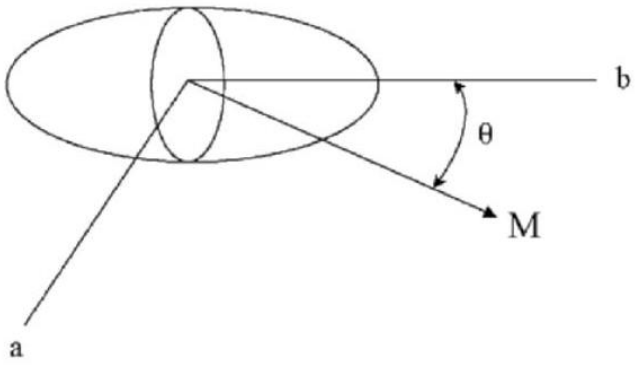


Fig. 4.19. Coordinate system used to describe an ellipsoidal particle

$$\nabla^2 \varphi = 4\pi \nabla \cdot \boldsymbol{M}. \tag{4.71}$$

This has the solution

$$H_D(\boldsymbol{r}) = -\int \frac{\nabla' \cdot \boldsymbol{M}'(\boldsymbol{r}-\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|^3} d^3\boldsymbol{r}' \tag{4.72}$$

If the magnetization is assumed to remain uniform, $\nabla \cdot \boldsymbol{M}$ is nonzero only at the surface of the particle. This "magnetic charge" produces a "demagnetization" field inside given by

$$\boldsymbol{H}_D = -\boldsymbol{N} \cdot \boldsymbol{M}, \tag{4.73}$$

where

$$\underset{\sim}{N} = \begin{pmatrix} N_a & 0 & 0 \\ 0 & N_b & 0 \\ 0 & 0 & N_c \end{pmatrix} \tag{4.74}$$

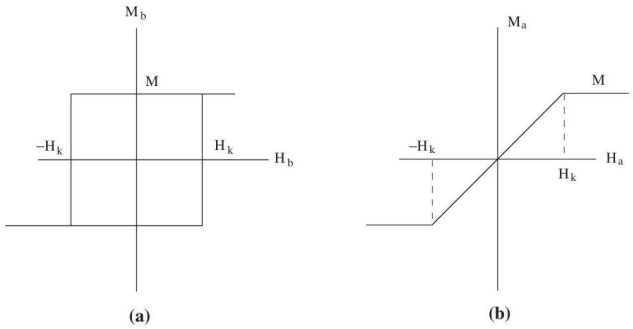


Fig. 4.20. Hysteresis loop for fields parallel (a) and perpendicular (b) to the axis of an ellipsoidal particle

It can be shown that the demagnetization factors N_i satisfy

$$\sum_{i=1}^{3} N_i = 4\pi. \tag{4.75}$$

The fact that the demagnetization field is uniform inside the particle is due to its ellipsoidal shape. The demagnetization field inside a uniformly magnetized bar, for example, is not uniform.

Since the particle has axial symmetry we may assume that the magnetization and the applied field lie in the $a-b$-plane. If the magnetization makes an angle θ with the b-axis then the total energy density of the system is

$$\begin{aligned} \mathcal{E} &= -\boldsymbol{M}\cdot\boldsymbol{H}_0 - \frac{1}{2}\boldsymbol{M}\cdot\boldsymbol{H}_D \\ &= -MH_a\sin\theta - MH_b\cos\theta + \frac{1}{2}\left(N_a - N_b\right)M^2\sin^2\theta + \frac{1}{2}N_bM^2 \end{aligned} \tag{4.76}$$

In the absence of the applied field the direction of $\boldsymbol{M}$ depends upon the demagnetization factors. Let us assume $N_a > N_b$. The magnetization will then point along the b-direction. Suppose we now apply a field along the b-direction opposite to $\boldsymbol{M}$. The energy density is then given by (ignoring the constant term)

$$\mathcal{E} = M\left|H_b\right|\cos\theta + \frac{1}{2}\left(N_a - N_b\right)M^2\sin^2\theta. \tag{4.77}$$

For small H_b this energy density has minima at $\theta = 0$ and $\theta = \pi$. When H_b exceeds $H_K \equiv 2K_u/M$, where $K_u \equiv \frac{1}{2}\left(N_a - N_b\right)M^2$, the minimum at $\theta = 0$ vanishes leaving only the minimum at $\theta = \pi$. At this point $\boldsymbol{M}$ switches to the negative b-direction (Fig. 4.20a).

If the field is applied along the a-direction then the energy density becomes

$$\mathcal{E} = -MH_a\sin\theta + \frac{1}{2}\left(N_a - N_b\right)M^2\sin^2\theta \tag{4.78}$$

In this case the minimum shifts continuously from $\theta = 0$ to $\theta = \pi/2$ and the magnetization rotates smoothly into the a-direction (Fig. 4.20b).

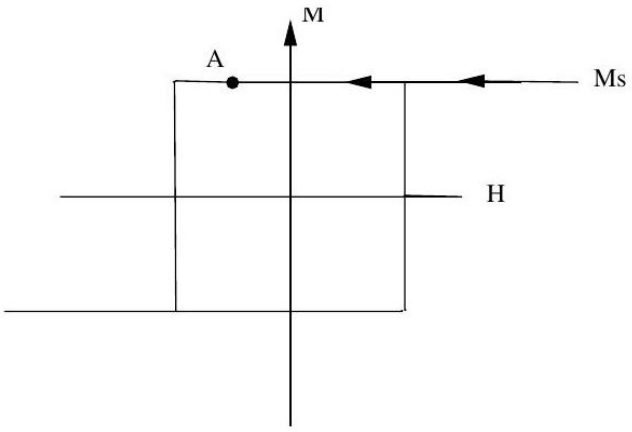
If we now assume that we have an array of identical particles, but that they are randomly oriented with respect to the applied field, and if we neglect the dipolar fields among the particles, then Stoner and Wohlfarth [99] showed that the magnetization of the array has a form similar to Fig. 4.18 with a coercivity $H_c = H_K/2$ and a remanence $M_r = M/2$. In this case the coercivity is governed by the shape anisotropy of the particles. If the particles also possess crystalline

anisotropy arising from the spin-orbit interaction described in Sect.2.3.1, this energy must be added to (4.74). Cobalt, for example, has an electronic structure that generally produces a large crystalling anisotropy. Consequently, one often finds cobalt as the main constituent in high coercivity materials.

Coercitivities can range from a fraction of an oersted, such as in permalloy, an alloy of iron and nickel, to tens of thousands of oersteds, as in transition metal-rare earth alloys. Materials having low coercivities are referred to as magnetically "soft", while those with large coercivities are "hard".

10.5.7 4.7 Dynamic Coercivity

Let us now consider the consequences of incorporating temperature in the Stoner-Wohlfarth model. Suppose we start in the fully saturated state and suddenly reverse the applied field to the point A on the hysteresis loop shown below. The sum of the anisotropy energy and the Zeeman energy as a function



of the angle the magnetization makes with the easy axis is shown in Fig. 4.21. Until H reaches $-H_K$ there is still a local energy minimum at $\theta = 0$. The energy difference between this minimum and the maximum between $\theta = 0$ and $\theta = \pi$ is

$$\Delta E = K_u V \left(1 - H/H_K\right)^2 \tag{4.79}$$

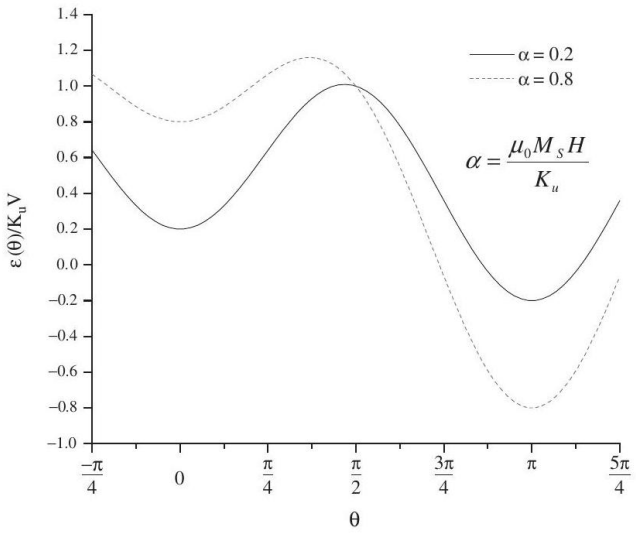


Fig. 4.21. Energy of a magnetic particle as a function of the direction that an applied field makes with the axis of the particle

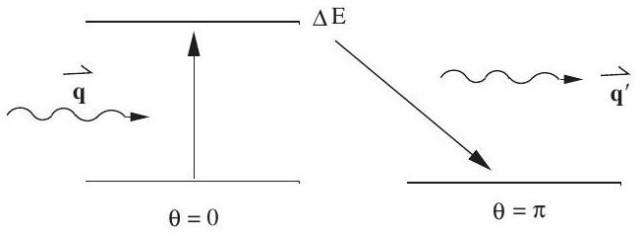


Fig. 4.22. Phonon-assisted magnetization reversal

where V is the "switching volume" which might be a magnetic particle or a grain in a film or just a semi-isolated spin cluster. If temperature is included then there is a probability that this volume may be excited over the barrier into the local minimum at $\theta = \pi$. Physically this reversal occurs through coupling between the spins of the magnetic system and the thermal vibrations of the lattice. As the lattice vibrates it modulates the exchange (exchangestriction) and the magnetostatic energy (magnetostriction). To lowest order these coupling energies are proportional to the amplitude of the lattice displacement, Δx. The displacement Δx may be expanded in terms of quantized lattice vibration, or phonon, amplitudes $b_q, b_q^\dagger$. The coupling energy then has the form $gS_\mu S_\nu \left(b_q + b_q^\dagger\right)$ where g characterizes the strength of the coupling. Second order perturbation theory gives the rate at which a phonon is absorbed by the magnetic system, exciting it to an energy ΔE, and then reemitting a phonon and dropping into the lower minimum as illustrated in Fig. 4.22,

$$\frac{1}{\tau} = \frac{2\pi}{\hbar^2} \sum_{\boldsymbol{q},\boldsymbol{q}'} |g|^2 n_{\boldsymbol{q}} \left(n_{\boldsymbol{q}} + 1\right) \delta \left(\Delta E - \hbar\omega_q\right), \tag{4.80}$$

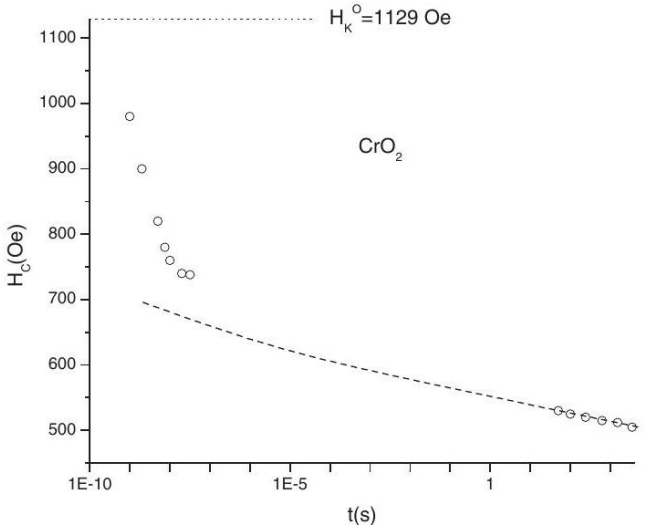


Fig. 4.23. Dynamic coercivity as a function of the time taken to do the measurement. The very short times are measurements taken with a pulsed field. The cluster of longer times are magnetometer measurements [101]

where

$$n_q = \left\langle b_q^\dagger b_q \right\rangle = \frac{1}{e^{\hbar\omega q/k_B T} - 1}. \tag{4.81}$$

Since $\hbar\omega_q = \Delta E$ which is, in general, larger than $k_B T$,

$$\frac{1}{\tau} = f_0 e^{-\Delta E/k_B T}. \tag{4.82}$$

f_0 is often referred to as an "attempt" frequency. In this model it is proportional to the spin-phonon coupling. This process corresponds to "thermally assisted" reversal of the magnetization, or superparamagnetism.

Equation (4.82) implies that the probability that a cluster of spins reverses by hopping over the barrier is $1 - e^{-t/\tau}$. Then the magnetization will reverse according to

$$M(t) = M_s \left(2e^{-t/\tau} - 1\right). \tag{4.83}$$

The magnetization will pass through zero at a time $t_0 = -\left(\ln \frac{1}{2}\right)\tau$. If we define the field at which this occurs as the coercivity $H_c(t_0)$, then

$$H_c(t_o) = H_K \left[1 - \frac{k_B T}{K_u V} \ln\left(\frac{f_0 t_0}{0.693}\right)\right], \tag{4.84}$$

where t_0 is identified as the "measurement time". Thus, (4.85) says that the coercivity depends upon the measurement time. Equation (4.85) is called Sharrock's law after M. Sharrock who applied it to particulate media [100]. The coercivity is plotted as a function of the measurement time in Fig. 4.23. The dashed line is (4.85). We see that this does not describe the data at very short times. We shall see what is happening at short times in Chap. 6.

10.5.8 4.8 Magnetic Viscosity

Since the barrier discussed in the previous section is proportional to the volume of the spin cluster, we expect, in general, to have a distribution of barrier heights. Let us define $N(E)$ as the density of spin clusters with barrier height E. As a result of thermal activation the rate of change of the number of clusters, $N(E)dE$, with barriers between E and $E + dE$, is given by

$$\frac{d}{dt} N(E)dE = -f_0 e^{-E/k_B T} N(E)dE \tag{4.85}$$

This has the solution

$$N(E) = N_0(E) e^{-t f_0 \exp(-E/k_B T)}, \tag{4.86}$$

where $N_0(E)$ is the initial density distribution. Suppose, for example, that the initial distribution has a square shape with the value N_0 up to $E_{\max}$, i.e.,

$$N_0(E) = N_0, \quad 0 \le E \le E_{\max}.$$

As time evolves after application of a reversing field those clusters with small barriers will reverse, depleting the low energy end of the distribution as shown in Fig. 4.24. The reverse magnetization is given by the area of this "hole". Since this is an exponential process the curve will sweep out equal areas in equal logarithmic increases in time. That is, the two shaded areas in the figure are the same. This means that the magnetization decays logarithmically with time,

$$M(t) = M(0) - S \ln t \tag{4.87}$$

The coefficient S is known as the viscosity parameter [102].
Since the thermal fluctuations lead to a change in the magnetization characterized by the viscosity S, one may describe this by an effective fluctuation field H_f which is related to S by an "irreversible" susceptibility,

$$S = \chi_{\rm irr} H_f.$$

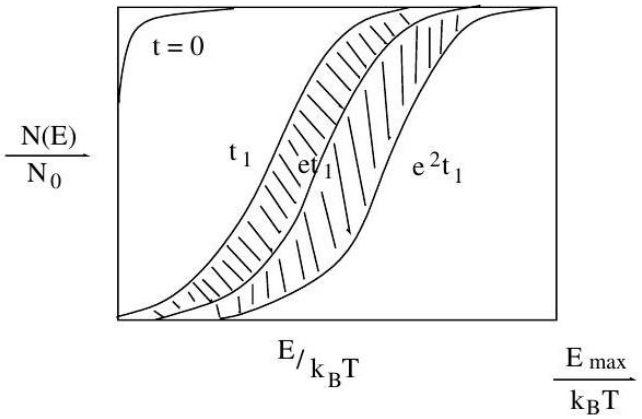


Fig. 4.24. Distribution of barrier heights for different times after a reversal field has been applied

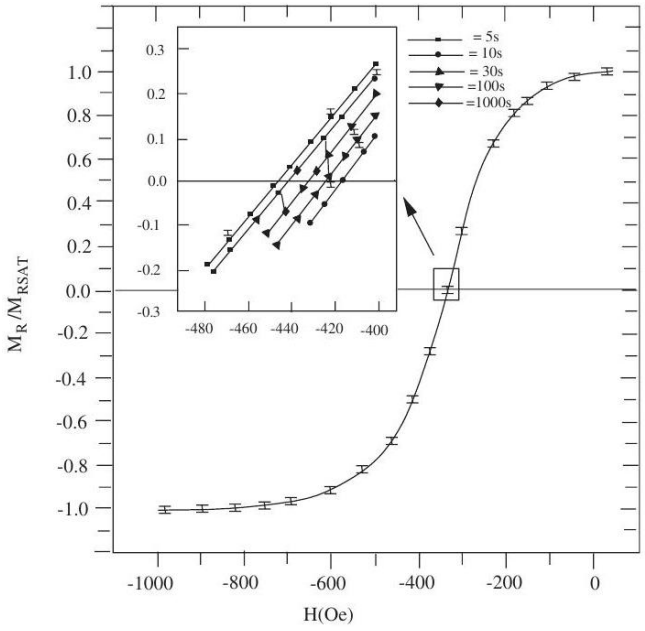


Fig. 4.25. Remanence curves for $\mathrm{CrO
_{2}$ with data for different values of pulse length shown on an expanded scale [103]}

This susceptibility may be obtained from the reverse remanence curve. That is, we initially saturate the sample and then apply a reverse field for some time τ. At the end of this time the remanent moment is measured enabling us to construct a remanence curve as shown in Fig. 4.25. The irreversible susceptibility $\chi_{\rm irr}$ is defined as the maximum slope, which occurs at the remanent coercivity.

Problems

4.1. Consider an impurity spin S' which couples to its nearest neighbors through an exchange interaction J'. In the molecular field approximationg this impurity represents an effective field

$H^z_{\text{eff}} = 2\left(J'S' - JS\right)/g\mu_B$. Using the concept of the generalized susceptibility calculate the moment induced around the impurity as the temperature approaches T_c from above.
4.2. Add an anisotropy energy of the form $-D\sum_i \left(S_i^z\right)^2$ to a ferromagnet and
(a) determine how this affects T_c;
(b) calculate $\chi_\perp(0)$.
4.3. Entropy is given by the temperature derivative of the free energy, $S = -\partial F/\partial T$.
(a) How does the entropy vary below the transition according to Landau Theory.
(b) What does Landau Theory predict for the specific heat at T_c.
4.4. Differentiate (4.63) with respect to H and use the definition

$$\partial G(\epsilon, H)/\partial H = -M$$

to derive a relation between x, y, and β, where β is defined by $M(\epsilon, 0) \sim (-\epsilon)^\beta$. Similarly, derive a relation between x, y, and δ, where δ is defined by $M(0, H) \sim H^{1/\delta}$. Now differentiate twice with respect to H and use the definition of the isothermal susceptibility, $\chi_T = -\left(\partial^2 G/\partial H^2\right)_T$ and, $\chi_T(\epsilon, 0) \sim (-\epsilon)^{-\gamma'}$ to derive a relation between γ', β, and δ. This is know as the Widom equality.
4.5. Compute $\chi_\perp(0)$ for the Stoner-Wohlfarth model and thereby obtain a relationship between the microscopic anisotropy constant, D in Problem 4.2, and the macroscopic anisotropy constant, K_1.

10.6 The Static Susceptibility of Interacting Systems: Metals

The long-range magnetic order in metals is very similar to that observed in insulators as illustrated by the magnetization curve of nickel shown in Fig. 5.1. However, the electrons participating in this magnetic state are itinerant as determined by the existence of a Fermi surface; that is, they also have translational degrees of freedom. How such a system of interacting electrons responds to a magnetic field is a many-body problem with all its attendant difficulties.

The many-body corrections to the Landau susceptibility and the Pauli susceptibility must be treated separately. Kanazawa and Matsudaira [104] found that the many-body corrections to the Landau susceptibility are small (less than one per cent) for high electron densities.

We shall approach the effect of electron-electron interactions on the spin susceptibility in two ways. The first, Fermi liquid theory, is a phenomenological approach. It involves parameters completely analogous to the parameters entering the spin Hamiltonian. These parameters may be determined experimentally or they may be obtained from the second approach which is to assume a specific microscopic model from which various physical properties can be calculated.

10.6.1 5.1 Fermi Liquid Theory

The phenomenological theory of an interacting fermion system was developed by Landau in 1956 [106]. Although Landau was mainly interested in the properties of liquid He^3, his theory may also be applied to metals. Modifications of this theory in terms of the introduction of a magnetic field have been made by Silin [107].

Let us begin by considering the ground state of a system of N electrons. For a noninteracting system the ground state corresponds to a well-defined Fermi sphere. Landau assumed that as the interaction between the electrons is gradually "turned on" the new ground state evolves smoothly out of the

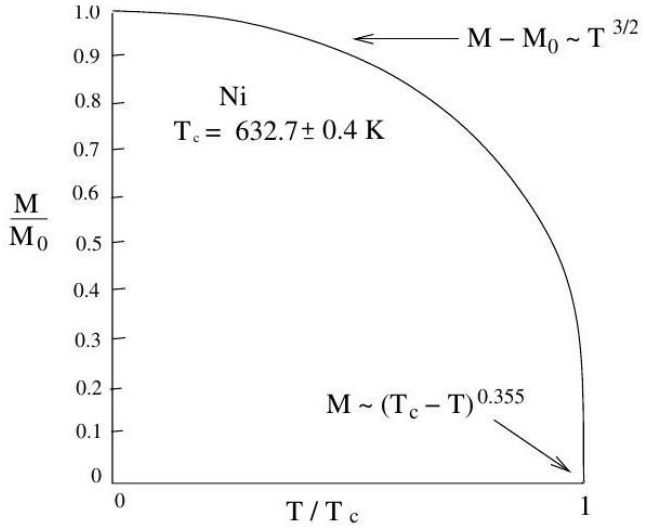


Fig. 5.1. Magnetization of nickel as a function of temperature. The original data of Weiss and Forrer [105] taken at constant pressure has been corrected to constant volume to eliminate the effects of thermal expansion

original Fermi sphere; if $|0\rangle$ is this new ground state, it is related to the original Fermi sphere $|FS\rangle$ by a unitary transformation,

$$|0\rangle = U|FS\rangle. \tag{5.1}$$

Let us denote the energy associated with $|0\rangle$ as E_0.
Landau also applied this assumption to the excitations of the interacting system. For example, suppose we add one electron, with momentum $\hbar\boldsymbol{k}$, to the non-interacting system. This state has the form $a^\dagger_{\boldsymbol{k}\sigma}|FS\rangle$, where $a^\dagger_{\boldsymbol{k}\sigma}$ is the creation operator for an electron. If the interactions are gradually turned on, let us approximate the new state as

$$|\boldsymbol{k}\sigma\rangle = U a^\dagger_{\boldsymbol{k}\sigma}|FS\rangle. \tag{5.2}$$

Because the electron possesses spin, this wave function is a spinor.
Let us define the difference between the energy of $|\boldsymbol{k}\sigma\rangle$ and $|0\rangle$ as $\epsilon^0(\boldsymbol{k},\boldsymbol{\sigma})$. Since the wave function is a spinor, this energy will be 2×2 matrix. If the system is isotropic, and in particular if there is no external magnetic field, then this energy is independent of the spin,

$$\epsilon^0(\boldsymbol{k},\boldsymbol{\sigma})_{\alpha\beta} = \epsilon^0(\boldsymbol{k})\delta_{\alpha\beta}.$$

Because the whole Fermi sphere has readjusted itself as a result of the interactions, the energy $\epsilon^0(\boldsymbol{k})$ will be quite different from the energy of a free particle. As we do not know this energy, we shall assume that $\boldsymbol{k}$ is close to k_F and expand in powers of $k - k_F$. Thus we obtain

$$\epsilon^0(\boldsymbol{k}) = \mu + \frac{\hbar^2 k_F}{m^*}(k - k_F) + \ldots, \tag{5.3}$$

where

$$\frac{\hbar^2 k_F}{m^*} \equiv \left.\frac{\partial\epsilon^0(\boldsymbol{k})}{\partial k}\right|_{k=k_F}. \tag{5.4}$$

This electron, "dressed" by all the other electrons, is called a quasiparticle. Notice that the energy required to create a quasiparticle at the Fermi surface is μ, the chemical potential.

Its increase in energy as it moves away from the Fermi surface is characterized by its effective mass m^*. We restrict ourselves to the region close to the Fermi surface because it is only in this region that quasiparticle lifetimes are long enough to make their description meaningful.

We could just as well have removed an electron from some point within the Fermi sphere. This would have created a "hole", which the interactions would convert into a quasi-hole. The energy associated with a hole is the energy required to remove an electron at the Fermi surface, $-\mu$, plus the energy it takes to move the electron at $\boldsymbol{k}$ up to the surface, $(\hbar^2 k_F/m^*)\,(k_F - k)$. However, if we define the total energy of the system containing a quasi-hole as $E_0 - \epsilon^0(\boldsymbol{k})$, then

$$\epsilon^0(\boldsymbol{k}) = \mu + \frac{\hbar^2 k_F}{m^*}\,|k - k_F|\,.$$

Thus the excitation spectrum associated with our Fermi liquid has the form shown in Fig. 5.2.

Suppose that other quasiparticles are now introduced into the system. This could occur, for example, as a result of an external field producing electronhole pairs. Since the energy of a quasiparticle depends on the distribution of all the other quasiparticles, any change in distribution will lead to a change in the quasiparticle energy. Let us denote the change in the distribution by $\delta n(\boldsymbol{k}, \boldsymbol{\sigma})$.

The quasiparticle distribution is essentially the density matrix associated with the quasiparticle. In particular, it is a 2×2 matrix. For example, $\delta n(\boldsymbol{k}, \boldsymbol{\sigma})_{11}$ gives the probability of finding an electron of momentum $\hbar\boldsymbol{k}$ with

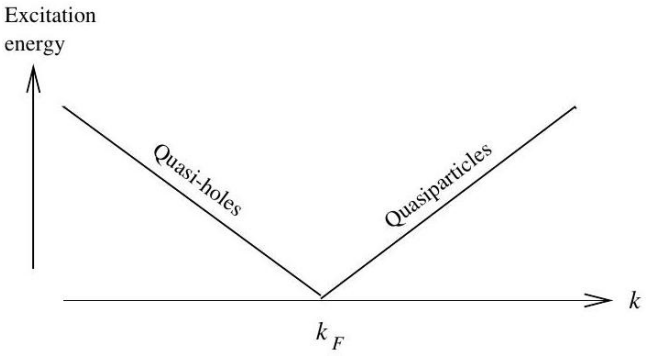


Fig. 5.2. Single-particle excitation spectrum of a Fermi liquid

spin up. Therefore the quasiparticle energy may be written in phenomenological terms as

$$\epsilon(\boldsymbol{k}, \boldsymbol{\sigma}) = \epsilon^0(\boldsymbol{k}, \boldsymbol{\sigma}) + \frac{1}{V}\,\mathrm{Tr}_{\sigma'}\left\{\sum_{\boldsymbol{k}'} f\,(\boldsymbol{k}, \boldsymbol{\sigma}; \boldsymbol{k}', \boldsymbol{\sigma}')\,\delta n\,(\boldsymbol{k}', \boldsymbol{\sigma}')\right\}. \tag{5.5}$$

The quantity $f\,(\boldsymbol{k}, \boldsymbol{\sigma}; \boldsymbol{k}', \boldsymbol{\sigma}')$ is a product of 2×2 matrices analogous to a dyadic vector product. Again, if the system is isotropic, the most general form this quantity can have is

$$f\,(\boldsymbol{k}, \boldsymbol{\sigma}; \boldsymbol{k}', \boldsymbol{\sigma}') = \varphi\,(\boldsymbol{k}, \boldsymbol{k}')\,\mathbf{1}\mathbf{1}' + \psi\,(\boldsymbol{k}, \boldsymbol{k}')\,\boldsymbol{\sigma} \cdot \boldsymbol{\sigma}', \tag{5.6}$$

where $\mathbf{1}$ is the 2×2 unit matrix. Furthermore, since this theory is valid only near the Fermi surface, we may take $|\boldsymbol{k}'| \simeq |\boldsymbol{k}| = k_F$. Then φ and ψ depend only upon the angle θ between $\boldsymbol{k}'$ and $\boldsymbol{k}$, and we may expand φ and ψ in Legendre polynomials:

$$\varphi\left(\boldsymbol{k},\boldsymbol{k}'\right)=\frac{\pi^2\hbar^2}{m^*k_F}A\left(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{k}}'\right)=\frac{\pi^2\hbar^2}{m^*k_F}\left[A_0+A_1P_1(\cos\theta)+\ldots\right] \tag{5.7}$$

$$\psi\left(\boldsymbol{k},\boldsymbol{k}'\right)=\frac{\pi^2\hbar^2}{m^*k_F}B\left(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{k}}'\right)=\frac{\pi^2\hbar^2}{m^*k_F}\left[B_0+B_1P_1(\cos\theta)+\ldots\right] \tag{5.8}$$

If we know the quasiparticle distribution function, then we can compute, just as for the electron gas, all the relevant physical quantities. These will involve the parameters A_n and B_n. The beauty of this theory is that some of the same parameters enter different physical quantities. Therefore by measuring certain quantities we can predict others. The difficulty, of course, is in determining the distribution function $\delta n(\boldsymbol{k},\boldsymbol{\sigma})$. For static situations this is relatively easy. However, for dynamic situations, as we shall see in the following chapters, we have to solve a Boltzmann-like equation.

Since the $\boldsymbol{k}$ dependence of the quasiparticle energy is a result of interactions, there should be a relation between m^* and the parameters A_n and B_n. To obtain this relation let us consider the situation at $T=0$ in which we have one quasiparticle at $\boldsymbol{k}$ with spin up, as illustrated in Fig. 5.3a. Now, suppose the momentum of this system is increased by $\hbar\boldsymbol{q}$, giving us the situation in

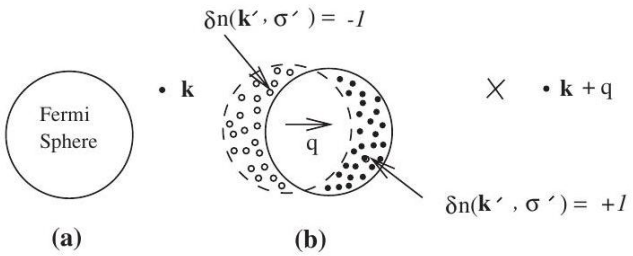


Fig. 5.3. Effect of a uniform translation in momentum space on a state containing one extra particle

Fig. 5.3b. This corresponds to placing the whole system on a train moving with a velocity $\hbar\boldsymbol{q}/m$. To an observer at rest with respect to the train it will appear that the quasiparticle has acquired an additional energy

$$\delta\epsilon(\boldsymbol{k},\boldsymbol{\sigma})_{11}=\hbar^2\boldsymbol{k}\cdot\boldsymbol{q}/m \tag{5.9}$$

for small $\boldsymbol{q}$. However, the quasiparticle itself experiences a change in energy associated with its own motion in momentum space, which, from (5.3), is just $\hbar^2\boldsymbol{k}\cdot\boldsymbol{q}/m^*$. In addition, it sees the redistribution of quasiparticles indicated in Fig. 5.3b. Since this momentum displacement does not produce any spin flipping, $\delta n(\boldsymbol{k},\boldsymbol{\sigma})_{\alpha\beta}$ will have the form $\delta n(\boldsymbol{k})\delta_{\alpha\beta}$, where $\delta n(\boldsymbol{k})$ is $+1$ for the quasiparticles and -1 for the quasi-holes. This gives a contribution of

$$\frac{2}{V}\sum_{\boldsymbol{k}'}\varphi\left(\boldsymbol{k},\boldsymbol{k}'\right)\delta n\left(\boldsymbol{k}'\right)$$

to the 1,1 component of the energy, where the factor 2 arises from the spin trace. Equating these two changes in energy, which is what is meant by Galilean invariance, and converting the sum over $\boldsymbol{k}'$ to an integral leads to our desired relation,

$$m^*=m\left(1+\frac{A_1}{3}\right). \tag{5.10}$$

It can be shown that the specific heat of a Fermi liquid has the same form as that for an ideal Fermi gas, with m replaced by m^*. Thus by measuring the specific heat we can determine the Fermi liquid parameter A_1.

Exchange Enhancement of the Pauli Susceptibility. We are now ready to consider our original question of the response of a Fermi liquid to a magnetic field. In the presence of a magnetic field the noninteracting quasiparticle energy $\epsilon^0(\boldsymbol{k}, \boldsymbol{\sigma})$ is no longer independent of the spin, but contains a Zeeman contribution,

$$\epsilon^0(\boldsymbol{k}, \boldsymbol{\sigma}) = \epsilon^0(\boldsymbol{k})\mathbf{1} + \mu_B H \sigma_z. \tag{5.11}$$

We shall assume that any field-induced contributions to the interaction term are small. Therefore the total quasiparticle energy is

$$\epsilon(\boldsymbol{k}, \boldsymbol{\sigma}) = \epsilon^0(\boldsymbol{k})\mathbf{1} + \mu_B H \sigma_z + \frac{1}{V} \operatorname{Tr}_{\sigma'} \left\{ \sum_{\boldsymbol{k}'} f\left(\boldsymbol{k}, \boldsymbol{\sigma}; \boldsymbol{k}', \boldsymbol{\sigma}'\right) \delta n\left(\boldsymbol{k}' \boldsymbol{\sigma}'\right) \right\}. \tag{5.12}$$

It is energetically more favorable for the quasiparticles to align themselves opposite to the field, since their gyromagnetic ratio is negative. However, each time a quasiparticle flips over it changes the distribution, thereby bringing in contributions from the last term in (5.12). Thus, if we start with two equal spin

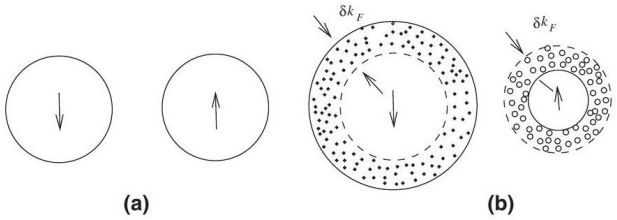


Fig. 5.4. Effect of a dc magnetic field on the spin-up and the spin-down Fermi spheres

distributions, as shown in Fig. 5.4a, an equilibrium situation will eventually be reached, as illustrated in Fig. 5.4b, in which the energy of a quasiparticle on the up-spin Fermi surface is equal to that of a quasiparticle on the down-spin surface; that is,

$$\epsilon\left(k_F + \delta k_F, \boldsymbol{\sigma}\right)_{22} = \epsilon\left(k_F - \delta k_F, \boldsymbol{\sigma}\right)_{11}. \tag{5.13}$$

From (5.3), (5.12) this condition becomes

$$\begin{array}{l} \frac{\hbar^2 k_F}{m^*} \delta k_F - \mu_B H + \frac{1}{V} \operatorname{Tr}_{\sigma'} \left\{ \sum_{\boldsymbol{k}'} \left[\varphi\left(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{k}}'\right) \mathbf{1}' - \psi\left(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{k}}'\right) \sigma_z' \right] \delta n\left(\boldsymbol{k}', \boldsymbol{\sigma}'\right) \right\} \\ = -\frac{\hbar^2 k_F}{m^*} \delta k_F + \mu_B H + \frac{1}{V} \operatorname{Tr}_{\sigma'} \left\{ \sum_{\boldsymbol{k}'} \left[\varphi\left(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{k}}'\right) \mathbf{1}' + \psi\left(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{k}}'\right) \sigma_z' \right] \delta n\left(\boldsymbol{k}', \boldsymbol{\sigma}'\right) \right\}. \end{array} \tag{5.14}$$

The change in quasiparticle distribution shown in Fig. 5.4b is characterized by

$$\delta n(\boldsymbol{k}, \boldsymbol{\sigma}) = \begin{cases} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} & k_F < |\boldsymbol{k}| < k_F + \delta k_F \\ \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} & k_F - \delta k_F < |\boldsymbol{k}| < k_F. \end{cases} \tag{5.15}$$

Therefore

$$\mathrm{Tr}_{\sigma'}\left\{\sum_{\boldsymbol{k}'}\varphi\left(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{k}}'\right)\mathbf{1}\delta n\left(\boldsymbol{k}',\sigma'\right)\right\}=0, \tag{5.16}$$

while

$$\mathrm{Tr}_{\sigma'}\left\{\sum_{k'}\psi\left(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{k}}'\right)\sigma_z'\delta n\left(\boldsymbol{k}',\boldsymbol{\sigma}\right)\right\}=-\frac{4\pi V}{(2\pi)^3}k_F^2\delta k_F\int_{-1}^{+1}d(\cos\theta)\psi\left(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{k}}'\right). \tag{5.17}$$

Equation (5.14) then reduces to

$$\frac{2\hbar^2k_F}{m^*}\delta k_F-2\mu_BH+\frac{2\hbar^2k_F\delta k_F}{m^*}B_0=0. \tag{5.18}$$

Since the magnetization is

$$M_z=-\frac{\mu_B}{V}\,\mathrm{Tr}_\sigma\left\{\sum_{\boldsymbol{k}}\sigma_z\delta n(\boldsymbol{k},\boldsymbol{\sigma})\right\}=2\mu_B\frac{\pi k_F^2\delta k_F}{(2\pi)^3} \tag{5.19}$$

the uniform susceptibility of a Fermi liquid at $T=0$ is

$$\chi(0)=\frac{1+\frac{1}{3}A_1}{1+B_0}\chi_{\mathrm{Pauli}} \tag{5.20}$$

Thus we find that in addition to the appearance of the effective mass in place of the bare mass, the susceptibility is also modified by the factor $(1+B_0)^{-1}$. In the Hartree-Fock approximation

$$B_0=-\frac{me^2}{\pi\hbar^2k_F}=-0.166r_s \tag{5.21}$$

and we speak of the susceptibility as being exchange enhanced. As the electron density decreases and $r_s\to 6.03$ the susceptibility diverges. This is usually taken to imply that such a material will be ferrogmagnetic. There has been a great deal of discussion [108] about the magnetic state of an interacting electron system, and it is generally agreed that such a system will not become ferromagnetic at any electron density. That is, the Hartree-Fock approximation favors ferromagnetism. The reason is that in this approximation parallel spins are kept apart by the exclusion principle while antiparallel spins are spatially uncorrelated. Thus the antiparallel spins have a relatively large Coulomb energy to gain by becoming parallel. In an exact treatment one would expect the antiparallel spins to be somewhat correlated, thereby reducing the Coulomb difference. The differences between the exact properties of an interacting electron system and those obtained in the Hartree-Fock approximation are referred to as correlation effects. Estimates of these correlation corrections indicate that the nonmagnetic ground state of the electron gas has a lower energy than the ferromagnetic one.

The predictions of Fermi-liquid theory, namely that the low temperature heat capcity varies as γT and that the resistivity varies as T^2 are found to describe most metals. During the last decade, however, non-Fermi-liquid behavior has been observed in a number of systems. One of the most studied is the high-temperature superconductor, $La_{z-x}Sr_xCuO_4$. The phase diagram for this system is shown in Fig. 4.16. The "normal" region above the superconducting region shows anamolous features. Anderson, for example, argues that the absence of a residual resistivity in the *ab* plane invalidates a Fermiliquid description. It is known that a Fermi-liquid approach certainly fails in one dimension, for in one dimension the Fermi "surface" consists of

only two points at $k = \pm k_F$. Any interaction with momentum transfer $q = 2k_F$ leads to an instability that produces a gap in the energy spectrum. In 1963

Luttinger introduced a model of 1D interacting Fermions. His solution does not give quasi-particles, but rather spin and charge excitations that propagate independently. Whether the 2D CuO_2 planes in $La_{2-x}Sr_xCuO_4$ can be described by such a "Luttinger liquid" is a subject of debate. Other materials exhibiting non-Fermi-liquid behavior are the so-called heavy Fermions.

10.6.2 5.2 Heavy Fermion Systems

At low temperatures the electrons in a "normal" metal contribute a term to the specific heat that is linear in the temperature, i.e., $C = \gamma T$ as we mentioned above. Simple theoretical considerations give $\gamma = \frac{2}{3}\pi^2 k_B^2 N(E_F)$. In particular, for a free electron metal the density of states is given by $N(E_F) = (2m/\hbar^2)^{3/2} E_F^{1/2}$ which gives a value of γ of the order of one mJ/K^{-2} mol^{-1}. There exist, however, metallic systems with low temperature heat capacity coefficients of the order of 1000 mJ/K^{-2} mol^{-1}. Examples include $CeCu_2Si_2$, UBe_{13}, and UPt_3. Almost all the examples involve rare earths, such as Ce , or actinides, such as U . These elements have their f^n and $f^{n\pm1}$ electronic configurations close enough in energy to allow valence fluctuations with hybridization. This can lead to Kondo behavior (see Sect.3.4.3) and is why some refer to heavy Fermion systems as "Kondo lattices". In this description the large electronic mass is associated with a large density of states at the Fermi level that derives from the many-body resonance we found in our treatment of the Kondo impurity. When d-electron ions are used to create a Kondo lattice their large spatial extent results in too strong a hybridization to show heavy Fermion behavior. LiV_2O_4 appears to be an exception.

Once one has a heavy Fermion system it is subject to the same sorts of Fermi surface instabilities found in more normal metals. In particular, $CeCu_2Si_2$ shows both a spin density wave and superconductivity as one changes the 4f-conduction electron coupling by substituting Ge for the Si.

The Kondo lattice is not the only mechanism that may lead to heavy Fermions. When Nd_2CuO_4 is doped with electrons by introducing Ce for Nd , i.e., $Nd_{2-x}Ce_xCuO_4$, the linear specific heat coefficient is $\gamma = 4$ J/K^{-2} mol^{-1}. The conduction occurs through hopping among the Cu sites. As a result of the double exchange we described in Sect.2.2.10, these hopping electrons see an effective antiferromagnetic field. The conduction electron Hamiltonian is therefore taken to be

$$-t\sum_{i,j,\sigma}\left(a^\dagger_{i\sigma}a_{j\sigma} + h.c.\right) + h\sum_{i\sigma}\sigma e^{i\boldsymbol{Q}\cdot\boldsymbol{R}_i}a^\dagger_{i\sigma}a_{i\sigma}$$

where $\boldsymbol{Q}$ is a reciprocal lattice vector ($\pi/a, \pi/a$) and h is the effective field that accounts for the antiferromagnetic correlations. The 4 f electrons are described by the term

$$\epsilon_f \sum_{i\sigma} f^\dagger_{i\sigma} f_{i\sigma}$$

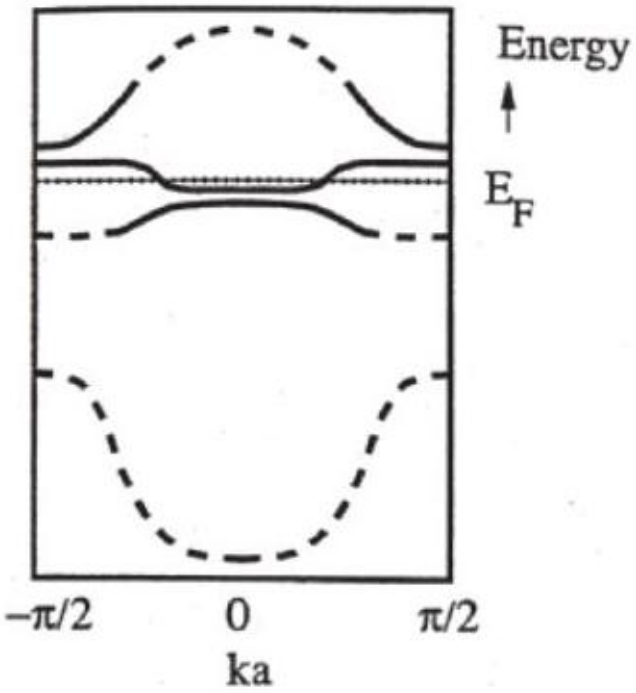


Fig. 5.5. Schematic plot of the quasiparticle bands of $\mathrm{Nd
{z-x} C e{x} \mathrm{CuO}_{4}$ for $x \neq 0$. The Fermi energy is indicated by a dotted line. Solid lines: f-like excitations, and dashed lines: d-like excitations [109]}

Finally, we add a hybridization

$$V \sum_{i\sigma} \left(a_{i\sigma}^{\dagger} f_{i\sigma} + \text{ h.c. } \right).$$

This is a very simplified model since it only considers one 4 f orbital instead of seven. Nevertheless, when this Hamiltonian is diagonalized [109] one obtains the four bands shown in Fig. 5.5. The Fermi level in the doped case sits in the narrow f-band giving rise to a large electronic mass.

10.6.3 5.3 Itinerant Magnetism

The appearance of ferromagnetism in real metals is related to the presence of the ionic cores which tend to localize the intinerant electrons and introduce structure in the electronic density of states. We shall now consider two models that incorporate these features.

5.3.1 The Stoner Model

In the 1930s both Slater [110] and Stoner [111] combined Fermi statistics with the molecular field concept to explain itinerant ferromagnetism. This one-electron approach is now generally referred to as the Stoner model. It bears similarities to Landau's Fermi liquid theory in that the effect of the electron-electron interactions is to produce a spin-dependent potential that simply shifts the original Bloch states.

Stoner's result is contained within the generalized susceptibility $\chi(\boldsymbol{q})$ of an interacting electron system. As a model Hamiltonian we take a form similar to (1.139) where ϵ_k now refers to the Bloch band energy. Since the Coulomb interaction in a metal is screened, let us take a delta-function interaction of

the form $I\delta(\boldsymbol{r}_i - \boldsymbol{r}_j)$. In this case one need not add a compensating background charge density and (1.139) becomes

$$\mathcal{H}_0 = \sum_{\boldsymbol{k}\sigma} \epsilon_k a^\dagger_{k\sigma} a_{k\sigma} + \frac{I}{V} \sum_{\boldsymbol{k}} \sum_{\boldsymbol{k}} \sum_{\boldsymbol{k}'} \sum_{\boldsymbol{q}} \sum_{\sigma} a^\dagger_{k-\boldsymbol{q},\sigma} a^\dagger_{\boldsymbol{k}'+\boldsymbol{q},-\sigma} a_{\boldsymbol{k},\sigma} a_{\boldsymbol{k}'\sigma'} \tag{5.22}$$

If we now add a spatially varying field $H\hat{\boldsymbol{z}}\cos(\boldsymbol{q}\cdot\boldsymbol{r})$, the Zeeman Hamiltonian becomes

$$\mathcal{H}_Z = -\frac{H}{2}\left[\mathcal{M}_z(\boldsymbol{q}) + \mathcal{M}_z(-\boldsymbol{q})\right] \tag{5.23}$$

where

$$\mathcal{M}_z(\boldsymbol{q}) = \frac{1}{2} g\mu_B \sum_{\boldsymbol{k}} \left(a^\dagger_{\boldsymbol{k}-\boldsymbol{q},\uparrow} a_{k\uparrow} - a^\dagger_{\boldsymbol{k}-\boldsymbol{q},\downarrow} a_{k\downarrow} \right) \tag{5.24}$$

Susceptibility. The susceptibility is obtained by calculating the average value of $\mathcal{M}_z(\boldsymbol{q})$ to lowest order in H . In particular, we must calculate the average of

$$m_{\boldsymbol{k},\boldsymbol{q}} \equiv a^\dagger_{k-\boldsymbol{q},\uparrow} a_{k,\uparrow} - a^\dagger_{k-\boldsymbol{q},\downarrow} a_{k,\downarrow}$$

Following Wolff [112] we shall do this by writing the equation of motion for $m_{\boldsymbol{k},\boldsymbol{q}}$ and using the fact that in equilibrium $\partial \langle m_{\boldsymbol{k},\boldsymbol{q}} \rangle \partial t = 0$. Thus,

$$\langle [m_{\boldsymbol{k},\boldsymbol{q}}, \mathcal{H}_0 + \mathcal{H}_Z] \rangle = 0 \tag{5.25}$$

This commutator involves a variety of twofold and fourfold products of electron operators. These are simplified by making a random-phase approximation in which we retain only those pairs which are diagonal or have the forms appearing in $m_{\boldsymbol{k},\boldsymbol{q}}$ itself. Furthermore, the diagonal pairs are replaced by their average in the noninteracting ground state. This is equivalent to a HartreeFock approximation. The result is

$$(\epsilon_{\boldsymbol{k}} - \epsilon_{\boldsymbol{k}+\boldsymbol{q}}) \langle m_{\boldsymbol{k},\boldsymbol{q}} \rangle - \frac{2I}{V} (n_{\boldsymbol{k}+\boldsymbol{q}} - n_{\boldsymbol{k}}) \sum_{\boldsymbol{k}'} \langle m_{\boldsymbol{k}',\boldsymbol{q}} \rangle - g\mu_B H (n_{\boldsymbol{k}+\boldsymbol{q}} - n_{\boldsymbol{k}}) = 0 \tag{5.26}$$

Dividing by $(\epsilon_{\boldsymbol{k}+\boldsymbol{q}} - \epsilon_{\boldsymbol{k}})$ and summing over $\boldsymbol{k}$ gives

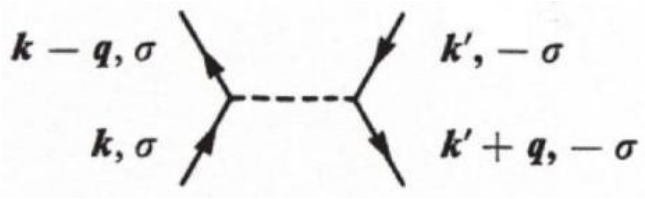
$$\sum_k \langle m_{\boldsymbol{k},\boldsymbol{q}} \rangle = \frac{g\mu_B H \sum_{\boldsymbol{k}} \frac{n_{\boldsymbol{k}} - n_{\boldsymbol{k}+\boldsymbol{q}}}{\epsilon_{\boldsymbol{k}+\boldsymbol{q}} - \epsilon_{\boldsymbol{k}}}}{1 - \frac{2I}{V} \sum_{\boldsymbol{k}} \frac{n_{\boldsymbol{k}} - n_{\boldsymbol{k}+\boldsymbol{q}}}{\epsilon_{\boldsymbol{k}+\boldsymbol{q}} - \epsilon_{\boldsymbol{k}}}} \tag{5.27}$$

Since the equilibrium occupation number $n_{\boldsymbol{k}}$ is just the Fermi function f_k, we recognize the sum appearing in this expression as the same as that which appears in the noninteracting susceptibility (3.93), which we shall denote as $\chi_0(\boldsymbol{q})$. Thus,

$$\chi(\boldsymbol{q}) = \frac{\chi_0(\boldsymbol{q})}{1 - \frac{2I}{g^2\mu_B^2}\chi_0(\boldsymbol{q})}. \tag{5.28}$$

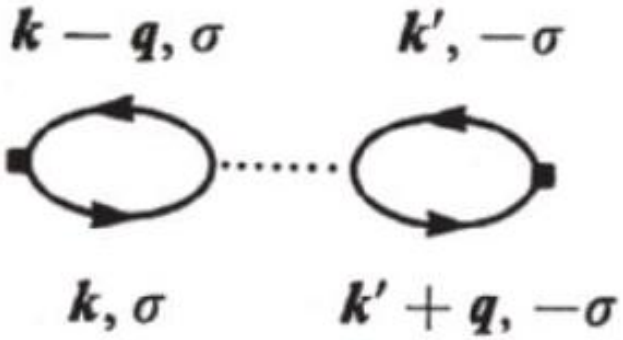
We could also have calculated $\langle m_{\boldsymbol{k},\boldsymbol{q}} \rangle$ directly by the formulation indicated in (1.81). However, the appearance of the interaction term in the exponentials which enter $m_{\boldsymbol{k},\boldsymbol{q}}(t)$ requires many-body perturbation techniques which are beyond the scope of this monograph. We shall use the diagrams introduced in Chap. 1 to make the results of such a treatment plausible. The

interaction term in (5.22) has the diagramatic form

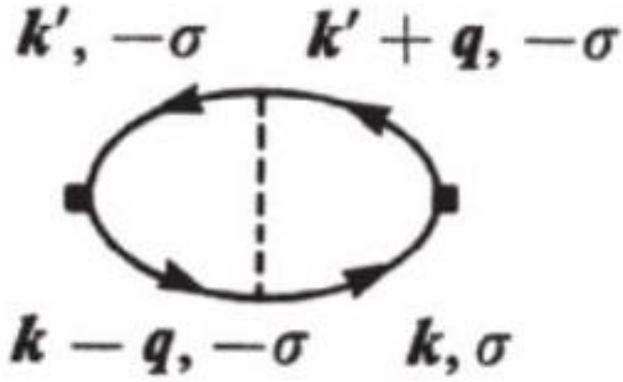


In calculating average values such as the energy or the magnetization the electron and hole lines must be closed.

There are two ways one can close the electron and hole lines:



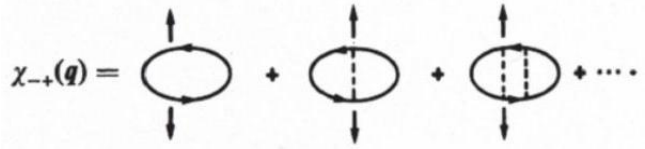
(a)



(b)

In (a) the vertices indicated by the small squares involve a momentum change $\boldsymbol{q}$ but no spin flip. This longitudinal spin fluctuation represents the first-order correction to χ_{zz}. The susceptibility (5.28) corresponds to the summation of all such diagrams:

In diagram (b) the vertices also involve a spin flip. This diagram is the firstorder correction to the transverse susceptibility $\chi_{-+}(\boldsymbol{q})$ whose noninteracting form is given by (3.97). Summing all such transverse spin fluctuations gives the transverse susceptibility.



In the paramagnetic state rotational invariance requires $\chi_{-+}(\boldsymbol{q}) = 2\chi_{zz}(\boldsymbol{q})$. Returning to (5.28), since $\lim_{\boldsymbol{q}\to 0}\chi_0(\boldsymbol{q}) = \frac{1}{3}g^2\mu_B^2\mathcal{D}(\epsilon_F)$, we see that

$$\chi(0) = \frac{\chi_0(0)}{1 - I\mathcal{D}(\epsilon_F)} \tag{5.29}$$

This has the same form as our Fermi liquid result (5.20). Thus we have an expression, at least in the random-phase approximation, for the Fermi liquid parameter B_0 in terms of the intraatomic exchange integral I and the density of states.

Notice that the criterion for the appearance of ferromagnetism is that $I\mathcal{D}(\epsilon_F) \geq 1$. This is referred to as the Stoner criterion.

One might wonder how ferromagnetism can occur with only an intraatomic Coulomb interaction. To see the physical origin of this, suppose we have a spin up at some site α. If the spin on a neighboring site α' is also up, it is forbidden by the exclusion principle from hopping onto site α. Therefore the two electrons do not interact, and we might define the energy of such a configuration as 0 . However, if the spin of site α' is down, it has a nonzero probability of hopping onto site α. Thus the energy of this configuration is higher than that of the "ferromagnetic" one. However, hopping around can lower the kinetic energy of the electrons. This is reflected in the appearance of the density of states in the Stoner criterion. The occurrence of ferromagnetism depends, therefore, on the relative values of the Coulomb interaction and the kinetic energy.

Equation (5.29) has interesting consequences for metals which are paramagnetic but have a large enhancement factor. For example, an impurity spin placed in such a host produces a large polarization of the conduction electrons in its vicinity. Such giant moments have been observed in palladium and certain of its alloys. This seems reasonable, since palladium falls just below nickel in the periodic table and has similar electronic properties. Thus we might suspect that although palladium is not ferromagnetic like nickel, it at least possesses a large exchange enhancement.

Spin-Density Waves. Just as in the case of localized moments, a divergence of $\chi(\boldsymbol{q})$ for $\boldsymbol{q} \neq 0$ would imply a transition to a state of nonuniform magnetization. In fact, Overhauser [113] has shown that the $\chi(\boldsymbol{q})$ associated with an unscreened Coulomb interaction in the Hartree-Fock approximation does diverge as $q \to 2k_F$, as shown in Fig. 5.6. This would lead to a ground state characterized by a period spin density, called a spin-density wave. The effects of screening and electron correlations, however, tend to suppress this divergence. Consequently, a spin-density wave can form only under rather special conditions. We get some feeling for these conditions by considering the behavior of the non-interacting susceptibility $\chi_0(\boldsymbol{q})$ in one and two dimensions

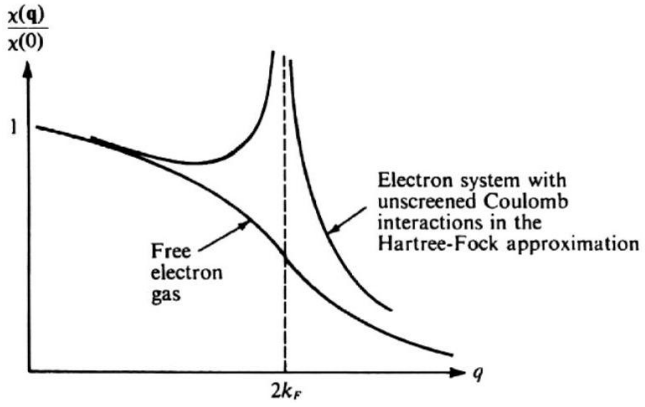


Fig. 5.6. Effect of electron-electron interactions on the susceptibility (3 dimensions)

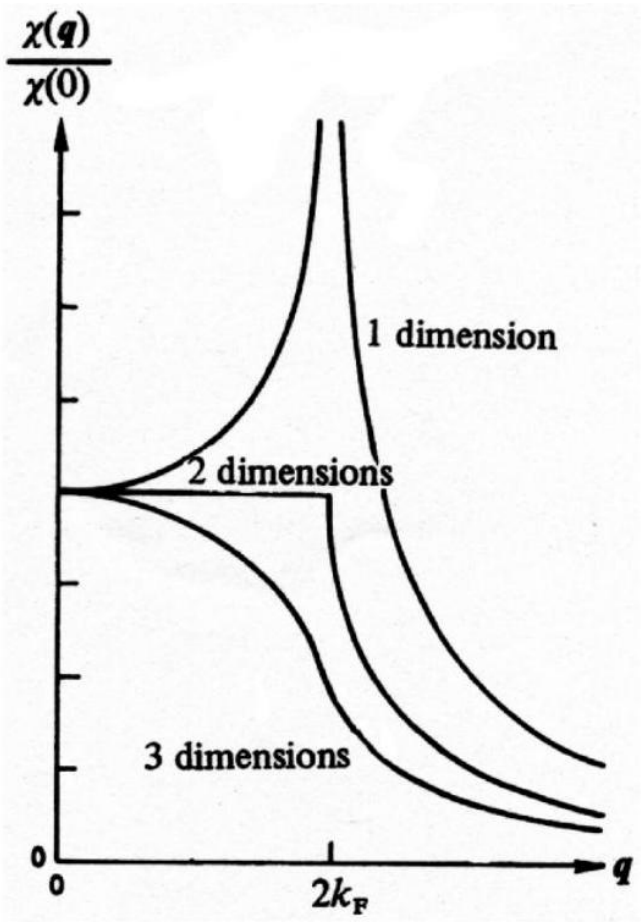


Fig. 5.7. Effect of dimensionality on the free electron susceptibility

as shown in Fig. 5.7. We see that lower dimensional systems are more likely to become unstable with respect to spin-density wave formation.

The reason for this has to do with the fact that in a state characterized by a wave vector $\boldsymbol{q}$ electrons with wave vectors which differ by $\boldsymbol{q}$ become correlated. This effectively removes them from the Fermi sea. This is often described as a "nesting" of the corresponding states. In one and two dimensions Fermi surfaces are geometrically simpler, which means that nesting will have more dramatic effects. These same considerations, however, also apply
to charge-density instabilities. And, in fact, most of the materials which do show such Fermi-surfaced-related instabilities show charge-density waves.

To date, chromium and its alloys are the best examples of materials possessing a spin-density-wave ground state. The reason for this has to do with the band structure of chromium.

If the susceptibility does not actually diverge at some nonzero wave vector, but nevertheless becomes very large, the system may be said to exhibit antiferromagnetic exchange enhancement. Experiments on dilute alloys of Sc : Gd indicate that scandium metal may be an example of such a type [114].

Exchange Splitting. If the system is ferromagnetic, then $\langle m_{\boldsymbol{k},\boldsymbol{q}=0}\rangle = (n_{\boldsymbol{k}\uparrow} - n_{\boldsymbol{k}\downarrow}) \neq 0$ even in the absence of an applied field. In this case the Hamiltonian may be written in a particularly revealing form by considering only the diagonal terms in (5.22):

$$\frac{1}{2}n_{k'\downarrow}a^{\dagger}_{k\uparrow}a_{k\uparrow} + \frac{1}{2}n_{k'\uparrow}a^{\dagger}_{k\downarrow}a_{k\downarrow} \tag{5.30}$$

Since

$$M = \frac{g\mu_B}{2V} \sum_k \left(n_{k\uparrow} - n_{k\downarrow}\right), \tag{5.31}$$

the Hamiltonian becomes

$$\mathcal{H}_0' = E_0 + \sum_{k,\sigma} \epsilon_{k\sigma} a_{k\sigma}^\dagger a_{k\sigma} \tag{5.32}$$

where

$$\epsilon_{k\sigma} = \epsilon_k + \frac{NI}{4V} - \frac{IM}{2g\mu_B}\sigma. \tag{5.33}$$

Thus the spin-up and spin-down energy bands are split by an amount proportional to the magnetization. This was the basic idea in Stoner's original theory of ferromagnetism. More generally, this splitting arises from the difference between the spin-up and spin-down exchange-correlation potentials seen by the electrons. Since the Stoner model was first proposed there has been a great deal of progress in specifying these potentials.

In 1951 Slater [115] suggested that we approximate the effect of exchange by the potential

$$V_x(\boldsymbol{r}) = -6[(3/8\pi)\rho(\boldsymbol{r})]^{1/3}$$

This approximation form has been used extensively both in atomic and solidstate calculations. Physically, this density to the one-third power arises from the fact that in the Hartree-Fock approximation parallel spins are kept farther apart than antiparallel spins. Therefore, there is an "exchange hole" around any particular spin associated with a deficiency of similar spins. The radius of this hole must be such that $\left(\frac{3}{4}\right)\pi r^3\rho = 1$. Since the potential associated with this deficiency is proportional to $1/r$, we obtain an exchange potential proportional to $\rho^{1/3}$.

In 1965 Kohn and Sham [116] rederived the exchange potential by a different method and obtained a value two thirds that of Slater's. This led workers to multiply $V_x(\boldsymbol{r})$ by an adjustable constant, α, which can be determined for each atom by requiring that the so-called $X\alpha$ energy be equal to the HartreeFock energy for that atom. An interesting application of the $X\alpha$ method has been made by Hattox et al. [117]. They calculated the magnetic moment of bcc vanadium as a function of lattice spacing. The result is shown in Fig. 5.8. We see that the moment falls suddenly to zero at a spacing 20% larger than the actual observed spacing. This decrease is due to the broadening of the 3d band as the lattice spacing decreases. This is consistent with the fact that bcc vanadium, where the vanadium-vanadium distance is $\sqrt{3}a_0/2 = 2.49$Å, is observed to be nonmagnetic. In Au_4 V, however, the vanadium-vanadium distance has increased to 3.78Å and the vanadium has a moment near one Bohr magneton.

If one uses this approach to find α for the magnetic transition metals one finds that the resulting magnetic moments are not in agreement with those observed.

This is not surprising when we consider that we have replaced a nonlocal potential (the Hartree-Fock potential) by a local potential (the Slater $\rho^{1/3}$). Furthermore, the fact that these are different means that the Slater potential must, by definition, include some correlation. However, we do not know if it is in the right direction.

The arbitrariness inherent in the $X\alpha$ method may be avoided by using the "spin-density functional" formalism. This essentially enables one to utilize the results of many-body calculations for the homogeneous electron gas in determining the exchange and correlation potentials in transition metals. This

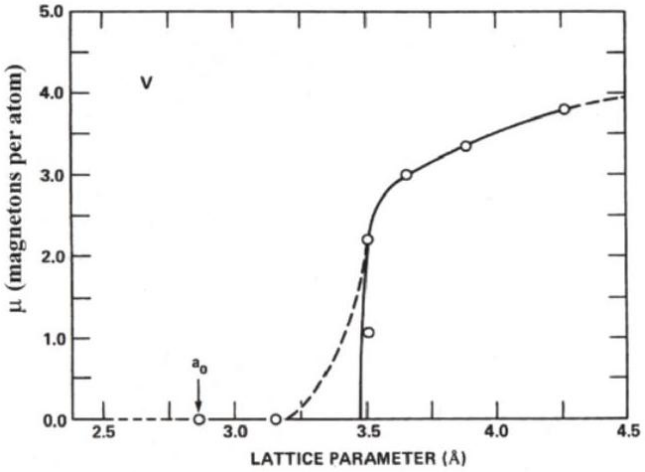


Fig. 5.8. Calculated magnetic moment of vanadium metal as a function of lattice parameter. The two points for $a = 3.5$Å correspond to two distinct self-consistent solutions associated with different starting potentials. These two solutions are the result of a double minimum in the total energy versus magnetization curve. At 4.25Å and 3.15Å the calculations converged to unique values [117]

approach is based on a theorem by Hohenberg and Kohn [118] which states that the ground-state energy of an inhomogeneous electron gas is a functional of the electron density $\rho(\boldsymbol{r})$ and the spin density $m(\boldsymbol{r})$. The effective potential, for example, is given by

$$V_{\text{eff}} = v(\boldsymbol{r}) + e^2 \int \frac{\rho\,(\boldsymbol{r}')\,d\boldsymbol{r}'}{|\boldsymbol{r} - \boldsymbol{r}'|} + \frac{\delta E_{xc}\{\rho\}}{\delta\rho(\boldsymbol{r})} \tag{5.34}$$

where $v(\boldsymbol{r})$ is the one-electron potential, the second term is the Hartree term, and $E_{xc}\{\rho\}$ is the exchange and correlation energy. The importance of this theorem is that it reduces the many-body problem to a set of one-body problems for the one-electron wave functions $\phi_{k\sigma}(\boldsymbol{r})$ which make up the density,

$$\rho(\boldsymbol{r}) = \sum_{\boldsymbol{k},\sigma} \phi^*_{\boldsymbol{k}\sigma}(\boldsymbol{r})\phi_{\boldsymbol{k}\sigma}(\boldsymbol{r}) \tag{5.35}$$

The local density approximation consists of replacing the unknown functional $E_{xc}\{\rho\}$ by $\int d^3r\rho(\boldsymbol{r})\epsilon^h_{xc}[\rho(\boldsymbol{r})]$ where $\epsilon^h_{xc}[\rho(\boldsymbol{r})]$ is the exchange and correlation contribution to the energy of a homogeneous interacting electron gas of density $\rho(\boldsymbol{r})$. Although one might question the use of such an approximation in transition metals where there are rapid variations in the charge density, the fact that it is only the spherical average of the exchange-correlation hole which enters the calculation makes $E_{xc}\{\rho\}$ fairly insensitive to the details of this hole.

The density functional approach leads to a set of self-consistent Hartree equations. It is tempting to identify the eigenvalues as effective single particle energies. However, detailed studies of the energy bands indicate that this identification may not be appropriate in certain cases. Photoemission has become a powerful technique for probing the electronic states of solids. In this technique, light (actually ultraviolet or X-rays) is absorbed by a solid. Those electrons excited above the vacuum level leave the solid and are collected. Initially, all the electrons emitted were collected and analyzed. However, with the availability of high-intensity synchrotron X-ray sources, it became possible to resolve the direction of the emitted electrons. This enables one to reconstruct the energymomentum relation of the electrons in the solid. Figure 5.9a compares the results of such an experiment on copper [119] with the calculated band structure. The agreement is remarkable. This technique has now been extended to determine

the spin polarization of the photoemitted electrons. This extension was pioneered by H.C. Siegmann in Zürich. The photoemitted electrons are accelerated to relativistic velocities and then scattered from a gold foil. Any initial spin polarization is reflected in asymmetric, or Mott, scattering. Figure 5.9b shows the results of such measurements on Ni [120]. Although the overall features are described by the theory, the detailed fit is not nearly as good as in Cu .

Despite this problem with the identification of single-particle energies, one can calculate ground-state properties such as the bulk modulus and the magnetic moment. The results are in very good agreement with the experimental values. What makes this agreement all the more impressive is that the only input to such calculations are the atomic numbers and the crystal structures.

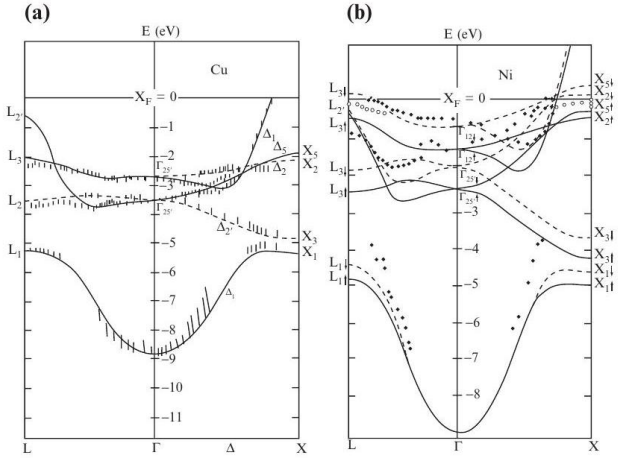


Fig. 5.9. Comparison of angle-resolved photoemission data with calculated band structures for (a) copper and (b) nickel (open data points correspond to spin up, solid points to spin down). The triangles are de Haas van Alphen data

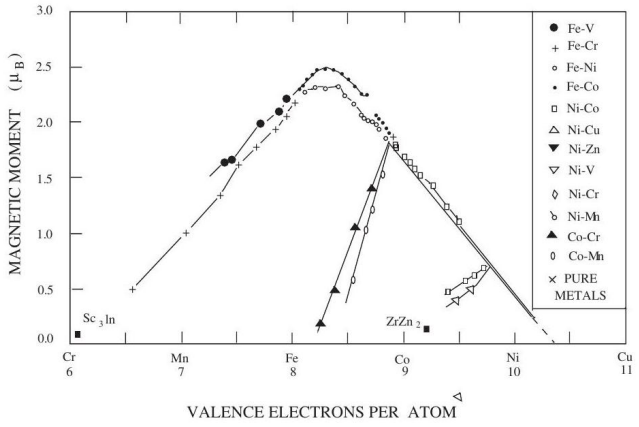


Fig. 5.10. Saturation magnetization as a function of electron concentration

Alloys. A band description also works well for alloys. Fig. 5.10 shows the magnetic moments for various transition metal alloys. This curve is known as the Slater-Pauling curve. Slater [121] noted that the magnetic properties of 3d solid solutions, particularly their moments, could be averaged over the periodic table and plotted as a function of the filling of the *d*-band. If the pairs of atoms have too much charge contrast, then the electronic states tend to become localized and the simple averaging fails as indicated by the lines branching off the main trend.

The Slater-Pauling curve may be understood by considering
the d-electron bands to be exchange split as shown for Ni in Fig. 5.9b. In the "rigid-band" approximation, one assumes that the only effect of alloying is to shift the Fermi level. The majority spin band in nickel is completely full (in recognition of which it is sometimes referred to as "strong" meaning further splitting cannot increase the magnetism). The maximum moment between cobalt and iron in the Slater-Pauling curve reflects the composition where the majority band is just completely full. Further removal of electrons as one moves across the compositional axis towards iron results in depletion of both majority and minority bands. Pauling [122] offered an alternative explanation based on the idea that 2.56d-orbitals hybridize with s - and p-orbitals to form nonmagnetic bonding orbitals. The remaining 2.44d-orbitals fill according to Hund's rule to give the magnetic moments.

There is no doubt that the magnetic electrons partake in the transport properties of the transition metal ferromagnets, i.e., that they are itinerant. Nevertheless, there are some physical properties where a localized model is a reasonable and tractable approximation. The situation that pertains is, of course, somewhere between completely localized and free. Using neutron diffraction, which is discussed in Chap. 10, Mook [123] has apportioned the moment, in μ_B, in nickel as follows:

3 d spin $(n_\uparrow - n_\downarrow) =$	$+0.656$
3 d orbital	$= +0.055$
4 s polarization $=$	-0.105
	$+0.606$

The moment density is quite asymmetric about the lattice sites. About 80% of the 3d magnetic electrons occupy t_{2g} orbitals. For hexagonal cobalt Moon [124] finds:

3 d spin $(n_\uparrow - n_\downarrow) =$	$+1.86$
3 d orbital $=$	$+0.13$
4 s polarization $=$	-0.28
	$+1.71$

In this case, the magnetic moment looks like an almost spherical distribution of positive moment localized around each atomic site decreasing to a negative level in the region between atoms.

Although the Stoner theory works reasonably well for magnetic properties at $T = 0$, it fails when applied to finite temperature properties. For example, the only place that temperature enters the susceptibility (5.28) is in the arguments of the Fermi functions. The Curie temperature is calculated from the relation

$$M = \frac{g\mu_B}{2V} \sum_k (f_{k\uparrow} - f_{k\downarrow}) \tag{5.36}$$

where $f_{k\sigma}$ is the Fermi function with $\epsilon_{k\sigma} = \epsilon_k + NI/4V - (IM/2g\mu_B)\,\sigma$. Since at $T_c M$ is very small, the Fermi functions may be expanded about $M = 0$. The condition for T_c then becomes

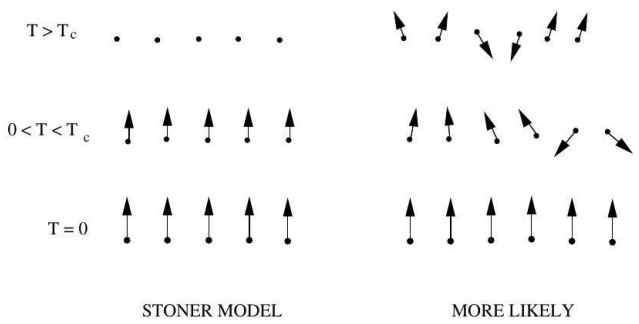


Fig. 5.11. Pictorial comparison of the exchange fields seen by an electron in the Stoner model and what is more likely the case

$$I \int d\epsilon \frac{\partial f\left(T_c\right)}{\partial \epsilon} \mathcal{D}(\epsilon) + 1 = 0. \tag{5.37}$$

Using the values of I that give the correct moments at $T = 0$ for Fe, Co, and Ni, we obtain T_c 's from (5.37) that are about 5 times larger than the observed values.

The problem with this application of the Stoner model is that the introduction of "up" and "down" spin directions destroys the rotational symmetry. That is, at nonzero temperatures the direction of the effective field arising from the electron-electron interactions as well as its magnitude will vary from site to site as illustrated in Fig. 5.11. Independent calculations by Hubbard [125] and Heine and collaborators [126] show that the energy associated with such local changes in the direction of the magnetization is much less than that associated with changes in the magnitude of the magnetization. That is, the exchange stiffness which characterizes the directional variation is less than the Stoner parameter I. The problem remains, however, to relate this observation to the thermodynamic properties.

5.3.2 The Hubbard Model

The one-electron Stoner model described above is expected to apply to systems with fairly broad bands. As the bands become narrower intraionic correlation effects become more important. In this case it is convenient to work in the Wannier representation which emphasizes the atomic aspect of the problem. In this representation the one-electron terms becomes

$$\sum_{\alpha\alpha'\sigma} t_{\alpha\alpha'} a^{\dagger}_{\alpha'\sigma} a_{\alpha\sigma}, \tag{5.38}$$

where

$$t_{\alpha\alpha'} = \frac{1}{V} \sum_{\boldsymbol{k}} \epsilon_{\boldsymbol{k}} \exp\left[i\boldsymbol{k} \cdot \left(\boldsymbol{R}_{\alpha} - \boldsymbol{R}_{\alpha'}\right)\right] \tag{5.39}$$

describes the hopping of an electron from site α to site α'. The interaction terms are given by (2.89). Since we are now dealing with an itinerant situation, we shall assume that screening effects restrict the interaction to one site. If there is a single nondegenerate orbital $\varphi_0\left(\boldsymbol{r} - \boldsymbol{R}_{\alpha}\right)$ associated with each site, then the interaction becomes

$$\sum_{\alpha,\sigma} U_0 n_{\alpha\sigma} n_{\alpha,-\sigma} \tag{5.40}$$

where

$$U_0 = \langle 00|V|00\rangle = \iint d\boldsymbol{r} d\boldsymbol{r}' \varphi_0^*(\boldsymbol{r}-\boldsymbol{R}_\alpha)\,\varphi_0^*(\boldsymbol{r}'-\boldsymbol{R}_\alpha)\,V(\boldsymbol{r}-\boldsymbol{r}') \times \varphi_0(\boldsymbol{r}'-\boldsymbol{R}_\alpha)\,\varphi_0(\boldsymbol{r}-\boldsymbol{R}_\alpha) \tag{5.41}$$

Because of the screening, which we can think of as $s-d$ correlation effects, the value of U_0 is of the order of several eV .

Equations (5.38), (5.40) constitute the Hubbard Hamiltonian. It contains the same physics as the Stoner model. In fact, it gives the same susceptibility (5.28) in the mean-field approximation with U_0 in place of I. In a series of papers Hubbard [127] investigated the effects of correlation within this model. He found, for example, that for large correlation the electronic band is split into two subbands separated by U_0. The transition-metal oxides such as NiO and CoO are generally cited as examples of materials where such correlation effects are responsible for their insulating properties.

In order to understand the variation in magnetic properties as one moves across the transition-metal series, it is necessary to generalize the model above to include orbital degeneracy. This obviously introduces many more Coulomb and exchange integrals. The first simplification is to neglect interactions involving more than two orbitals. One then assumes that all the off-diagonal Coulomb and exchange integrals are the same, i.e.,

$$\left.\begin{array}{l}\langle mm'|\,V\,|mm'\rangle = U \\ \langle mm'|\,V\,|m'm\rangle = J\end{array}\right\} m' \neq m \tag{5.42}$$

The exchange integral J is smaller than U. We also take all the diagonal integrals to have the same value. If we require that our model Hamiltonian preserve the rotational invariance of the original Hamiltonian, then the diagonal integral is related to the off-diagonal integrals by

$$\langle mm|V|mm\rangle = U + J. \tag{5.43}$$

The resulting generalized Hubbard Hamiltonian becomes

$$\mathcal{H}' = \frac{1}{2}(U+J)\sum_{\alpha,m,\sigma} n_{\alpha m\sigma} n_{\alpha m,-\sigma} + \frac{1}{2}\sum_{\alpha,m,m'\sigma}(1-\delta_{mm'}) \times \left[U n_{\alpha m\sigma} n_{\alpha m',-\sigma} + (U-J) n_{\alpha m\sigma} n_{\alpha m'\sigma} - J a^\dagger_{\alpha m\sigma} a_{\alpha m-\sigma} a^\dagger_{\alpha m'-\sigma} a_{\alpha m'\sigma}\right] \tag{5.44}$$

By making use of the spin representation (2.85)-(2.87) the last term may be written

$$\left(U - \frac{1}{2}J\right)\sum_{\alpha,m<m',\sigma,\sigma'} n_{\alpha m\sigma} n_{\alpha m'\sigma'} - 2J \sum_{\alpha,m<m'} \boldsymbol{S}_{\alpha m}\cdot\boldsymbol{S}_{\alpha m'}, \tag{5.45}$$

which clearly reveals the Hund's rule exchange coupling. Let us now consider how these exchange interactions modify the generalized susceptibility. As in the derivation (5.28) we shall employ the Hartree-Fock approximation. This amounts to writing

$$n_{\alpha m\sigma} = \langle n_{\alpha m\sigma}\rangle + (n_{\alpha m\sigma} - \langle n_{\alpha m\sigma}\rangle) \tag{5.46}$$

and assuming the term in parentheses is small. Thus,

$$n_{\alpha m\sigma} n_{\alpha m'\sigma'} = \langle n_{\alpha m\sigma}\rangle\, n_{\alpha m'\sigma'} + \langle n_{\alpha m'\sigma'}\rangle\, n_{\alpha m\sigma} - \langle n_{\alpha m\sigma}\rangle\,\langle n_{\alpha m'\sigma'}\rangle\,, \tag{5.47}$$

and the Hartree-Fock Hamiltonian becomes

$$\begin{aligned}\mathcal{H}_{HF} = &\sum_{\alpha,\alpha' m,\sigma} t_{\alpha\alpha'} a^\dagger_{\alpha m\sigma} a_{\alpha' m\sigma} + (U+J)\sum_{\alpha,m\sigma} \langle n_{\alpha m\sigma}\rangle\, n_{\alpha m,-\sigma} \\ &+ U \sum_{\alpha,\alpha' m,\sigma} \left(\langle n_{\alpha m\sigma}\rangle\, n_{\alpha m\sigma,-\sigma} + \langle n_{\alpha m',-\sigma}\rangle\, n_{\alpha m\sigma}\right) \\ &+ (U-J) \sum_{\alpha,m<m',\sigma} \left(\langle n_{\alpha m\sigma}\rangle\, n_{\alpha m'\sigma} + \langle n_{\alpha m'\sigma}\rangle\, n_{\alpha m\sigma}\right). \end{aligned} \tag{5.48}$$

The magnetic moment per ion is

$$m_\alpha = -\frac{g\mu_B}{2} \sum_m \left(\langle n_{\alpha m\uparrow}\rangle - \langle n_{\alpha m\downarrow}\rangle\right). \tag{5.49}$$

If the average number of electrons per ion is n, then

$$n = \sum_m \left(\langle n_{\alpha m\uparrow}\rangle + \langle n_{\alpha m\downarrow}\rangle\right). \tag{5.50}$$

If we take the case of a transition metal where $m = 1, \ldots 5$, then

$$\begin{aligned}\langle n_{\alpha m\uparrow}\rangle &= \frac{1}{10}\left(n - \frac{2m_\alpha}{g\mu_B}\right), \\ \langle n_{\alpha m\downarrow}\rangle &= \frac{1}{10}\left(n + \frac{2m_\alpha}{g\mu_B}\right), \end{aligned} \tag{5.51}$$

and the Hartree-Fock Hamiltonian becomes

$$\begin{aligned}\mathcal{H}_{HF} = &\sum_{\alpha,\alpha' m,\sigma} t_{\alpha\alpha'} a^\dagger_{\alpha m\sigma} a_{\alpha' m\sigma} + \frac{1}{10} N(9U - 3J)n^2 \\ &+ \frac{U+5J}{g\mu_B} \sum_{\alpha,m} m_\alpha \left(n_{\alpha m\uparrow} - n_{\alpha m\downarrow}\right) \end{aligned} \tag{5.52}$$

In the presence of an external field $H\cos(\boldsymbol{q}\cdot\boldsymbol{r})\hat{\boldsymbol{z}}$ the Zeeman interaction has the form

$$\mathcal{H}_Z = \frac{g\mu_B H}{2} \sum_{\alpha,m} \left(n_{\alpha m\uparrow} - n_{\alpha m\downarrow}\right) \cos \boldsymbol{q}\cdot\boldsymbol{R}_\alpha. \tag{5.53}$$

If we assume that the induced moments m_α have the same spatial variation as the applied field, i.e.,

$$m_\alpha = m\cos\left(\boldsymbol{q}\cdot\boldsymbol{R}_\alpha\right),$$

then by comparing the last term in (5.52) with (5.53) we see that the effect of the interactions is to give an effective field

$$H(\boldsymbol{r})_{\text{eff}} = \frac{2m}{g\mu_B}(U+5J)\cos(\boldsymbol{q}\cdot\boldsymbol{r}). \tag{5.54}$$

Taking the Fourier transform of the total effective field and using the fact that

$$M(\boldsymbol{q}) = \chi_0(\boldsymbol{q}) H(\boldsymbol{q})_{\text{eff}}, \tag{5.55}$$

where $\chi_0(\boldsymbol{q})$ is the susceptibility of the noninteracting electron system, we find for the susceptibility of this Hubbard model

$$\chi(\boldsymbol{q}) = \frac{\chi_0(\boldsymbol{q})}{1 - \frac{2(U+5J)}{g^2\mu_B^2}\chi_0(\boldsymbol{q})}. \tag{5.56}$$

This has the same form as the Stoner susceptibility with an effective Stoner parameter $I_{\text{eff}} = U + 5J$. The corresponding Stoner criterion becomes $I_{\text{eff}}\,\mathcal{D}(\epsilon_F) > 1$. Thus the presence of intraatomic exchange favors ferromagnetism. However, again, we expect correlation effects to be very important. There have been many calculations of correlation effects within the Hubbard model but their description is beyond the scope of the monograph. The general result of including correlation is to reduce the region of parameter space where ferromagnetism, or antiferromagnetism, is expected.

Problems

5.1. To obtain some familiarity with the Fermi liquid formalism, this problem asks you to determine the particle current associated with a quasiparticle of momentum k. From the definition of the particle current,

$$j_k = v_{\boldsymbol{k}} - \sum_{\boldsymbol{k}'} f_{kk'}\frac{\partial n_k}{\partial \epsilon'_k} v_F.$$

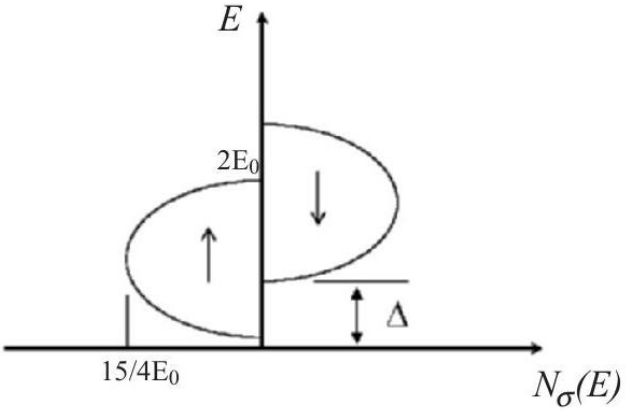
show that

$$j_k = v_k + \left(\frac{m^*}{m} - 1\right) v_F$$

The last term on the right is a "drag" term associated with the flow of the other quasiparticles.

5.2. Evaluate (5.28) for 1 -dimension at $T = 0K$.

5.3. Considering the exchange split bands having the form



$$N_\uparrow(\varepsilon) = \frac{15}{4\varepsilon_0}\left[1 - \left(\frac{\varepsilon - \varepsilon_0}{\varepsilon_0}\right)^2\right]$$

Compute the magnetic moment as a function of ε_F for $\varepsilon_0 < \varepsilon_F < 4\varepsilon_0$ for the two exchange splittings $\Delta = \varepsilon_0$ and $\Delta = 3\varepsilon_0/2$.

5.4. For the band structure shown in Problem 5.3, calculate the total energy as a function of the number of electrons per atom, n. Find the value of Δ which minimizes the total energy and plot $\Delta(n)$.

10.7 The Dynamic Susceptibility of Weakly Interacting Systems: Local Moments

We now turn to the response of magnetic systems to time-dependent excitations. In this chapter we shall restrict our consideration to systems in which the interactions among the constituents are not strong enough to produce a spontaneous magnetization. The study of the frequency response of such systems is essentially the study of paramagnetic resonance and relaxation phenomena. This is obviously an enormous subject, and we shall not be able to go into it in great detail. However, we shall examine some of the basic ideas within the framework of our generalized susceptibility.

10.7.1 6.1 Equation of Motion

Let us begin by considering a system of identical localized spins characterized by a noninteracting Hamiltonian $\mathcal{H}_0$. In Sect. 3.1 we found that in the presence of a uniform static field $H_0\hat{z}$ such a system develops a magnetization, which we shall denote by $M_0\hat{z}$. Let us now apply an additional time-dependent field $\boldsymbol{H}_1 \cos\omega t$ and investigate the response to this field.

Since the applied field is uniform, we shall drop explicit reference to spatial considerations. Thus we may write the magnetization as

$$\boldsymbol{M} = \frac{1}{V}\,\mathrm{Tr}\{\rho\boldsymbol{\mathcal{M}}\}. \tag{6.1}$$

The presence of the volume in this relation and not in (1.49) is due to our definition of the space-dependent operator in (1.48). Because the Zeeman Hamiltonian is time dependent, the density matrix, and hence the magnetization, will be time dependent. Differentiating (6.1) with respect to time and making use of (1.47), which implies $d\rho/dt = 0$, we obtain

$$\frac{d\boldsymbol{M}}{dt} = -\frac{i}{\hbar V}\,\mathrm{Tr}\{\rho[\boldsymbol{\mathcal{M}}, \mathcal{H}]\}. \tag{6.2}$$

For the present let us assume that the Hamiltonian consists of a part which commutes with $\boldsymbol{\mathcal{M}}$ plus a Zeeman part

$$\mathcal{H}_Z = -\mathcal{M}\cdot\boldsymbol{H}. \tag{6.3}$$

Equation (6.2) then becomes

$$\frac{d\boldsymbol{M}}{dt} = -\gamma\boldsymbol{M}\times\boldsymbol{H} \tag{6.4}$$

where $\gamma\hbar = g\mu_B$. The minus sign arises because we are explicitly dealing with electrons which have a negative gyromagnetic ratio.

If the dynamic field $\boldsymbol{H}_1(t) = \boldsymbol{H}_1\cos\omega t$ is applied in the z direction, then, according to (6.4), it exerts no torque on the equilibrium magnetization $M_0\hat{\boldsymbol{z}}$. Therefore the response to such a field is 0 . This raises an interesting point. If the frequency ω goes to 0 , (6.4) tells us that there is no response to such a static field. But in Chap. 3 we found that the magnetization does respond to a static field, with the resulting magnetization given by Curie's law. The answer to this paradox lies in the fact that in Chap. 3 we assumed that the spin system was always in equilibrium. This implies that there is a coupling between the individual spins and their environment which enables them to reach equilibrium. The time it might take the spin system to do this is not important in the static case, since we can always keep the field on until equilibrium has been achieved. In the dynamic case, however, this assumption is not valid, since we may want the response at a frequency that is much faster than this relaxation frequency. In fact, this is generally the experimental situation. In the dynamic case we must

actually solve for the nonequilibrium density matrix. We shall see later that the only way to get a response in the z direction is to introduce a relaxation mechanism (see Problem 6.1).

Let us now consider a system which has come to equilibrium in an applied *dc*-field, H_0, and ask what the response will be to a time-varying transverse field $\boldsymbol{H}_1$. In particular, let $\boldsymbol{H}_1(t) = H_1(t)\hat{\boldsymbol{x}}$. If we write the magnetization as

$$\boldsymbol{M} = m_x\hat{\boldsymbol{x}} + m_y\hat{\boldsymbol{y}} + (M_0 - m_z)\,\hat{\boldsymbol{z}} \tag{6.5}$$

(6.4) becomes, in component form,

$$\frac{dm_x}{dt} = -\gamma H_0 m_y \tag{6.6a}$$
$$\frac{dm_y}{dt} = -\gamma H_1(t)\,(M_0 - m_z) + \gamma H_0 m_x \tag{6.6b}$$
$$\frac{dm_z}{dt} = -\gamma H_1(t) m_y \tag{6.6c}$$

These equations are nonlinear. As a consequence, magnetic systems exhibit a number of interesting features, such as "spin echoes", which we shall discuss
later. For the present, let us linearize (6.6) by neglecting terms which are quadratic in H_1 or the components of $\boldsymbol{m}$. The result is

$$\frac{dm_x}{dt} = -\omega_0 m_y \tag{6.7a}$$
$$\frac{dm_y}{dt} = -\gamma H_1(t) M_0 + \omega_0 m_x \tag{6.7b}$$
$$\frac{dm_z}{dt} = 0 \tag{6.7c}$$

where $\omega_0 \equiv \gamma H_0$.
We see from (6.7c) than m_z is a constant. Since $[\mathcal{M}\cdot\mathcal{M}, \mathcal{M}\cdot\boldsymbol{H}] = 0$, this means that the magnitude of the magnetization is a conserved quantity. Therefore once m_x and m_y are known, m_z may be obtained from the condition that

$$\mathcal{M}\cdot\mathcal{M} = M_0^2. \tag{6.8}$$

The components m_x and m_y may be determined from (6.7a, b). Differentiating (6.7a) with respect to time and using (6.7b), we obtain the equation for m_x,

$$\frac{d^2m_x}{dt^2} + \omega_0^2 m_x = \gamma H_1 \omega_0 M_0 \cos\omega t \tag{6.9}$$

To find the susceptibility we must solve this differential equation. We may write it in a more general form as

$$\mathcal{L}m_x(t) = F(t), \tag{6.10}$$

where $\mathcal{L}$ is the differential operator, in this case $d^2/dt^2 + \omega_0^2$. The solution is then given symbolically by

$$m_x(t) = \mathcal{L}^{-1}F(t), \tag{6.11}$$

where $\mathcal{L}^{-1}$ is the operator inverse of $\mathcal{L}$, provided that is exists. It is customary to rewrite this as

$$m_x(t) = \int_{-\infty}^{\infty} dt' \mathcal{L}^{-1} F(t') \delta(t - t') = \int_{-\infty}^{\infty} dt' \mathcal{G}(t, t') F(t') \tag{6.12}$$

where $\mathcal{G}(t, t')$ is the Green's function of the differential operator $\mathcal{L}$, defined as $\mathcal{L}^{-1}$ operating on the delta function.

A very important relationship between the Green's function and the susceptibility is easily obtained from (6.12). Taking the Fourier transform, we have

$$m_x(\omega) = \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} dt' \mathcal{G}(t, t') F(t') e^{i\omega t} \tag{6.13}$$

If the medium is stationary, it can be shown that the Green's function depends only on the relative time $t - t'$. Inserting the factor $\exp(-i\omega t') \exp(i\omega t')$ in (6.13), we obtain

$$m_x(\omega) = \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} dt' \mathcal{G}(t - t') e^{i\omega(t-t')} F(t') e^{i\omega t'} \tag{6.14}$$

Converting the integral over t into the integral over $\tau \equiv t - t'$ and recalling our definition of the susceptibility (1.58), we obtain

$$\chi_{xx}(\omega) = \gamma M_0 \omega_0 \mathcal{G}(\omega) \tag{6.15}$$

The factor $\gamma M_0 \omega_0$ arises because the forcing function $F(t)$ differs from the applied field by this quantity. The important result is that the Fourier transform of the Green's function is, essentially, the frequency-dependent susceptibility. Notice that if the so-called double-time Green's function defined on page 18 is introduced into equation (1.81), we obtain (6.15) directly, aside from the proportionality factor. This should not be surprising, for both Green's functions represent the linear response to a time-dependent magnetic field.

Let us now evaluate the Green's function. This is easily done by using the integral representation for the delta function.

$$\delta(t - t') = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega e^{-i\omega(t-t')}$$

Thus

$$\mathcal{G}(t - t') = \mathcal{L}^{-1} \delta(t - t') = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \frac{e^{i\omega(t-t')}}{-\omega^2 + \omega_0^2} \tag{6.16}$$

This integral is not defined until we specify how to treat the singularity at $\omega = \omega_0$. To resolve this difficulty we invoke causality and require that $\mathcal{G}(t - t') = 0$ for $t < t'$. This enables us to evaluate (6.16) unambiguously. The result is

$$\mathcal{G}(t - t') = \frac{\sin \omega_0 (t - t')}{\omega_0} \theta(t - t') \tag{6.17}$$

where $\theta(t - t')$ is 0 for $t < t'$ and 1 for $t > t'$. Taking the Fourier transform of (6.17) and multiplying the result by $\gamma M_0 \omega_0$, we finally obtain

$$\chi_{xx}(\omega) = \frac{\gamma M_0 \omega_0}{\omega_0^2 - \omega^2} + i \frac{\pi \gamma M_0}{2} \left[\delta(\omega - \omega_0) - \delta(\omega + \omega_0) \right]. \tag{6.18}$$

It is interesting to note that the real part of this susceptibility comes from the particular solution to (6.9), while the imaginary part, which is necessary to satisfy causality, comes from the solutions of the homogeneous version of (6.9). It is easily demonstrated that (6.18) satisfies the Kramers-Kronig relations (1.64), (1.65).

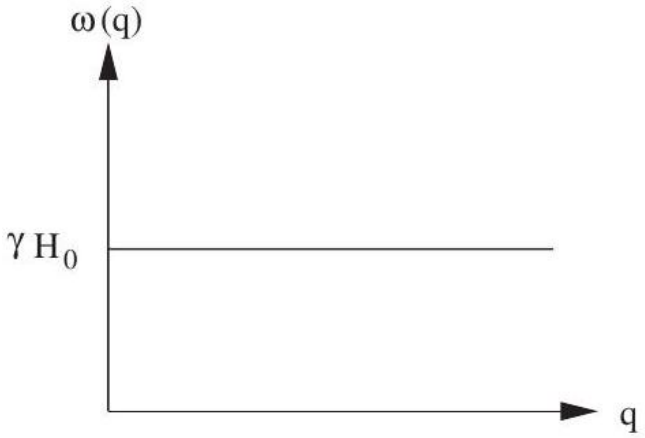


Fig. 6.1. The excitation spectrum of a system of uncoupled spins

The power absorbed by the magnetic system from this time-varying source is given by

$$P = \overline{-\boldsymbol{M} \cdot \frac{d\boldsymbol{H}_1(t)}{dt}}, \tag{6.19}$$

where the bar indicates a time average. From (1.55),

$$M_x(t) = H_1 \left[\chi'_{xx}(\omega) \cos \omega t + \chi''_{xx}(\omega) \sin \omega t\right], \tag{6.20}$$

and so (6.19) becomes

$$P = \frac{1}{2} \omega H_1^2 \chi''_{xx}(\omega). \tag{6.21}$$

From (6.18) we see that the power absorption of our noninteracting magnetic system occurs only at the frequency $\omega_0 = \gamma H_0$. Furthermore, this response will be infinite. In reality, of course, interactions within the system make this response finite and spread it over a distribution of frequencies. It is the determination of this response function to which we now address ourselves.

The poles in the complex response function define the excitation spectrum associated with the system. The real part of the pole gives the frequency of the excitation, and the imaginary part gives its damping. If the poles move too far from the real axis, the concept of an excitation is less well defined. Since we are dealing here with noninteracting localized moments, these poles are independent of wave vector. Therefore the excitation spectrum is as indicated in Fig. 6.1.

10.7.2 6.2 The Bloch Equations

If the magnetization is excited to a nonequilibrium value under the influence of an external field, when the field is suddenly removed the magnetization will relax back to its equilibrium value. The details of the relaxation depend on the nature of the interactions in the system. However, if we assume that this
relaxation, whatever its origin, has an exponential form, then we can develop a phenomenological description of the response function. In general, the longitudinal and transverse components of the magnetization may relax with different rates. The equations of motion, called the Bloch equations, are [128]

$$\frac{dm_z}{dt} = -\gamma(\boldsymbol{M} \times \boldsymbol{H})_z - \frac{m_z}{T_1} \tag{6.22a}$$

$$\frac{dm_{x,y}}{dt} = -\gamma(\boldsymbol{M} \times \boldsymbol{H})_{x,y} - \frac{m_{x,y}}{T_2} \tag{6.22b}$$

If we again assume a driving field of the form $\boldsymbol{H}_1(t) = H_1 \cos(\omega t)\hat{\boldsymbol{x}}$, the linearized equation for m_x is

$$\frac{d^2 m_x}{dt^2} + \frac{2}{T_2}\frac{dm_x}{dt} + \left(\omega_0^2 + \frac{1}{T_2^2}\right) m_x = \gamma H_1 \omega_0 M_0 \cos \omega t \tag{6.23}$$

By assuming that m_ν relaxes to 0 we have essentially incorporated causality into our solution, for the poles which enter the expression for the Green's function associated with this system, analogous to (6.16), are now displaced off the real axis, as shown in Fig. 6.2. Thus we speak of the susceptibility as being analytic in the upper half of the complex plane. The Green's function is then

$$\mathcal{G}\,(t - t') = \frac{\sin \omega_0\,(t - t')}{\omega_0} e^{-(t-t')/T_2} \theta\,(t - t') \tag{6.24}$$

Taking the Fourier transform leads us to

$$\chi_{xx}(\omega) = \chi'_{xx}(\omega) + i\chi''_{xx}(\omega) \tag{6.25}$$

where

$$\chi'_{xx}(\omega) = \frac{1}{2}\gamma M_0 \left[\frac{\omega_0 - \omega}{(\omega_0 - \omega)^2 + (1/T_2)^2} + \frac{\omega + \omega_0}{(\omega + \omega_0)^2 + (1/T_2)^2}\right] \tag{6.26}$$

and

$$\chi''_{xx}(\omega) = \frac{\gamma M_0}{2T_2} \left[\frac{1}{(\omega_0 - \omega)^2 + (1/T_2)^2} - \frac{1}{(\omega_0 + \omega)^2 + (1/T_2)^2}\right] \tag{6.27}$$

As $1/T_2$ becomes very small the functions in (6.27) become higher and narrower in such a way as to maintain their area. Therefore, in the limit $1/T_2 \to 0$, they may be considered as representations of delta functions; that is,

$$\lim_{\epsilon \to 0} \frac{\epsilon}{x^2 + \epsilon^2} = \pi\delta(x) \tag{6.28}$$

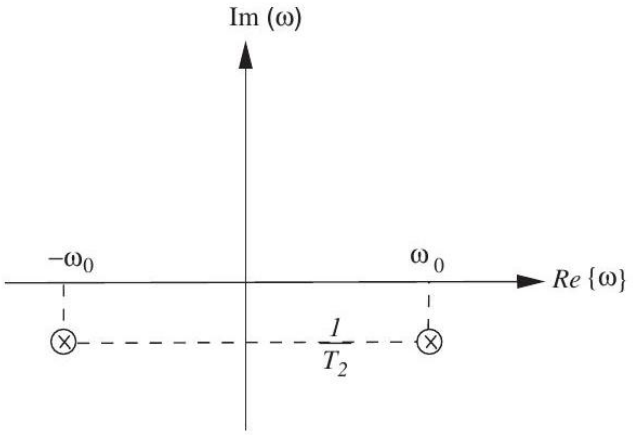


Fig. 6.2. Location of the poles in the imaginary part of the susceptibility

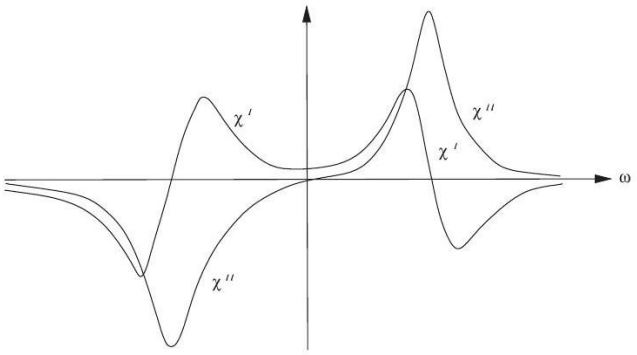


Fig. 6.3. Plot of the real and imaginary parts of the susceptibility, $\chi_{x x (\omega)}$

and (6.26), (6.27) reduce to our previous result (6.18). The real and imaginary part of the susceptibility are shown in Fig. 6.3.

Let us keep in mind that we are dealing with the response to a linearly polarized driving field. Some experiments employ a circularly polarized field. In such cases the response involves the other components of the susceptibility. For example, consider the circularly polarized field

$$\boldsymbol{H}_1(t) = H_1 \cos(\omega t)\hat{\boldsymbol{x}} + H_1 \sin(\omega t)\hat{\boldsymbol{y}}$$

When ω is positive this corresponds to counter-clockwise rotation. This is the direction $\boldsymbol{M}$ precesses in the presence of the dc field $H_0\hat{\boldsymbol{z}}$. From (1.55), the responses in the x and y directions to such a field may be written as

$$\begin{aligned} m_x(t) &= 2\pi H_1 \left[\left(\chi'_{xx}(\omega) - \chi''_{xy}(\omega)\right)\cos\omega t + \left(\chi''_{xx}(\omega) + \chi'_{xy}(\omega)\right)\sin\omega t\right] \\ m_y(t) &= 2\pi H_1 \left[\left(\chi'_{yx}(\omega) - \chi''_{yy}(\omega)\right)\cos\omega t + \left(\chi''_{yx}(\omega) + \chi'_{yy}(\omega)\right)\sin\omega t\right] \end{aligned}$$

Solving (6.22a,b) for m_y leads to the results

$$\begin{aligned} \chi'_{yx}(\omega) &= -\frac{\gamma M_0}{2T_2}\left[\frac{1}{(\omega_0-\omega)^2 + (1/T_2)^2} + \frac{1}{(\omega+\omega_0)^2 + (1/T_2)^2}\right] \\ \chi''_{yx}(\omega) &= \frac{\gamma M_0}{2}\left[\frac{\omega_0-\omega}{(\omega_0-\omega)^2 + (1/T_2)^2} - \frac{\omega_0+\omega}{(\omega+\omega_0)^2 + (1/T_2)^2}\right] \end{aligned}$$

If a field is applied in the y direction, the equation for m_y is identical to (6.23). Therefore $\chi_{yy} = \chi_{xx}$. However, $\chi_{xy} = -\chi_{yx}$. Combining these results, we find that the rotational component $m_+ = m_x + im_y$ has the solution

$$m_+(t) = \left(\chi'_+(\omega) + i\chi''_+(\omega)\right) H_1 e^{i\omega t},$$

where

$$\chi'_+(\omega) = \gamma M_0 \frac{\omega_0-\omega}{(\omega-\omega_0)^2 + (1/T_2)^2} \quad \text{and} \tag{6.29a}$$

$$\chi''_+(\omega) = -\frac{\gamma M_0}{T_2}\frac{1}{(\omega-\omega_0)^2 + (1/T_2)^2} \equiv -\pi\gamma M_0 f_L(\omega). \tag{6.29b}$$

Figure 6.4 shows the power absorbed as predicted by the Bloch equation for a circularly polarized field. While the Bloch equations offer a nice description of resonance the form given in (6.22a,b) has an interesting unphysical consequence. Namely, if we drive the system in

the clockwise (anti-Larmor) direction the power absorption becomes negative! The reason for this is that in (6.22a,b) the magnetization relaxes to the z-direction, i.e., the transverse components relax to zero. But, in reality, the magnetization should relax to the direction of the instantaneous field $\boldsymbol{H}_0 + \boldsymbol{H}_1(t)$.

The line shape $f_L(\omega)$ defined by (6.29b) is the familiar Lorentzian curve. This shape reflects the fact that it is the lifetime of the quantum states participating in the transition, in this case the Zeeman levels, that govern the profile. It is often found experimentally that the shape of the absorption more nearly resembles the Gaussian function

$$f_G(\omega) = \frac{1}{\Delta\sqrt{2\pi}} e^{-(\omega-\omega_0)^2/2\Delta^2}, \tag{6.30}$$

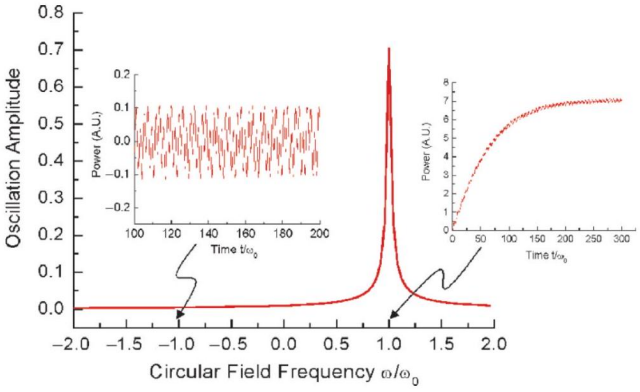


Fig. 6.4. Frequency response to a circularly polarized field ($H_{1 =0.0141 \omega_{0} / \gamma$) (Courtesy of Jian-gang Zhu)}

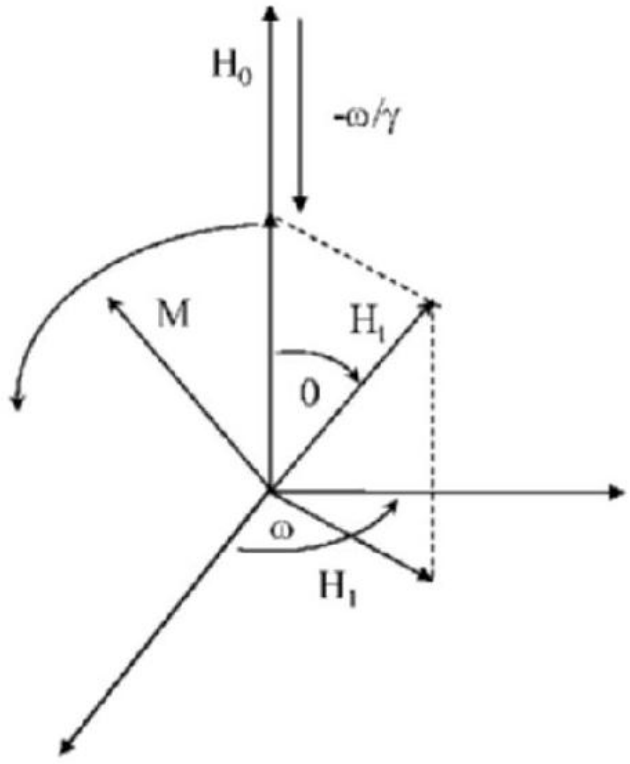


Fig. 6.5. Coordinate system rotating at a rate ω showing that the *dc* field is reduced by ω/γ

where Δ characterizes the width of the Gaussian. The appearance of such a shape is the result of inhomogeneous broadening. That is, the moments find themselves in different fields which give rise to a distribution of resonant frequencies. In general, the absorption may have some arbitrary shape $f(\omega)$. This shape tells us a great deal about the dynamics of the magnetic system. Adiabatic Rapid Passage. This discussion of susceptibility has been based on linearized equations which result in m_z being constant. The Bloch equations (6.22a,b), however, contain a rich variety of solutions. For example, in the expressions above we have assumed that the frequency of the driving field is fixed. Experimentally, however, one generally studies the resonance in a magnetic system by sweeping the value of the frequency or, equivalently, the *dc* field $\boldsymbol{H}_0$. This raises the question as to the relationship between the sweep rate and the relaxation times in the Bloch equations. Suppose, for example, we wish to reverse the magnetization by using resonance. We apply a circularlypolarized field and sweep through the resonance. This is most conveniently described by using the rotating coordinate system shown in Fig. 6.5. In this frame the total field is

$$\boldsymbol{H}_t = \left(\boldsymbol{H}_0 - \frac{\omega}{\gamma}\right) + \boldsymbol{H}_1.$$

Suppose we start at a frequency $\omega \ll \gamma H_0$. Then $\boldsymbol{H}_t$ is largely aligned with $\boldsymbol{H}_0$. As we increase $\omega, \boldsymbol{H}_t$ rotates out away from $\boldsymbol{H}_0$. The magnetization $\boldsymbol{M}$ will process around $\boldsymbol{H}_t$ at a rate γH_t. If we want $\boldsymbol{M}$ to follow $\boldsymbol{H}_t$ in a tight spiral, then the sweep rate, $d\theta/dt$, must be slow compared to $\gamma \boldsymbol{H}_t$. The smallest value of $\gamma \boldsymbol{H}_t$ occurs at resonance (i.e., $\boldsymbol{H}_t = \boldsymbol{H}_1$). Therefore this condition becomes

$$\frac{1}{\boldsymbol{H}_1}\frac{d\boldsymbol{H}_0}{dt} \ll \gamma \boldsymbol{H}_1.$$

When this condition holds we say that the motion is adiabatic. On the other hand, the sweep rate must be faster than the time it takes for the magnetization to relax back to its equilibrium direction along $\boldsymbol{H}_0, T_1$, i.e.

$$\frac{1}{\boldsymbol{H}_1}\frac{d\boldsymbol{H}_0}{dt} \gg \frac{1}{T_1}.$$

These conditions define "adiabatic rapid passage", the technique used by Bloch, Hansen, and Packard [129a] in their historic demonstration of nuclear induction.
Precessional Switching. The value of T_1 for the nuclear spins associated with the protons in water is approximately 3 seconds. Therefore it is relatively easy to reverse the nuclear magnetization by sweeping the field $\boldsymbol{H}_0$ in the presence of an rf field. Bloembergen [129b] pointed out that Bloch's equations, which Bloch developed with nuclear moments in mind, apply equally well to electronic moments, particularly those in a ferromagnet where they are exchange-coupled together to form a macroscopic magnetization, M. For electron spins, however, T_1 is very short, making it much more difficult to satisfy the adiabatic rapid passage conditions. It is, nevertheless, possible to exploit the gyromagnetic nature of the Bloch equations to reverse the magnetization of a ferromagnet. This is done by applying a dc field, H_{pulse} , at right angles to the magnetization for a time equal to half a precessional cycle. Consider a thin magnetic film initially magnetized in the plane of the film. The field H_{pulse} applied in the plane of the film at a right angle to the magnetization causes the magnetization to tip up out of the plane. This produces a relatively large demagnetization field which causes the magnetization to precess in a conical path. Such a "quasiballistic" magnetization reversal has been observed in a film of CoFe ($4\pi\text{M} = 10,800\text{oe}$) with a pulse having an amplitude of 81 oe and a duration of 175 psec [129c].

10.7.3 6.3 Resonance Line Shape

The shape of the resonance contains information about the interactions which govern the time dependence of the system. There are two approaches that have been very fruitful in understanding the nature of the magnetic-resonance line shape. One is the method of moments, which was employed so successfully by Van Vleck [130]. The other is what we might call the relaxation-function method, developed by Kubo and Tomito [131]. These approaches enable us to relate the phenomenological description given in the previous section to the microscopic physics.

6.3.1 The Method of Moments

If the resonance curve is described by a normalized shape function $f(\omega)$ centered at ω_0, then the nth moment with respect to ω_0 is defined by

$$M_n = \int_{-\infty}^{\infty} (\omega - \omega_0)^n f(\omega) d\omega \tag{6.31}$$

If the function $f(\omega)$ is symmetric about ω_0, it is convenient to introduce the function $F(\Omega) = f(\omega_0 + \Omega)$. In terms of this function, the nth moment is

$$M_n = \int_{-\infty}^{\infty} (\Omega)^n F(\Omega) d\Omega \tag{6.32}$$

If $\Gamma(t)$ is the Fourier transform of $F(\Omega)$,

$$\Gamma(t) = \int_{-\infty}^{\infty} F(\Omega) e^{-i\Omega t} d\Omega = 2 \int_0^{\infty} F(\Omega) \cos(\Omega t) d\Omega \tag{6.33}$$

then the $2n$th moment may be written as

$$M_{2n} = (-1)^n \frac{d^{2n}\Gamma(t)}{dt^{2n}}\bigg|_{t=0} \tag{6.34}$$

In order to obtain an explicit expression for the $2n$th moment let us assume that the width of the absorption at ω_0 is much less than ω_0 itself. We may then approximate the shape function by $\chi''(\omega)/\omega$. Let us also make the high-temperature approximation $g\mu_B H \ll k_B T$. With this approximation the correlation function introduced in Sect. 1.3.2 becomes [132]

$$\langle \mathcal{M}_x(t) \mathcal{M}_x \rangle \rightarrow \frac{\mathrm{Tr}\left\{\mathcal{M}_x(t)\mathcal{M}_x\right\}}{\mathrm{Tr}\{1\}} \equiv G(t) \tag{6.35}$$

Under these conditions the fluctuation-dissipation theorem (1.87) reduces to

$$\langle \mathcal{M}_x(t) \mathcal{M}_x \rangle = \frac{k_B T}{\pi} \int_{-\infty}^{\infty} d\omega \frac{\chi''_{xx}(\omega)}{\omega} e^{-i\omega t} \tag{6.36}$$

Therefore, since $\chi''(\omega)/\omega$ is proportioned to the lineshape $f(\omega)$,

$$G(t) = C \int_0^{\infty} d\omega f(\omega) \cos \omega t \tag{6.37}$$

where C is the coefficient of proportionality. This may be rewritten as

$$G(t) = C \int_{-\omega_0}^{\infty} d\Omega F(\Omega) \cos\left(\omega_0 + \Omega\right) t \tag{6.38}$$

If $F(\Omega)$ falls to 0 by the time Ω reaches $-\omega_0$, so that the lower limit may be extended to $-\infty$, then

$$G(t) = \left[2C \int_0^\infty d\Omega F(\Omega) \cos \Omega t\right] \cos \omega_0 t = G_1(t) \cos \omega_0 t \tag{6.39}$$

Therefore

$$M_{2n} = \frac{(-1)^n}{G_1(0)} \frac{d^{2n} G_1(t)}{dt^{2n}}\bigg|_{t=0} \tag{6.40}$$

Differentiating (6.35), (6.39), we obtain the explicit expressions for the second and fourth moments,

$$M_2 = -\frac{\operatorname{Tr}\left\{[\mathcal{H}, \mathcal{M}_x]^2\right\}}{\hbar^2 \operatorname{Tr}\left\{\mathcal{M}_x^2\right\}} \tag{6.41}$$

$$M_4 = \frac{\operatorname{Tr}\left\{\left[\mathcal{H}, [\mathcal{H}, \mathcal{M}_x]^2\right\}\right.}{\hbar^4 \operatorname{Tr}\left\{\mathcal{M}_x^2\right\}} - 6\omega_0^2 M_2 - \omega_0^4 \tag{6.42}$$

Let us consider what these moments would be for Lorentizian and Gaussian lines. Consider first the normalized Gaussian given by (6.30). From definition (6.31) we find

$$M_2 = \Delta^2 \quad \text{and} \quad M_4 = 3\Delta^4. \tag{6.43}$$

The halfwidth $\Delta\omega$, as defined by $f_G(\omega_0 + \Delta\omega) = \frac{1}{2} f_G(\omega_0)$, is

$$\Delta\omega = \sqrt{2 \ln 2}\Delta = 1.18\Delta \tag{6.44}$$

Therefore the square root of the second moment gives a good approximation to the width of a Gaussian line.

Now consider the normalized Lorentzian,

$$f_L(\omega) = \frac{1}{\pi T_2} \frac{1}{(\omega - \omega_0)^2 + (1/T_2)^2} \tag{6.45}$$

The integrals for the second and higher moments diverge for this function. Therefore a cutoff is usually introduced at $|\omega - \omega_0| = \omega_m$. The moments are then given by

$$M_2 = \frac{2\omega_m}{\pi T_2} \quad \text{and} \quad M_4 = \frac{2\omega_m^3}{3\pi T_2} \tag{6.46}$$

Notice that

$$\frac{M_4}{M_2^2} = \begin{cases} 3 & \text{for a Gaussian} \\ \frac{\pi\omega_m T_2}{6} & \text{for a cutoff Lorentzian.} \end{cases} \tag{6.47}$$

If we compute this ratio using (6.41) and (6.42) for some relaxation process defined by $\mathcal{H}$ and find that the result is very large, we suspect the line shape will be Lorentzian, whereas if the result is near 3 , the line shape will be closer to a Gaussian.

The problem to which Van Vleck addressed himself was that of a system of magnetic moments located in an external field and interacting weakly through
both the dipole-dipole interaction (2.50) and the exchange interaction (2.89). Thus the system is characterized by the Hamiltonian

$$\mathcal{H} = \mathcal{H}_Z + \mathcal{H}_{\text{dip}} + \mathcal{H}_{\text{ex}} \tag{6.48}$$

Notice that the last three terms in (2.50) have matrix elements between states of the system in which the total magnetic quantum number is changed by ±1 and ±2. These terms lead to transitions which are separated from the main resonance by multiples of $g\mu_B H_0$ and appear as sidebands on the main Zeeman line at $g\mu_B H_0$. As long as the sidebands do not overlap our main line, we may neglect the last three terms in (2.50), in so far as we are concerned only with the main line. The truncated Hamiltonian is thus

$$\mathcal{H} = \mathcal{H}_Z + \mathcal{H}'_{\text{dip}} + \mathcal{H}_{\text{ex}} \tag{6.49}$$

where

$$\mathcal{H}'_{\text{dip}} = g^2\mu_B^2 \sum_{i,j\neq i} \left[-\frac{3\cos^2\theta_{ij} - 1}{r_{ij}^3} S_i^z S_j^z + \frac{3\cos^2\theta_{ij} - 1}{4r_{ij}^3} \left(S_i^+ S_j^- + S_i^- S_j^+ \right) \right] \tag{6.50}$$

Since

$$\mathcal{M}_x = -g\mu_B \sum_i S_i^x \quad \text{and} \quad \mathcal{H}_Z = g\mu_B H \sum_i S_i^z$$

we find that

$$[\mathcal{H}_Z, \mathcal{M}_x] = i\hbar\omega_0 \mathcal{M}_y \tag{6.51}$$

It is also easy to show that $[\mathcal{H}_{\text{ex}}, \mathcal{M}_x] = 0$. Therefore the second moment reduces to

$$M_2 = -\frac{\text{Tr}\left\{ \left[\mathcal{H}'_{\text{dip}}, \mathcal{M}_x \right]^2 \right\}}{\hbar^2 \,\text{Tr}\left\{ \mathcal{M}_x^2 \right\}} \tag{6.52}$$

To evaluate the traces Van Vleck used the fact that the trace is independent of the basis. Therefore, even though this is truly a many-body system, we may use a basis consisting of products of individual spin states, $|M_{S_1}, M_{S_2}, \ldots\rangle$. We can then readily carry out the traces and find that

$$M_2 = \frac{3}{4} \frac{g^4\mu_B^4}{\hbar^2} S(S+1) \sum_i \frac{\left(1 - 3\cos^2\theta_{ij}\right)^2}{r_{ij}^6} \tag{6.53}$$

If the sample consists of randomly oriented crystallites, then the angular part of (6.53) averages to $\frac{4}{5}$. Furthermore, if we have a moment at every site in a simple cubic lattice with a lattice parameter a, then

$$\sum_i \frac{1}{r_{ij}^6} = \frac{8.32}{a^6} \quad \text{and} \quad M_2 = \frac{5\mu^4 S(S+1)}{\hbar^2 a^6}$$

For nuclear moments separated by, say 3Å, the corresponding line width $\Delta H = \sqrt{M_2}\gamma_N$ is about 6 gauss, which is in reasonable agreement with observations in nuclear magnetic-resonance experiments. For electronic moments, however, this line width is about 5 kG . This is enormous. But when we have electronic moments as close as 3°Å, the exchange interaction becomes very important, and the second moment, which does not contain exchange effects, is not sufficient to describe the situation. In particular, for a Lorentzian line (6.46) gives

$$\frac{1}{T_2} = \frac{\pi M_2}{2} \sqrt{\frac{M_2}{3M_4}} \tag{6.54}$$

To study the effects of exchange Van Vleck also computed the fourth moment. What he found was that when the exchange interaction exceeds the dipoledipole interaction, the fourth

moment exceeds the square of the second moment. Equation (6.47) then tells us that as the exchange interaction increases, the shape of the absorption line changes from Gaussian to Lorentzian. Furthermore, if the fourth moment increases while the second moment remains unchanged, the intensity in the wings of the absorption curve increases. If the total integrated intensity is to remain the same, there must be a decrease in intensity closer to the center of the line. This results in what is called exchange narrowing, illustrated in Fig. 6.6. The origin of this narrowing may be better understood by considering the effective field acting on a particular spin. In the absence of any exchange, this spin experiences not only the applied field $\boldsymbol{H}_0$, but also a dipolar field arising from the other spins in its environment. Since each spin experiences a slightly different environment, this leads to a distribution of resonant frequencies which manifests itself as an inhomogeneously broadened absorption. The effect of the exchange interaction is to modulate the dipolar field. As this modulation increases the average dipolar field decreases, with the result that the distribution of resonant frequencies peaks more strongly near ω_0.

6.3.2 The Relaxation-Function Method

Guided by this interpretation of the results of the moment calculations, let us now outline the relaxation-function approach to obtain the detailed shape of the resonance curve. Our discussion follows that of Abragam [133]. Since we are assuming that the absorption spectrum is proportional to $\chi''(\omega)/\omega$, we must compute the high-temperature correlation function introduced in (6.35),

$$G(t) = \frac{\mathrm{Tr}\,\{\mathcal{M}_x(t)\mathcal{M}_x\}}{\mathrm{Tr}\{1\}} \tag{6.55}$$

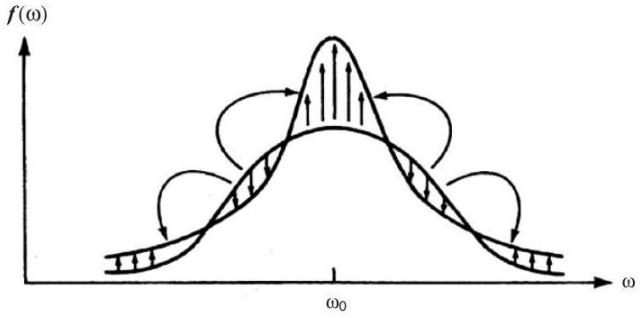


Fig. 6.6. Schematic represenation of exchange narrowing. The arrows indicate the transfer of intensity

We shall not include the exchange interaction explicitly, but shall merely assume that it leads to a time-dependent dipolar interaction. The total Hamiltonian is then $\mathcal{H} = \mathcal{H}_Z + \mathcal{H}_{\mathrm{dip}}$. The dipole-dipole interaction is assumed to be small in comparison with the Zeeman interaction. Therefore we look for a solution to $\mathcal{M}_x(t)$ which is an expansion in powers of the dipoledipole interaction. This is very much different from the method of moments, which is essentially an expansion of $G(t)$ in powers of t.

In order to develop an expansion of $\mathcal{M}_x(t)$ in the dipole interaction it is convenient to remove the Zeeman interaction by the transformation

$$\tilde{\mathcal{M}}_x \equiv \exp\left(\frac{-i\mathcal{H}_Z t}{\hbar}\right)\mathcal{M}_x(t)\exp\left(\frac{i\mathcal{H}_Z t}{\hbar}\right). \tag{6.56}$$

The equation of motion for this operator is

$$-i\hbar\frac{d\tilde{\mathcal{M}}_x}{dt} = \left[\exp\left(\frac{-\mathcal{H}_Z t}{\hbar}\right)\mathcal{H}_{\text{dip}}\exp\left(\frac{i\mathcal{H}_Z t}{\hbar}\right), \tilde{\mathcal{M}}_x\right]. \tag{6.57}$$

The corresponding equation for the matrix element of $\tilde{\mathcal{M}}_x$ between eigenstates of $\mathcal{H}_Z$ is

$$\begin{aligned}-i\hbar\frac{d}{dt}\left(\tilde{\mathcal{M}}_x\right)_{nn'} = \sum_{n''}\Big\{&\left(\mathcal{H}_{\text{dip}}\right)_{nn''}\left(\tilde{\mathcal{M}}_x\right)_{n''n'}\exp\left[i\left(E_{n''}-E_n\right)t/\hbar\right] \\ &-\left(\tilde{\mathcal{M}}_x\right)_{nn''}\left(\mathcal{H}_{\text{dip}}\right)_{n''n'}\exp\left[-i\left(E_{n''}-E_{n'}\right)t/\hbar\right]\Big\}\end{aligned} \tag{6.58}$$

Since the transformation (6.56) has removed the rapid Zeeman precession from $\widetilde{\mathcal{M}}_x$, we may assume that remaining time dependence is relatively slow. Then, since $E_{n''}-E_n$ and $E_{n''}-E_{n'}$ are multiples of $\hbar\omega_0$, those terms in (6.58) for which these differences are not 0 oscillate rapidly, giving a very small contribution which we shall neglect. This leaves us with matrix elements of the form $\left(\mathcal{H}_{\text{dip}}\right)_{nn''}$ between degenerate states. We shall assume that these states are chosen such that these matrix elements vanish unless $n''=n$. Thus (6.58) reduces to

$$-i\hbar\frac{d}{dt}\left(\tilde{\mathcal{M}}_x\right)_{nn'} = \hbar\Delta\omega(t)_{nn'}\left(\tilde{\mathcal{M}}_x\right)_{nn'} \tag{6.59}$$

where

$$\hbar\Delta\omega(t)_{nn'} \equiv \left(\mathcal{H}_{\text{dip}}\right)_{nn} - \left(\mathcal{H}_{\text{dip}}\right)_{n'n'} \tag{6.60}$$

The time dependence of this quantity reflects the exchange modulation. That is, the exchange interaction causes the spin system to fluctuate among the degenerate states within $|n\rangle$ giving rise to different values of $\left(\mathcal{H}_{\text{dip}}\right)_{nn}$. Equation (6.59) has the solution

$$\left(\tilde{\mathcal{M}}_x\right)_{nn'} = \left(\tilde{\mathcal{M}}_x\right)^0_{nn'}\exp\left[i\int_0^t\Delta\omega\left(t'\right)_{nn'}dt'\right] \tag{6.61}$$

If we invert (6.56) and use the solution (6.61), the correlation function becomes

$$G(t) = \sum_{n,n'}\exp\left[i\left(E_n-E_{n'}\right)t/\hbar\right]\left|\left\langle n\right|\mathcal{M}_x\left|n'\right\rangle\right|^2\exp\left[i\int_0^t\Delta\omega\left(t'\right)_{nn'}dt'\right] \tag{6.62}$$

Let us consider only the main absorption line, for which $E_n - E_{n'} = \hbar\omega_0$. This restricts the sum over n and n' to those pairs whose eigenvalues differ by $\hbar\omega_0$. If we define

$$\int_0^t\Delta\omega\left(t'\right)_{nn'}dt' \equiv x(t) \tag{6.63}$$

then for those states n and n' which give a particular $x(t)$ we also have a corresponding value for $\left|\langle n|\mathcal{M}_x|\,n'\rangle\right|^2$, which we write as $P[x(t)]$. Thus the correlation function may be thought of as the average of $\exp[ix(t)]$ using the probability distribution $P[x(t)]$; that is,

$$G(t) = e^{i\omega_0 t}\int_{-\infty}^{\infty}dxP(x)e^{ix} = e^{i\omega_0 t}\left\langle e^{ix}\right\rangle \tag{6.64}$$

where $P(x)$ includes a normalizing factor as well as the matrix element appearing in (6.62).

The assumption is now made that $\Delta\omega\left(t'\right)_{nn'}$ is a Gaussian function of time, with a mean value given by the second moment M_2, which we computed earlier. Thus we take

$$\langle\Delta\omega(t)_{nn'}\Delta\omega(t+\tau)_{nn'}\rangle = M_2\psi(\tau) \tag{6.65}$$

where $\psi(\tau)$ characterizes the fluctuations in the local dipolar field as it is randomly modulated by the exchange interaction. There is a theorem, called the central-limit theorem, which says, in effect, that if $\Delta\omega\left(t'\right)_{nn'}$ is a Gaussian
function, then $x(t)$, as defined by (6.63), is also a Gaussian function. This means that $P[x(t)]$ has the Gaussian form

$$P[x(t)] = \frac{1}{\sqrt{2\pi\left\langle x^2\right\rangle}} \exp\left(-x^2/2\left\langle x^2\right\rangle\right) \tag{6.66}$$

Thus

$$\left\langle e^{ix}\right\rangle = e^{-\left\langle x^2\right\rangle/2}$$

where

$$\left\langle x^2\right\rangle = \left\langle \int_0^t dt' \int_0^t dt'' \Delta\omega\left(t'\right)_{nn'} \Delta\omega\left(t''\right)_{nn'} \right\rangle = 2M_2 \int_0 d\tau (t-\tau)\psi(\tau) \tag{6.67}$$

Therefore we finally obtain the result that the correlation function whose Fourier transform gives the lineshape is given by

$$G(t) = e^{i\omega_0 t}\varphi(t) \tag{6.68}$$

where the relaxation function, $\varphi(t)$, is

$$\varphi(t) = \exp\left[-M_2 \int_0^t (t-\tau)\psi(\tau)d\tau\right] \tag{6.69}$$

The narrowing process may be seen from this expression. If the correlation time that characterizes $\psi(\tau)$ is long, then $\psi(\tau) \simeq 1$ over the range of the integral in (6.69) and

$$\varphi(t) \to \exp\left(-M_2 t^2/2\right) \tag{6.70}$$

This leads to a Gaussian line with the Van Vleck second moment. If, on the other hand, $\psi(\tau)$ decays before we reach the upper limit of the integral, then

$$M_2 \int_0^t (t-\tau)\psi(\tau)d\tau \simeq M_2 t \int_0^\infty \psi(\tau)d\tau \equiv M_2 t/\omega_{\text{ex}} \tag{6.71}$$

and

$$\varphi(t) \to \exp\left(-M_2 t/\omega_{\text{ex}}\right) \tag{6.72}$$

This leads to a Lorentzian line with a halfwidth proportional of M_2/ω_{ex}. Thus as the exchange increases, the dipolar broadened line changes from Gaussian to narrow Lorentzian.

An analytic form for $\psi(\tau)$ which contains both the Gaussian and Lorentzian limits is

$$\psi(\tau) = \exp\left(-\pi\omega_{\text{ex}}^2\tau^2/4\right) \tag{6.73}$$

6.3.3 Spin Diffusion

Detailed comparison with measured correlation functions shows that $\psi(\tau)$ actually decays more slowly than a Gaussian. To obtain a better description of the relaxation at long times we recognize that if we perturb the spin system at some point $\boldsymbol{r}$ this perturbation will propagate away by means of the exchange interaction. At long times this process will involve many spins and may be described by a diffusion process. This means that the fluctuations described by $\psi(t)$ will also be spatially dependent, i.e., $\psi(\boldsymbol{r}, t)$ and will obey the diffusion equation,

$$\Lambda \nabla^2 \psi = \frac{\partial \psi}{\partial t} \tag{6.74}$$

where the diffusion coefficient Λ is related to the correlation time ω_{ex}^{-1}. The difference between Gaussian behavior and diffusive behavior is illustrated in Fig. 6.7.

It is interesting to note from the diffusion equation that $\psi(\tau)$ must vary as $\tau^{-d/2}$ where d is the dimensionality. Thus, in one dimension the correlation time ω_{ex}^{-1} defined by (6.71) diverges and we expect an anomalous line shape. $(CH_3)_4\,NMnCl_3$, abbreviated TMMC, is a system in which the Mn^{2+} spins are exchange coupled only along chains approximating a one-dimensional spin system. Magnetic resonance studies indeed show [134] that the line shape is the Fourier transform of $\exp\left(-At^{3/2}\right)$, intermediate between a Gaussian and a Lorentzian.

Although we have considered here the case of exchange-narrowed dipolarbroadened lines, the relaxation-function approach may also be applied in a variety of other situations, such as exchange-narrowed hyperfine-broadened nuclear magnetic resonance lines [135] or motionally narrowed, dipolarbroadened nuclear resonance lines [136].

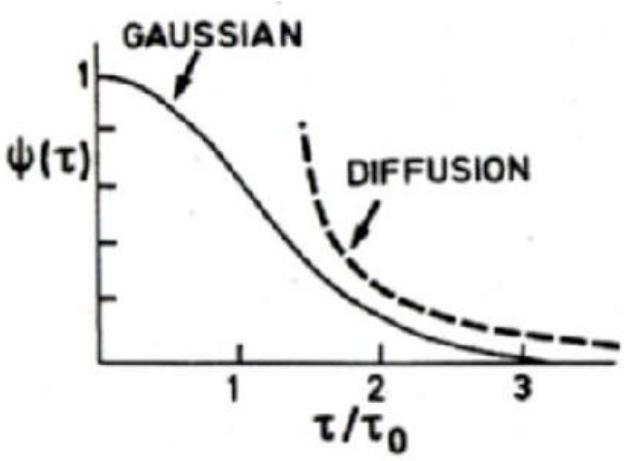


Fig. 6.7. Comparison of the Gaussian and diffusion approximations to the correlation function $\psi(\tau)$. The time is in units of $\left(1 / \tau_{0 \right)^{2}=2 z(J / \hbar)^{2} S(S+1) / 3$ where z is the number of of neighbors and J is the exchange}

10.7.4 6.4 Spin Echoes

6.4.1 Measurement of T_1

Notice that the longitudinal relaxation time T_1 does not enter the linear response function (6.27). Had we included the nonlinear contributions to this response function, the additional term $\gamma^2 H_1^2 T_1 T_2$ would have appeared in the denominator. As a consequence, when the amplitude of the driving field becomes comparable to $(\gamma^2 T_1 T_2)^{-1/2}$ the resonance begins to saturate. The onset of this saturation provides a measure of T_1.

The same nonlinearity is also employed to produce spin echoes [137]. This phenomenon is easily understood with the help of Fig. 6.8. A circularily polarized field H_1 at frequency ω_0 is applied transverse to the *dc* field H_0. We know from the Bloch equations that in a coordinate system rotating about the z-axis, at a rate ω_0, the magnetization will precess away from the z-axis as illustrated in Fig. 6.8a. If the driving field is turned off just as the magnetization reaches the xy-plane, a 90° pulse is produced. Leaving the driving field on for twice as long results in a 180° pulse, which just inverts the magnetization.

Suppose that at time $t = 0$ we apply a 90° pulse. After this pulse is turned off the spins process at their own resonant frequencies, which, because of the dipolar field, are distributed about ω_0. After some dephasing time T_2^* the spins produce a "pancake" in the rotating system, as shown in Fig. 6.8b. A 180° pulse applied at some time τ to the system in this configuration will cause this pancake to rotate 180° about the x' axis, as shown in Fig. 6.8c.

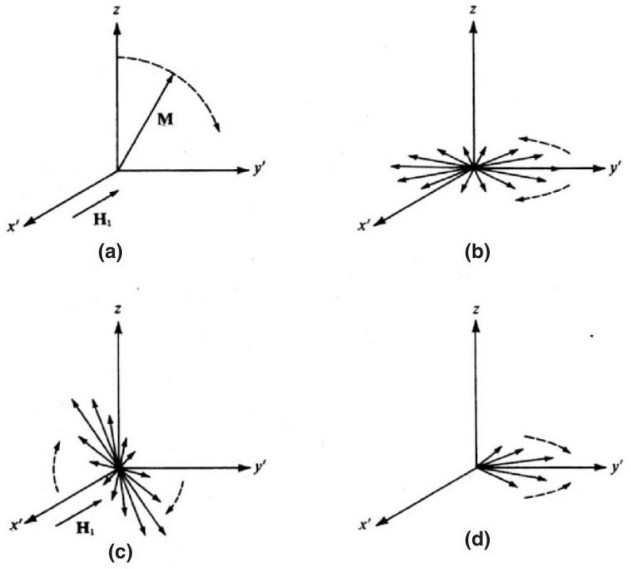


Fig. 6.8. Description of the origin of a spin echo in a coordinate frame rotating about the z axis with the Zeeman frequency

Now, as the spins precess they begin to rephase and at a time τ after the 180° pulse they again become coherent, producing an "echo". If we continue to apply 180° pulses at times $3\tau, 5\tau$, etc., we shall observe echoes at times 4τ, 6τ, etc. If there were no transverse relaxation, these echoes would all have the same height. However, because of this relaxation the height of the nth echo decreases as $\exp(-2n\tau/T_2)$, providing us with a measure of T_2. This scheme is shown in Fig. 6.9.

If $T_1 \gg T_2$, then the longitudinal relaxation time T_1 may be obtained as follows. First we apply a 90° pulse, which takes M_z to zero. After a time τ, M_z has returned to $M_0[1 - \exp(-\tau/T_1)]$. If we now apply a 90° pulse followed by a 180° pulse, the height of the resulting echo will be proportional to $M_z(\tau)$. By repeating this sequence we can follow M_z back to M_0.

Let us consider briefly still another way of determining T_1. This consists of measuring the response of the system to a low-frequency longitudinal driving field. In general, the magnetization relaxes toward the instantaneous field. Therefore the Bloch equation (6.22a) for a field $\boldsymbol{H} = H_0\hat{\boldsymbol{z}} + H_1 \cos\omega t\hat{\boldsymbol{z}}$ should be written as

$$\frac{dm_z}{dt} = \gamma(\boldsymbol{M} \times \boldsymbol{H})_z + \frac{\chi_0 H_1 \cos\omega t - m_z}{T_1}. \tag{6.75}$$

This is only an approximation, in the sense that we have used the static susceptibility χ_0. Solving (6.75) leads directly to the response functions

$$\chi'_{zz}(\omega) = \frac{\chi_0}{1 + \omega^2 T_1^2}, \tag{6.76a}$$

$$\chi''_{zz}(\omega) = \frac{\chi_0 \omega T_1}{1 + \omega^2 T_1^2}. \tag{6.76b}$$

These functions are plotted in Fig. 6.10. For electronic spins this broad longitudinal absorption is in the radio-frequency part of the spectrum. This is far below the transverse resonance, which is in the microwave region. Microwave sources were unavailable until World War II, and consequently the relaxation time T_1 was investigated much earlier than T_2. In fact, the response function (6.76b) was derived from thermodynamic considerations long before the development of Bloch's equations.

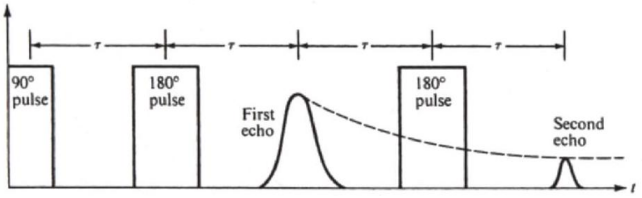


Fig. 6.9. Spin echo scheme for measuring $T_{2
$}

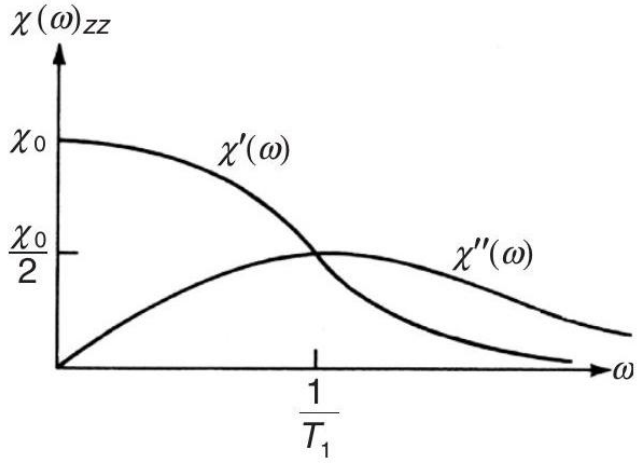


Fig. 6.10. Longitudinal susceptibility of a weakly interacting spin system

6.4.2 Calculation of T_1

Let us now derive an expression for the phenomenological relaxation time T_1. According to the Bloch equation (6.22a), if the z component of the magnetization is reduced from its equilibrium value M_0, it will return to this value according to

$$M_z = M_0 \left(1 - e^{t/T_1}\right). \tag{6.77}$$

To obtain an explicit expression for T_1 let us compute M_z for a simple twolevel spin system that has been displaced from equilibrium. If the number of spins in the state $\left|-\frac{1}{2}\right\rangle$ is $N_-$and the number in $\left|+\frac{1}{2}\right\rangle$ is N_+, then the total magnetization is

$$M_z = \frac{\frac{1}{2}g\mu_B\left(N_- - N_+\right)}{V} = \frac{1}{2}g\mu_B n. \tag{6.78}$$

Thus, since the z component of the magnetization depends on the relative population of the Zeeman states, any change in this population will result in a change in M_z.

Let us define $W(- \to +)$ as the probability per unit time that a spin in the state $\left|-\frac{1}{2}\right\rangle$ is "flipped" to the state $\left|+\frac{1}{2}\right\rangle$. Then the change in the populations may be described by the rate equations

$$\frac{dN_+}{dt} = N_- W(- \to +) - N_+ W(+ \to -), \tag{6.79a}$$

$$\frac{dN_-}{dt} = N_+ W(+ \to -) - N_- W(- \to +). \tag{6.79b}$$

In equilibrium,

$$N_-^0 W(- \to +) - N_+^0 W(+ \to -) = 0. \tag{6.80}$$

Therefore, if the levels are separated by $g\mu_B H$ and we define $W(+ \to -) \equiv W$, then

$$W(- \to +) = W \exp\left(-g\mu_B H/k_B T\right). \tag{6.81}$$

The equation for n becomes

$$\frac{dn}{dt} = \frac{2N_+ W}{V} = \frac{2N_- W \exp\left(-g\mu_B H/k_B T\right)}{V}. \tag{6.82}$$

In the high-temperature limit, valid for most experimental situations, the exponential in (6.82) may be expanded. If we replace the occupation numbers in (6.78) by their equilibrium values, we see that n_0 is approximately $(N_-/V)\left(g\mu_B H/k_B T\right)$ and

$$\frac{dn}{dt} = -2W\left(n - n_0\right) \tag{6.83}$$

Solving this equation and using (6.78), we obtain

$$M_z = M_0\left(1 - e^{-2Wt}\right) \tag{6.84}$$

Comparing this with (6.77), we have the result

$$\frac{1}{T_1} = 2W \tag{6.85}$$

There are various mechanisms that might contribute to the transition probability W. For example, in a paramagnet the spins may be coupled to the lattice vibrations. In this case the spin flips are accompanied by the emission or absorption of phonons. The resulting T_1 is called the spin-lattice relaxation time.

Another example is that of nuclear spins in a ferromagnet. In this case the nuclear spins are coupled to the electron spins by the hyperfine interaction. Fluctuations in the electron-spin system can then flip the nuclear spins, leading to a nuclear-spin relaxation time T_1. Let us consider this particular mechanism as an example of the calculation of T_1.

Suppose our unperturbed system is a single nuclear spin of $\frac{1}{2}$ plus the exchange-coupled electron system. Thus

$$\mathcal{H}_0 = \mathcal{H}_N + \mathcal{H}_{\text{el}} \tag{6.86}$$

where $\mathcal{H}_N = -g_N \mu_N H I_z$ and $\mathcal{H}_{\text{el}}$ is the electronic Hamiltonian. The eigenfunctions of $\mathcal{H}_0$ have the form $|m_I, n\rangle$, where n characterizes the state of the electron system. The perturbing Hamiltonian is the hyperfine interaction

$$\mathcal{H}_1 = A\boldsymbol{I} \cdot \boldsymbol{S}, \tag{6.87}$$

where $\boldsymbol{S}$ is the spin of the electron associated with the particular nucleus we are considering. If, at time $t = 0$, the system is in the state $|m_1, n\rangle$, then we want to know the probability of

finding the system in the state $|m_I', n'\rangle$ at some later time t. To find this probability we must solve the Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial t} = (\mathcal{H}_0 + \mathcal{H}_1)\,\psi. \tag{6.88}$$

We shall see in a moment that it is convenient to introduce

$$\psi = \exp\left(\frac{-i\mathcal{H}_0 t}{\hbar}\right)\phi, \tag{6.89}$$

where ϕ satisfies

$$i\hbar\frac{\partial\phi}{\partial t} = \mathcal{H}_1(t)\phi \tag{6.90}$$

and

$$\mathcal{H}_1(t) \equiv \exp\left(\frac{i\mathcal{H}_0 t}{\hbar}\right)\mathcal{H}_1 \exp\left(\frac{-i\mathcal{H}_0 t}{\hbar}\right). \tag{6.91}$$

Integrating this equation with the condition that $\phi(t=0) = |m_I, n\rangle$, we obtain

$$\phi = \exp\left[-\frac{i}{\hbar}\int_0^t dt'\mathcal{H}_1\left(t'\right)\right]|m_I, n\rangle\,. \tag{6.92}$$

The reason for introducing ϕ is that (6.92) may be expanded in powers of the perturbation $\mathcal{H}_1$,

$$\phi = |m_I, n\rangle - \frac{i}{\hbar}\int_0^t dt'\mathcal{H}_1\left(t'\right)|m_I, n\rangle + \ldots \tag{6.93}$$

Therefore the probability of finding the system in the state $|m_I', n'\rangle$ is

$$\begin{aligned}\left|\langle m_I', n' \mid \psi\rangle\right|^2 =& \frac{1}{\hbar^2}\int_0^t dt'\int_0^t dt''\,\langle m_I, n|\,\mathcal{H}_1\left(t'\,\right)|m_1', n'\rangle \\ &\times\langle m_1', n'|\,\mathcal{H}_1\left(t''\right)|m_I, n\rangle + \ldots\end{aligned} \tag{6.94}$$

The total probability per unit time that the system will make a transition is obtained by dividing (6.94) by t and letting $t \to \infty$. To find W we set $m_I = -\frac{1}{2}$ and $m_I' = +\frac{1}{2}$. We also average over all the electronic degrees of freedom. That is, we multiply (6.94) by the probability that the electronic system will be in the initial state n and then we sum over n as well as over all the final states n'. The sum over n' gives unity by closure, leaving us with just the thermal average of $\mathcal{H}_1\left(t'\right)\mathcal{H}_1\left(t''\right)$, which we denote by $\langle\ldots\rangle$. This is not to be confused with the bras and kets above. Thus

$$W = \lim_{t\to\infty}\frac{A^2}{4\hbar^2 t}\int_0^t dt''\exp\left[i\omega_N\left(t'-t''\right)\right]\left\langle S^-\left(t'\right)S^+\left(t''\right)\right\rangle \tag{6.95}$$

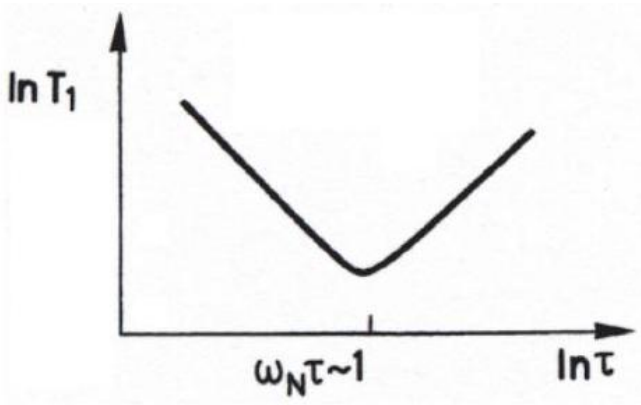


Fig. 6.11. Variation of $T_{1
$ with the correlation time τ according to the BPP model}

where $\omega_N = g_N \mu_N H$ is the nuclear magnetic resonance frequency. Changing variables leads to the result

$$\frac{1}{T_1} = \frac{A^2}{2\hbar^2} \int_{-\infty}^{\infty} dt \left\langle S^-(t) S^+ \right\rangle \exp\left(i\omega_N t\right) \tag{6.96}$$

Thus we see that the longitudinal relaxation frequency in this case is the Fourier component at the nuclear magnetic resonance frequency of the electronspin correlation function. If the correlation function relaxes exponentially, i.e.,

$$\left\langle S^-(t) S^+ \right\rangle = \left\langle S_\perp^2 \right\rangle e^{-t/\tau},$$

then

$$\frac{1}{T_1} = \frac{A^2 \left\langle S_\perp^2 \right\rangle}{\hbar^2} \frac{\tau}{1 + \omega_N^2 \tau^2}. \tag{6.97}$$

The variation of T_1 with τ is shown in Fig. 6.11. Thus as τ varies, due to a variation in temperature for example, T_1 goes through a minimum when $\omega_N \tau \sim 1$. This characteristic behavior was first derived by Bloembergen, Purcell, and Pound [138] and is often simply referred to as the BPP result.

Problems

6.1. Consider a moment M initially pointing along the $+z$ direction. If we apply a dc field H in the $-z$-direction describe the behavior of M_x, M_y, M_z according to the Bloch equations.
6.2. Show that the real and imaginary parts of the susceptibility (6.26) and (6.3) satisfy the Kramers-Kronig relations (1.64) and (1.65).
6.3. Calculate the second and fourth moments (i.e., (6.43)) of the Gaussian function (6.30)
6.4. Show that $[\mathcal{H}_{\text{ex}}, M_x] = 0$. What does this imply about the torque on the magnetization from exchange?
6.5. Would a single-ion anisotropy of the form

$$\mathcal{H}_K = \sum_i K \left(S_i^z\right)^2$$

contribute to the broadening of the resonance?
6.6. Solve the linearized Bloch equations,

$$\dot{m}_x = -\frac{1}{T_2}m_x - \gamma H_y m_y + \gamma h_y(t)$$
$$\dot{m}_y = -\frac{1}{T_2}m_y + \gamma H_x m_x - \gamma h_x(t)$$

where $H_{x,y}$ are called the Kittel "stiffness" fields, to obtain $\chi_{\alpha\beta}(\omega)$. Use the fluctuation-dissipation theorem to find the correlation functions $\langle m_\alpha(\tau)m_\beta(0)\rangle$ ($\alpha, \beta = x, y$). The "noise power" is given by setting $\tau = 0$. Use the result to calculate the energy $\langle\varepsilon\rangle = M_s V H_x \langle m_x^2\rangle$. What is the meaning of this result? [139]

10.8 The Dynamic Susceptibility of Weakly Interacting Systems: Metals

Let us now investigate the response of a metal to a time-dependent magnetic field. Experimentally, such investigations are very difficult because of problems associated with the fact that the driving field penetrates only a short distance into the metal. The first resonance experiments involving conductionelectron spins were carried out in the early 1950s. These experiments entailed measuring the surface impedance of samples located in a microwave cavity.

In 1964 electron spin resonance was observed in transmission. The analysis of both resonance techniques requires a knowledge of the nonlocal susceptibility function $\chi(z, z', \omega)$. This may be obtained from the generalized susceptibility $\chi(\boldsymbol{q}, \omega)$ along with a boundary condition on $\boldsymbol{M}$. If we assume that the boundaries do not relax the magnetization then the magnetization current $\boldsymbol{\nabla} \times \boldsymbol{M}$ must be conserved at the boundary. Let us now consider the generalized susceptibility. One approach that is particularly instructive makes use of the fluctuation-dissipation theorem. As we saw earlier, the presence of a time-dependent field leads to a coupling between the x and y components of the magnetization. Consequently, it is more convenient to work with the rotational components $M_\pm(\boldsymbol{q}, \omega)$. In terms of these components, the susceptibility $\chi''_{xx} + \chi''_{yy}$ is proportional to the correlation function $\langle \mathcal{M}_-(\boldsymbol{q}, t)\mathcal{M}_+(-\boldsymbol{q})\rangle$. This is just the susceptibility $\chi_{-+}(\boldsymbol{q}, \omega)$ we have already encountered.

The magnetization operator associated with the spin of a system of itinerant electrons is

$$\mathcal{M}(\boldsymbol{r}) = \sum_i \mu_B \boldsymbol{\sigma}_i \delta(\boldsymbol{r} - \boldsymbol{r}_i) \tag{7.1}$$

Second quantizing this operator in terms of the plane-waves states (1.130) gives

$$\mathcal{M}(\boldsymbol{r}) = \frac{\mu_B}{V}\sum_{\boldsymbol{k},\boldsymbol{q}}\sum_{\alpha,\beta}\boldsymbol{\sigma}_{\alpha\beta}e^{i\boldsymbol{q}\cdot\boldsymbol{r}}a^\dagger_{\boldsymbol{k}-\boldsymbol{q},\alpha}a_{\boldsymbol{k},\beta}. \tag{7.2}$$

Therefore

$$\mathcal{M}_+(-\boldsymbol{q}) = -2\mu_B\sum_{\boldsymbol{k}} a^\dagger_{\boldsymbol{k}+\boldsymbol{q},\uparrow}a_{\boldsymbol{k}\downarrow} \tag{7.3}$$

and

$$\mathcal{M}_-(\boldsymbol{q}) = -2\mu_B\sum_{\boldsymbol{k}} a^\dagger_{\boldsymbol{k}-\boldsymbol{q},\downarrow}a_{\boldsymbol{k}\uparrow} \tag{7.4}$$

The desired correlation function is then

$$\langle \mathcal{M}_-(\boldsymbol{q},t)\mathcal{M}_+(-\boldsymbol{q})\rangle = 4\mu_B^2 \sum_{\boldsymbol{k}} \sum_{\boldsymbol{k}'} \times \left\langle \exp\left(\frac{i\mathcal{H}t}{\hbar}\right) a^\dagger_{\boldsymbol{k}'-\boldsymbol{q},\downarrow} a_{\boldsymbol{k}'\uparrow} \exp\left(\frac{-i\mathcal{H}t}{\hbar}\right) a^\dagger_{\boldsymbol{k}+\boldsymbol{q},\uparrow} a_{\boldsymbol{k}\downarrow} \right\rangle \tag{7.5}$$

where

$$\mathcal{H} = \sum_{\boldsymbol{k}} \sum_{\sigma} \epsilon_{\boldsymbol{k}\sigma} a^\dagger_{\boldsymbol{k}\sigma} a_{\boldsymbol{k}\sigma}. \tag{7.6}$$

The eigenvalue $\epsilon_{\boldsymbol{k}\sigma}$ is $\hbar^2k^2/2m + mu_BH_0\sigma$, where σ is $+1$ for an up spin and -1 for a down spin.

At $T = 0$ K the ground state consists of up- and down-spin Fermi spheres. As in our calculation of the static susceptibility in Chap. 3, the wave vector $\boldsymbol{k}$ is restricted to those values within the down-spin Fermi sphere such that $\boldsymbol{k}+\boldsymbol{q}$ lies outside the up-spin Fermi sphere. The correlation function at $T = 0$ K is then

$$\langle \mathcal{M}_-(\boldsymbol{q},t)\mathcal{M}_+(-\boldsymbol{q})\rangle = 4\mu_B^2 \sum_{\boldsymbol{k}} f_{\boldsymbol{k}\downarrow}\left(1 - f_{\boldsymbol{k}+\boldsymbol{q},\uparrow}\right) \exp \frac{-i\left(\epsilon_{\boldsymbol{k}+\boldsymbol{q},\uparrow} - \epsilon_{\boldsymbol{k}\downarrow}\right)t}{\hbar}. \tag{7.7}$$

Taking the Fourier transform and applying (1.88) at $T = 0$ K leads to the susceptibility,

$$\chi''_{-+}(\boldsymbol{q},\omega)_s = \frac{8\pi\mu_B^2}{V} \sum_{\boldsymbol{k}} f_{\boldsymbol{k}\downarrow}\left(1 - f_{\boldsymbol{k}+\boldsymbol{q},\uparrow}\right) \delta\left(\hbar\omega - \epsilon_{\boldsymbol{k}+\boldsymbol{q},\uparrow} + \epsilon_{\boldsymbol{k}\downarrow}\right). \tag{7.8}$$

The real part of the susceptibility is obtained from the Kramers-Kronig relation (1.65),

$$\chi'_{-+}(\boldsymbol{q},\omega)_s = \frac{8\mu_B^2}{V} \sum_{\boldsymbol{k}} \frac{f_{\boldsymbol{k}\downarrow}\left(1 - f_{\boldsymbol{k}+\boldsymbol{q},\uparrow}\right)}{\epsilon_{\boldsymbol{k}+\boldsymbol{q},\uparrow} - \epsilon_{\boldsymbol{k}\downarrow} - \hbar\omega}. \tag{7.9}$$

Notice that χ_{-+}involves different spin indices whereas χ_{zz}, as given in (3.84), for example, involves the same spin indices. In zero field $\epsilon_{\boldsymbol{k}\uparrow} = \epsilon_{\boldsymbol{k}\downarrow} = \epsilon_{\boldsymbol{k}}$, and in the limit $\omega \to 0$ (7.9) correctly reduces to twice the static result.

The single-particle energies entering the delta function of (7.8) have the value

$$\epsilon_{\boldsymbol{k}+\boldsymbol{q},\uparrow} - \epsilon_{\boldsymbol{k}\uparrow} = 2\mu_B H_0 + \frac{\hbar^2}{m}\boldsymbol{k}\cdot\boldsymbol{q} + \frac{\hbar^2}{2m}q^2. \tag{7.10}$$

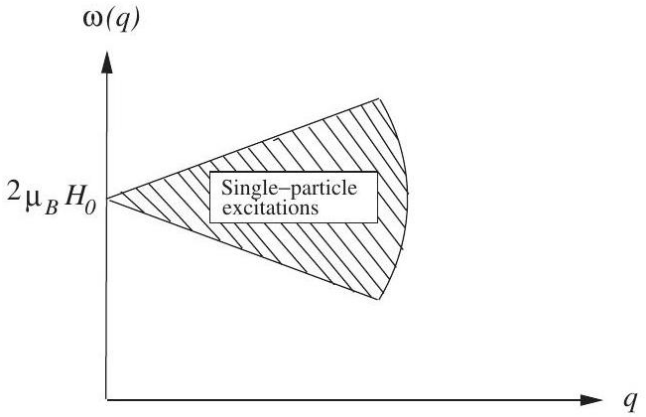


Fig. 7.1. The magnetic-excitation spectrum of a noninteracting electron gas

Therefore the excitation spectrum has the form shown in Fig. 7.1. The $\boldsymbol{q} = 0$ mode is the conduction-electron spin resonance at $\omega = \omega_s = g\mu_B H_0$.

10.8.1 7.1 Paramagnons

Let us now consider how the susceptibility is modified by the introduction of interactions among the electrons. The result for $\chi_{-+}(\boldsymbol{q}, \omega)$ in the presence of the delta-function interaction introduced in (4.81) is readily obtained by the equation-of-motion approach used in that section. In order to calculate $\chi_{-+}(\boldsymbol{q}, \omega)$, as opposed to $\chi_{zz}(\boldsymbol{q}, \omega)$, we must determine the response

$$M_+(\boldsymbol{q}) = -2\mu_B \sum_{\boldsymbol{k}} \left\langle a^\dagger_{\boldsymbol{k}-\boldsymbol{q},\uparrow} a_{\boldsymbol{k}\downarrow} \right\rangle \tag{7.11}$$

to a transverse field $H[\cos(\omega t)\hat{\boldsymbol{x}} + \sin(\omega t)\hat{\boldsymbol{y}}]\cos(\boldsymbol{q}\cdot\boldsymbol{r})$. The corresponding Zeeman interaction is

$$\mathcal{H}_Z = -\frac{1}{4} H \left\{ [\mathcal{M}_-(\boldsymbol{q}) + \mathcal{M}_-(-\boldsymbol{q})]\, e^{i\omega t} + [\mathcal{M}_+(\boldsymbol{q}) + \mathcal{M}_+(-\boldsymbol{q})]\, e^{i\omega t} \right\} \tag{7.12}$$

Adding this interaction to (4.81) the equation of motion for $\left\langle a^\dagger_{\boldsymbol{k}-\boldsymbol{q},\uparrow} a_{\boldsymbol{k}_\downarrow} \right\rangle$ in the random-phase approximation becomes

$$\begin{aligned} i\hbar \frac{d}{dt} \left\langle a^\dagger_{\boldsymbol{k}-\boldsymbol{q},\uparrow} a_{\boldsymbol{k}\downarrow} \right\rangle = {} & (\epsilon_{\boldsymbol{k}} - \epsilon_{\boldsymbol{k}-\boldsymbol{q}}) \left\langle a^\dagger_{\boldsymbol{k}-\boldsymbol{q},\uparrow} a_{\boldsymbol{k}\downarrow} \right\rangle \\ & + \frac{2I}{V} \sum_{\boldsymbol{k}'} \left(n_{\boldsymbol{k}'\uparrow} - n_{\boldsymbol{k}'\downarrow} \right) \left\langle a^\dagger_{\boldsymbol{k}-\boldsymbol{q},\uparrow} a_{\boldsymbol{k}\downarrow} \right\rangle \\ & - \frac{2I}{V} \left(n_{\boldsymbol{k}-\boldsymbol{q},\uparrow} - n_{\boldsymbol{k}\downarrow} \right) \sum_{\boldsymbol{q}'} \left\langle a^\dagger_{\boldsymbol{k}-\boldsymbol{q}'-\boldsymbol{q},\uparrow} a_{\boldsymbol{k}-\boldsymbol{q}',\downarrow} \right\rangle \\ & + \frac{1}{2}\mu_B H \left(n_{\boldsymbol{k}-\boldsymbol{q},\uparrow} - n_{\boldsymbol{k}\downarrow} \right) e^{i\omega t}. \end{aligned} \tag{7.13}$$

We see that the "spin fluctuation" $\left\langle a^\dagger_{\boldsymbol{k}-\boldsymbol{q},\uparrow} a_{\boldsymbol{k}\downarrow} \right\rangle$ is coupled to many other fluctuations through the third term on the right-hand side. If we assume that these spin fluctuations all have a time dependence of the form $\exp(i\omega t)$, then

$$\begin{aligned} \left\langle a^\dagger_{\boldsymbol{k}-\boldsymbol{q},\uparrow} a_{\boldsymbol{k}\downarrow} \right\rangle = {} & \frac{(2I/V)\left(n_{\boldsymbol{k}-\boldsymbol{q},\uparrow} - n_{\boldsymbol{k}\downarrow} \right) \sum_{\boldsymbol{q}'} \left\langle a^\dagger_{\boldsymbol{k}-\boldsymbol{q}'-\boldsymbol{q},\uparrow} a_{\boldsymbol{k}-\boldsymbol{q}',\downarrow} \right\rangle}{\hbar\omega - (\tilde{\epsilon}_{\boldsymbol{k}-\boldsymbol{q},\uparrow} - \tilde{\epsilon}_{\boldsymbol{k}\downarrow})} \\ & - \frac{\frac{1}{2}\mu_B H \left(n_{\boldsymbol{k}-\boldsymbol{q},\uparrow} - n_{\boldsymbol{k}\downarrow} \right)}{\hbar\omega - (\tilde{\epsilon}_{\boldsymbol{k}-\boldsymbol{q},\uparrow} - \tilde{\epsilon}_{\boldsymbol{k}\downarrow})}, \end{aligned} \tag{7.14}$$

where

$$\tilde{\epsilon}_{\underset{\sim}{\boldsymbol{k}\sigma}} = \epsilon_{\boldsymbol{k}} + \frac{I}{2V} \sum_{\boldsymbol{k}'} n_{\boldsymbol{k}'\sigma} \tag{7.15}$$

Since we're interested in the magnetization, (7.11), we must sum (7.14) over $\boldsymbol{k}$. The index $\boldsymbol{q}'$ then becomes a dummy index, which enables us to solve for the magnetization. From definition (1.55) of the susceptibility we obtain for the real part

$$\chi'_{-+}(\boldsymbol{q}, \omega) = \frac{(2\mu_B^2/V)\,\Gamma(\boldsymbol{q}, \omega)}{1 - (2/V)\Gamma(\boldsymbol{q}, \omega)} \tag{7.16}$$

where

$$\Gamma(\boldsymbol{q},\omega)=\sum_{\boldsymbol{k}}\frac{n_{\boldsymbol{k}-\boldsymbol{q},\uparrow}-n_{\boldsymbol{k}\downarrow}}{\hbar\omega-(\tilde{\epsilon}_{\boldsymbol{k}-\boldsymbol{q},\uparrow}-\tilde{\epsilon}_{\boldsymbol{k}\downarrow})} \tag{7.17}$$

In zero field and in the paramagnetic state $\epsilon_{\boldsymbol{k}\uparrow}=\epsilon_{\boldsymbol{k}\downarrow}$. In this case we can evaluate the real and imaginary parts of the susceptibility for simple parabolic bands. The imaginary part is shown in Fig. 7.2. When the interaction $I=0$, we obtain a broad response. However, as I increases toward the Stoner critical value, the response develops a sharp peak at low frequency whose position varies linearly with the wave vector $\boldsymbol{q}$. The presence of this low-frequency peak, or paramagnon, enhances the effective mass m^*. In particular, it is found [141] that the mass contains a contribution which varies as the logarithm of the Stoner enhancement factor.

The above treatment does not include the interactions between the electrons and their host lattice. These interactions are primarily responsible for the resistivity and are characterized by a relaxation time τ. If this relaxation time is much smaller than the inverse cyclotron frequency ω_c, i.e., $\omega_c\tau\ll 1$, then the electron's flight between collisions is approximately a straight line. The electron's overall motion will then be that of a random walk with a diffusion constant $D=\frac{1}{3}v_F^2\tau$. We shall see below that this behavior, while not shifting the conduction-electron spin resonance frequency, has a

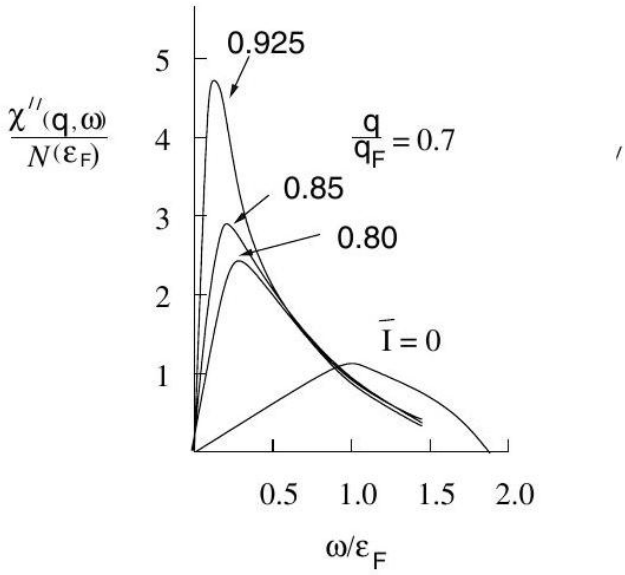


Fig. 7.2. Frequency response of the imaginary part of the susceptibility, showing the paramagnon peak [140]

dramatic effect on the line shape. It is possible to incorporate collision effects into the susceptibility (7.9) in analogy with Mermin's [142] treatment of the Lindhard dielectric function. We shall, however, do this within the framework of Fermi liquid theory.

10.8.2 7.2 Fermi Liquid Theory

As we saw in Sect. 4.2, the important quantity in the Fermi liquid approach is the distribution of quasiparticles $n(\boldsymbol{k},\boldsymbol{\sigma})$. Since we shall eventually be interested in the response to space-varying excitations, let us consider a region of space which is small in comparison with the wavelength of such excitations but large enough to contain many quasiparticles. Then $n(\boldsymbol{k},\boldsymbol{\sigma})$ becomes a function of position $n(\boldsymbol{r},\boldsymbol{k},\boldsymbol{\sigma})$. In Chap. 4 we were able to guess correctly how

the distribution function was changed by a constant velocity displacement or by the presence of a uniform magnetic field. In the case of a time-dependent perturbation, however, it is not obvious what the resulting change will be.

In the absence of any magnetic fields the quasiparticle energy is $\epsilon^0(\boldsymbol{k})\mathbf{1}$. When magnetic fields are applied the energy changes for two reasons. First, there are orbital and spin Zeeman terms. Second, these Zeeman terms alter the distribution of quasiparticles, thereby bringing interaction terms into the energy. If we write the change in the quasiparticle distribution in the general form

$$\delta n(\boldsymbol{r},\boldsymbol{k},\boldsymbol{\sigma}) = n_1(\boldsymbol{r},\boldsymbol{k})\mathbf{1} + \boldsymbol{n}_2(\boldsymbol{r},\boldsymbol{k})\cdot\boldsymbol{\sigma}, \tag{7.18}$$

then the change in the quasiparticle energy may be written as

$$\delta\epsilon(\boldsymbol{r},\boldsymbol{k},\boldsymbol{\sigma}) = \epsilon_1(\boldsymbol{r},\boldsymbol{k})\mathbf{1} + \boldsymbol{\epsilon}_2(\boldsymbol{r},\boldsymbol{k})\cdot\boldsymbol{\sigma}, \tag{7.19}$$

where

$$\epsilon_1(\boldsymbol{r},\boldsymbol{k}) = -\frac{e\hbar}{2m^*c}\boldsymbol{k}\times\boldsymbol{H}_0\cdot\boldsymbol{r} + \frac{2}{V}\sum_{\boldsymbol{k}'}\varphi\left(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{k}}'\right)n_1\left(\boldsymbol{r},\boldsymbol{k}'\right) \tag{7.20}$$

and

$$\boldsymbol{\epsilon}_2(\boldsymbol{r},\boldsymbol{k}) = \mu_B H_0\hat{\boldsymbol{z}} + \mu_B\boldsymbol{H}_1(\boldsymbol{r},t) + \frac{2}{V}\sum_{\boldsymbol{k}'}\psi\left(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{k}}'\right)\boldsymbol{n}_2\left(\boldsymbol{r},\boldsymbol{k}'\right), \tag{7.21}$$

where $\boldsymbol{H}_1(\boldsymbol{r},t)$ is the probing magnetic field and the quantities φ and ψ were defined in (4.66). The first term in $\epsilon_1(\boldsymbol{r},\boldsymbol{k})$ is the orbital Zeeman contribution.

The magnetization at $\boldsymbol{r}$ arising from the quasiparticles in state $\boldsymbol{k}$ is

$$\boldsymbol{m}(\boldsymbol{r},\boldsymbol{k}) = -\frac{\mu_B}{V}\,\mathrm{Tr}_\sigma\{[n(\boldsymbol{r},\boldsymbol{k},\boldsymbol{\sigma})\boldsymbol{\sigma}]\} = -\frac{2\mu_B}{V}\boldsymbol{n}_2(\boldsymbol{r},\boldsymbol{k}). \tag{7.22}$$

To determine the magnetization we must therefore determine the spindependent part of the change in the quasiparticle distribution function. This may be found by solving the equation of motion associated with $n(\boldsymbol{r},\boldsymbol{k},\boldsymbol{\sigma})$.

To set up this equation we ask for the different ways in which $n(\boldsymbol{r},\boldsymbol{k},\boldsymbol{\sigma})$ may change with time. First of all, it might depend explicitly on time. This rate of change has the form $\partial n/\partial t$. Second, quasiparticles may drift into and out of the region in real space defined by $\boldsymbol{r}$. The rate of change of $n(\boldsymbol{r},\boldsymbol{k},\boldsymbol{\sigma})$ associated with this process is

$$\frac{1}{2\hbar}\left(\frac{\partial n}{\partial\boldsymbol{r}}\frac{\partial\epsilon}{\partial\boldsymbol{k}} + \frac{\partial\epsilon}{\partial\boldsymbol{k}}\frac{\partial n}{\partial\boldsymbol{r}}\right) \tag{7.23}$$

Since $\epsilon(\boldsymbol{r},\boldsymbol{k},\boldsymbol{\sigma})$ and $n(\boldsymbol{r},\boldsymbol{k},\boldsymbol{\sigma})$ are matrices, this is written in the symmetrized form. A third way in which the quasiparticle distribution might change is for quasiparticles to change their momentum $\hbar\boldsymbol{k}$. Using Hamiltonian's equation $\partial\boldsymbol{p}/\partial t = -\partial\epsilon/\partial\boldsymbol{r}$, we have for this contribution

$$\frac{1}{2\hbar}\left(\frac{\partial n}{\partial\boldsymbol{k}}\frac{\partial\epsilon}{\partial\boldsymbol{r}} + \frac{\partial\epsilon}{\partial\boldsymbol{r}}\frac{\partial n}{\partial\boldsymbol{k}}\right) \tag{7.24}$$

Another way for $n(\boldsymbol{k},\boldsymbol{r},\boldsymbol{\sigma})$ to change is for the quasiparticle to change its spin state. This change is determined by the commutator $[\epsilon,n]$. Finally, the interaction with the lattice will also produce a change in the distribution. Therefore the total equation of motion becomes

$$\frac{\partial n}{\partial t} + \frac{1}{\hbar}\left\{\frac{\partial n}{\partial\boldsymbol{r}}\frac{\partial\epsilon}{\partial\boldsymbol{k}}\right\} + \frac{1}{\hbar}\left\{\frac{\partial n}{\partial\boldsymbol{k}}\frac{\partial\epsilon}{\partial\boldsymbol{r}}\right\} - \frac{i}{\hbar}[\epsilon,n] = \left.\frac{\partial n}{\partial t}\right|_{\text{collisions}} \tag{7.25}$$

where the braces indicate the symmetrized product. Notice that the second and third terms have been written as a symmetrized Poisson bracket.

The equation of motion for the magnetization is obtained by multiplying (7.25) by $-(\mu_B/V)\,\boldsymbol{\sigma}$ and taking the spin trace. Let us consider the second term. Since the unperturbed distribution function is spatially independent, the spatial derivative becomes

$$\frac{\partial n}{\partial \boldsymbol{r}} = \frac{\partial n_1}{\partial \boldsymbol{r}} + \frac{\partial}{\partial \boldsymbol{r}}\left(\boldsymbol{n}_2 \cdot \boldsymbol{\sigma}\right) \tag{7.26}$$

The derivative of the energy with respect to the wave vector is

$$\frac{\partial \epsilon}{\partial \boldsymbol{k}} = \frac{\partial \epsilon^0}{\partial \boldsymbol{k}} + \frac{\partial \epsilon_1}{\partial \boldsymbol{k}} + \frac{\partial}{\partial \boldsymbol{k}}\left(\boldsymbol{\epsilon}_2 \cdot \boldsymbol{\sigma}\right) = \hbar \boldsymbol{v} + \frac{\partial}{\partial \boldsymbol{k}}\left(\boldsymbol{\epsilon}_2 \cdot \boldsymbol{\sigma}\right) \tag{7.27}$$

Since $\mathrm{Tr}\{\boldsymbol{\sigma}\} = 0, \mathrm{Tr}\{(\boldsymbol{A}\cdot\boldsymbol{\sigma})\boldsymbol{\sigma}\} = 2A$, and $\mathrm{Tr}\{(\boldsymbol{A}\cdot\boldsymbol{\sigma})(\boldsymbol{B}\cdot\boldsymbol{\sigma})\boldsymbol{\sigma}\} = 2i\boldsymbol{A}\times\boldsymbol{B}$, we obtain

$$-\frac{\mu_B}{V}\,\mathrm{Tr}\left\{\frac{\partial n}{\partial \boldsymbol{r}}\frac{\partial \epsilon}{\partial \boldsymbol{k}}\right\}\boldsymbol{\sigma} = -\frac{2\mu_B}{V}\left(\frac{\partial n_1}{\partial \boldsymbol{r}}\frac{\partial}{\partial \boldsymbol{k}}\right)\boldsymbol{\epsilon}_2 + \hbar(\boldsymbol{v}\cdot\nabla)\boldsymbol{m} \tag{7.28}$$

But

$$\frac{\partial n_1}{\partial \boldsymbol{r}} = \frac{\partial n_0}{\partial \epsilon^0}\frac{\partial \epsilon_1}{\partial \boldsymbol{r}} = -\frac{\partial n_0}{\partial \epsilon^0}\left(\frac{e\hbar}{2m^*c}\right)\boldsymbol{k}\times\boldsymbol{H}_0 \tag{7.29}$$

Therefore

$$-\frac{\mu_B}{\hbar V}\,\mathrm{Tr}\left\{\frac{\partial n}{\partial \boldsymbol{r}}\frac{\partial \epsilon}{\partial \boldsymbol{k}}\right\}\boldsymbol{\sigma} = \frac{\mu_B}{V}\frac{\partial n_0}{\partial \epsilon^0}\frac{e}{m^*c}\left[(\boldsymbol{k}\times\boldsymbol{H}_0)\cdot\frac{\partial}{\partial \boldsymbol{k}}\right]\epsilon_2 + \hbar(v\cdot\nabla)\boldsymbol{m} \tag{7.30}$$

The third term in (7.25) is evaluated in the same manner, giving

$$\frac{\mu_B}{\hbar V}\,\mathrm{Tr}\left\{\frac{\partial n}{\partial \boldsymbol{k}}\frac{\partial \epsilon}{\partial \boldsymbol{r}}\right\}\boldsymbol{\sigma} = \frac{2\mu_B}{V}\frac{\partial n_0}{\partial \epsilon^0}(v\cdot\nabla)\epsilon_2 + \frac{e}{2m^*c}\left[(\boldsymbol{k}\times\boldsymbol{H}_0)\cdot\frac{\partial}{\partial \boldsymbol{k}}\right]\boldsymbol{m} \tag{7.31}$$

Notice that the terms in (7.30, 7.31) which do not explicitly involve $\boldsymbol{m}$ contain the factor $\partial n_0/\partial\epsilon^0$. This implies that $\boldsymbol{m}$ is proportional to $\partial n_0/\partial\epsilon^0$. Since $\partial n_0/\partial\epsilon^0$ is essentially nonzero only at the Fermi surface, the magnetization arises only from those quasiparticles at the Fermi surface. This, together with the fact that we are interested in the transverse response of the system, suggests that we introduce the quantity $g(\boldsymbol{r},\boldsymbol{k})$, defined by

$$m_+(\boldsymbol{r},\boldsymbol{k}) = m_x + i m_y = -\frac{\partial n_0}{\partial \epsilon^0} g(\boldsymbol{r},\boldsymbol{k}) \tag{7.32}$$

Since $\boldsymbol{k} \simeq k_F\hat{\boldsymbol{k}}$, the dependence of $g(\boldsymbol{r},\boldsymbol{k})$ upon $\boldsymbol{k}$ involves only the polar angles of $\boldsymbol{k}$ with respect to the z-axis.

The contribution to the equation of motion for $m_+$from the fourth term involving the commutator is

$$i\frac{\mu_B}{\hbar V}\,\mathrm{Tr}\left\{[\epsilon, n]\sigma_+\right\} = i\left[\left(\frac{\mu_B H_0}{\hbar} + \frac{V}{2\hbar\mu_B}G_z\right)m_+ + \frac{V}{2\hbar\mu_B}m_z G - \frac{\mu_B}{\hbar}m_z H_+\right] \tag{7.33}$$

where

$$G(\boldsymbol{r},\boldsymbol{k}) = \frac{2}{V}\sum_{\boldsymbol{k}'}\psi\left(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{k}}'\right)m_+\left(\boldsymbol{r},\boldsymbol{k}'\right) \tag{7.34}$$

and G_z involves m_z in place of m_+. To proceed further with this expression we must determine $m_z(\boldsymbol{r}, \boldsymbol{k})$. In Chap. 5 we found that the magnetization resulting from a uniform field arises from a spherically symmetric shell of states at the Fermi surface. Therefore we take

$$m_z(\boldsymbol{r}, \boldsymbol{k}) = -\frac{\partial n_0}{\partial \epsilon_0} g(\boldsymbol{r})_z \tag{7.35}$$

To find $g(\boldsymbol{r})_z$ we note that the magnetization at $\boldsymbol{r}$, which is given by

$$m_z(\boldsymbol{r}) = \sum_{\boldsymbol{k}} m_z(\boldsymbol{r}, \boldsymbol{k}) \tag{7.36}$$

is also related to the static susceptibility by

$$m_z(\boldsymbol{r}) = \frac{1}{V} \sum_{\boldsymbol{q}} \chi(\boldsymbol{q}) H(\boldsymbol{q}) = \chi(0) H_0. \tag{7.37}$$

We use the fact that

$$\frac{\partial n_0}{\partial \epsilon_0} = -\delta\left(\mu - \epsilon^0(\boldsymbol{k})\right) = -\frac{m^*}{\hbar^2 k_F} \delta\left(k - k_F\right) \tag{7.38}$$

along with our result (4.80) for the static susceptibility, which may be written as

$$\chi(0) = \frac{\mu_B^2 m^* k_F}{\pi^2 \hbar^2 \left(1 + B_0\right)} \tag{7.39}$$

We then obtain

$$m_z(\boldsymbol{r}, \boldsymbol{k}) = -\frac{\partial n_0}{\partial \epsilon^0} \frac{2\mu_B^2 H_0}{V\left(1 + B_0\right)} \tag{7.40}$$

The collision term on the right-hand side of (7.25) is handled by assuming that the distribution function $n(\boldsymbol{r}, \boldsymbol{k}, \boldsymbol{\sigma})$ relaxes to its local average value $\bar{n}(\boldsymbol{r}, \boldsymbol{\sigma})$, defined by

$$\bar{n}(\boldsymbol{r}, \boldsymbol{\sigma}) = \frac{1}{\frac{4}{3}\pi k_F^2} \int d\boldsymbol{k} n(\boldsymbol{r}, \boldsymbol{k}, \boldsymbol{\sigma}) \tag{7.41}$$

If this relaxation is characterized by a time τ, then

$$\left.\frac{\partial g}{\partial t}\right|_{\text{collisions}} = -\frac{1}{\tau}\left(g - \frac{1}{4\pi} \int d\Omega g\right) \tag{7.42}$$

Combining all these results in the equation of motion for $g(\boldsymbol{r}, \boldsymbol{k})$ yields

$$\begin{aligned} &\frac{\partial g}{\partial t} + \left[\boldsymbol{v} \cdot \boldsymbol{\nabla} + \frac{e}{m^* c}\left(\boldsymbol{k} \times \boldsymbol{H}_0\right) \cdot \frac{\partial}{\partial \boldsymbol{k}} + i\Omega_0\right](g + G) \\ &= -\frac{1}{\tau}\left(g - \frac{1}{4\pi} \int d\Omega g\right) + \frac{2\mu_B^2}{V}\left(\boldsymbol{v} \cdot \boldsymbol{\nabla} + i\Omega_0\right) H_+ \end{aligned} \tag{7.43}$$

where

$$\Omega_0 = \frac{2\mu_B H_0}{\hbar\left(1 + B_0\right)} \tag{7.44}$$

Equation (7.43) is a linear integral equation that is difficult to solve in general. Therefore certain simplifying approximations are made. For example, Silin has considered the spatially uniform normal modes of a collisionless system [143]. This corresponds to solving (7.43) without

the $\boldsymbol{v} \cdot \boldsymbol{\nabla}$ term on the left and with the right side set to 0 . The resulting solutions for $g(\theta, \varphi)$ are spherical harmonics, $Y_n^m(\theta, \varphi)$, with the corresponding eigenfrequencies.

$$\omega_{nm} = (\Omega_0 + m\omega_c)\left(1 + \frac{B_n}{2n+1}\right), \tag{7.45}$$

where $\omega_c = eH_0/m^*c$ is the cyclotron frequency. The mode $\omega_{00} \equiv \omega_s \equiv 2\mu_B H_0/\hbar$ is the conduction-electron spin resonance. We see that as a result of the interactions there are also modes involving combined spin and orbital excitations.

Platzman and Wolff have investigated the $\boldsymbol{q}$ dependence of the ω_{00} mode [144]. However, they also retained the collision term characterized by a relaxation time τ. Their result for the susceptibility, including also a spin-relaxation term T_2 is [1]

$$\chi(\boldsymbol{q}, \omega) = \frac{m^*}{m} \frac{\chi_{\text{Pauli}}}{1 + B_0} \frac{\omega_s}{\omega_s - \omega - iD^* q^2 - i/\bar{T}_2} \tag{7.46}$$

where

$$\begin{aligned} iD^* =& \frac{1}{3} v_F^2 \left(1 + B_0\right)\left(1 + B_1\right)\left(\bar{\omega}_s - \bar{\Omega}_0\right) \\ & \times \left[\frac{\sin^2 \Delta}{\bar{\omega}_c^2 - \left(\bar{\omega}_s - \bar{\Omega}_0\right)^2} - \frac{\cos^2 \Delta}{\left(\bar{\omega}_s - \bar{\Omega}_0\right)^2}\right]. \end{aligned} \tag{7.47}$$

Here Δ is the angle that $\boldsymbol{q}$ makes with the z axis, and

$$\bar{\omega}_s = \omega_s + \frac{i}{\tau}\left(1 + B_1\right) \tag{7.48}$$

$$\bar{\omega}_c = \frac{eH_0}{m^* c}\left(1 + B_1\right) \tag{7.49}$$

$$\bar{\Omega}_0 = \Omega_0\left(1 + B_1\right) \tag{7.50}$$

$$\bar{T}_2 = T_2\left(1 + B_0\right)^{-1}$$

Notice that as ω and $q \to 0$ and $\tau \to \infty$, (7.46) reduces to the result we obtained in Chap. 4.

10.8.3 7.3 Conduction-Electron Spin Resonance

An interesting limit is when there are no quasiparticle interactions, that is, $B_0 = B_1 = 0$. In this case $\bar{\omega}_s - \bar{\Omega}_0 = \frac{i}{\tau}$ and D^* is a real number. The susceptibility (7.46) then contains a pole at the conduction-electron spin resonance (CESR) frequence ω_s. The fact that the lifetime of this mode is partially governed by the diffusion term has an interesting consequence for the resonance spectrum. The point is that in a metal the electromagnetic field that excites the resonance is not uniform across the sample. To find the field distribution we must solve Maxwell's equations with the appropriate boundary conditions.

Let us consider the geometry shown in Fig. 7.3 where the electromagnetic wave is incident from the left on a metal with conductivity σ. Assuming the

[1] Feher and Kip [145] denote the spin relaxation time as T_1. In a metal the rapid orbital relaxation (τ) results in $T_1 \approx T_2$.

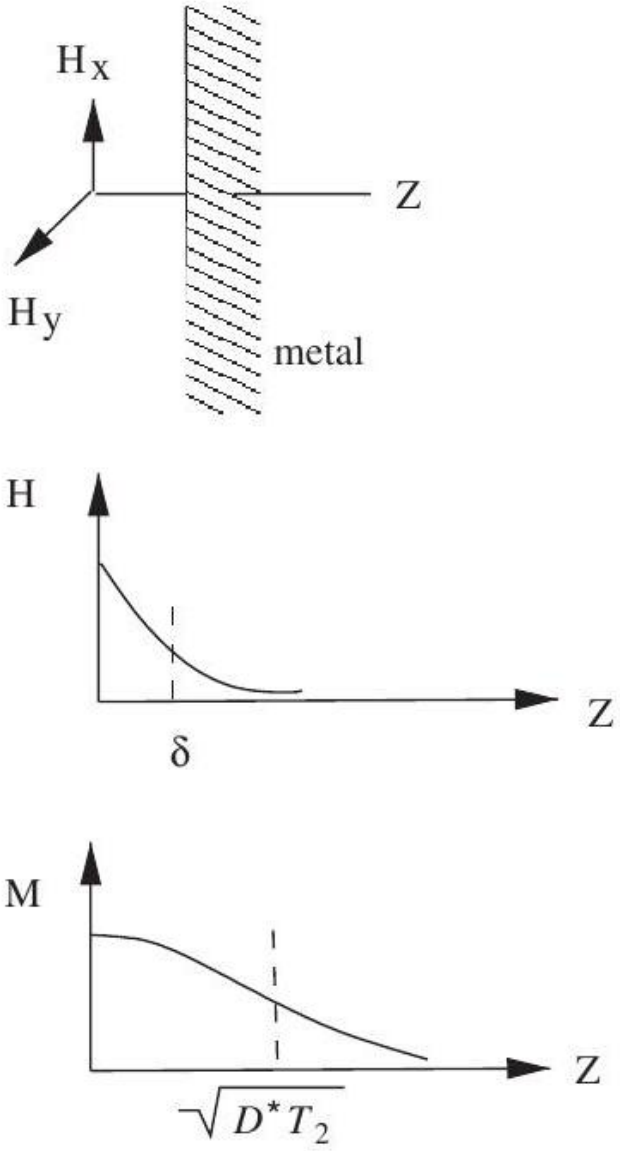


Fig. 7.3. Geometry and field distributions associated with conduction-electron spin resonance

fields vary only along the z-direction, the two Maxwell equations, together with $\boldsymbol{j} = \sigma \boldsymbol{E}$, reduce to

$$\nabla \times \boldsymbol{E} = \frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t} \begin{cases} \frac{\partial E_x}{\partial z} = -\frac{1}{c}\left(\frac{\partial H_y}{\partial t} + 4\pi \frac{\partial M_y}{\partial t}\right) \\ \frac{\partial E_y}{\partial z} = \frac{1}{c}\left(\frac{\partial H_x}{\partial t} + 4\pi \frac{\partial M_z}{\partial t}\right) \end{cases} \tag{7.51}$$

$$\nabla \times \boldsymbol{H} = \frac{4\pi}{c} j \begin{cases} \frac{\partial E_x}{\partial z} = \frac{4\pi}{c} j_y = \frac{4\pi}{c}\sigma E_y \\ \frac{\partial H_y}{\partial z} = -\frac{4\pi}{c} j_x = -\frac{4\pi}{c}\sigma E_x \end{cases} \tag{7.52}$$

Introducing the circular components $M_+ = M_x + iM_y, H_+ = H_x + iH_y$, and assuming solutions of the form

$$\begin{aligned} M_+ &= m e^{i\omega t - \kappa z} \\ H_+ &= h e^{i\omega t - \kappa z} \end{aligned} \tag{7.53}$$

we obtain

$$\frac{m}{h} = \chi = -\frac{1}{4\pi}\left(1 + \frac{1}{2} i\delta^2\kappa^2\right) \tag{7.54}$$

where $\delta = c/\sqrt{2\pi\sigma\omega}$ is the skin depth. Self-consistency between the susceptibility (7.46) derived from the equation of motion for the magnetization and that derived above from Maxwell's equations gives a quadratic equation for the parameter κ. This has two roots, one of the order of δ, the other of the order of $\sqrt{D^*T_2}$. The latter characterizes how far a spin diffuses before relaxing. The amplitudes of the fields associated with these solutions are determined from the boundary conditions. The results are sketched in Fig. 7.3. The magnetic field falls off in a distance of the order of the skin depth. The magnetization, however, penetrates a distance $\sqrt{D^*T_2}$ which may be much larger than δ.

In the standard reflection-type resonance experiment the metal is placed at one end of a microwave cavity. The power absorbed by the cavity, which is easily calculated from the above results, is shown in Fig. 7.4.

We immediately notice the asymmetric nature of the absorption. This is associated with the fact that although an electron may only spend a time of the order of $T_D = \delta^2/2D^*$ in the skin depth, it returns to the skin depth many times before its spin "memory" is destroyed. Thus it experiences a set of microwave pulses whose intervals are random but whose phases are coherent. This can lead to destructive as well as constructive interference with the result that the power absorbed at some values of the dc field may be less than that ordinarily absorbed in zero field. Since Dyson [145] was the first to derive this characteristic shape, such lines are referred to as Dysonian. The asymmetry may be quantified by the ratio A/B which depends upon the ratio of T_D to

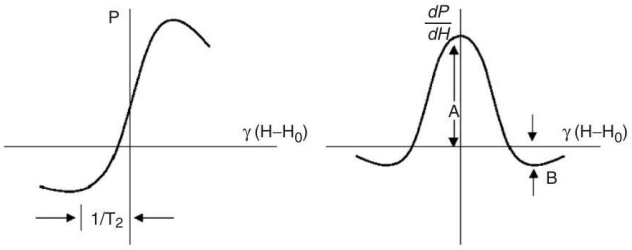


Fig. 7.4. Power absorption (a) and its derivative (b) due to electron spin resonance in a metal for a ratio of $T_D / T_2 \simeq 0.1$

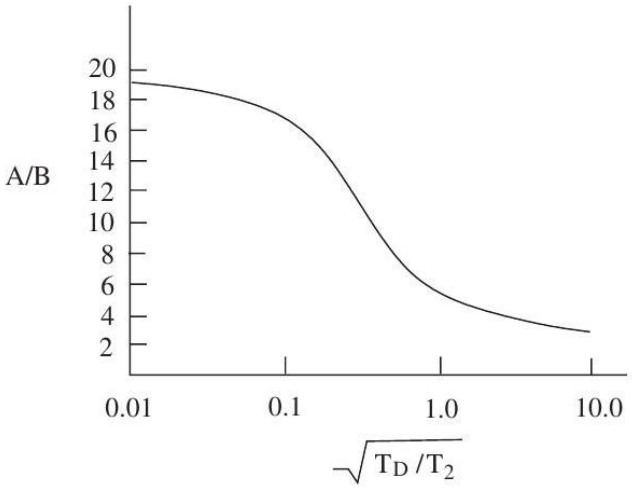


Fig. 7.5. Dependence of the asymmetry ratio A/B on the ratio of the spin diffusion time T_D to the spin relaxation time T_2

T_2 as shown in Fig. 7.5. The fact that the width of the line is proporitional to $1/T_2$ is also due to the fact that the electron returns to the skin depth many times.

The fact that the magnetization penetrates much farther into the metal than the driving magnetic field is the basis of the transmission technique. In this technique the sample takes the form of a thin film whose thickness is large compared with δ but smaller than $\sqrt{D^* T_2}$. This film is placed between two microwave cavities. Off resonance no power reaches the second cavity. At resonance, however, the precessing spins diffuse across the film and radiate into the second cavity. The intensity of the magnetization excited depends upon the susceptibility within the skin depth. Thus it has been found that the sensitivity of the transmission technique can be greatly enhanced by depositing a thin ferromagnetic film on the front surface of the film being studied.

10.8.4 7.4 Spin Waves

Let us now return to the susceptibility (7.46) and consider the effect of the electron-electron interactions. In particular, if

$$\left| \frac{(B_0 - B_1)\,\omega_s \tau}{1 + B_0} \right| \gg 1$$

then $\tilde{\omega}_s - \bar{\Omega}_0$ becomes real and D^* becomes pure imaginary and modes appear at $\omega = \omega_s - D^* k^2$. Thus, the interactions convert the transport of the magnetization from diffusive to propagating. These modes are referred to as spin waves, although they bear only a qualitative resemblance to the spin waves we shall discuss in the next chapter. In a slab of thickness L the wave vector takes on values which are multiples of π/L. The modes then appear as sidebands of the main conduction-spin resonance line. Such spinwave sidebands have been observed in various metals [146]. The transmission spectrum for sodium is shown in Fig. 7.6. From careful measurements of the angular dependence of the spin-wave positions the Fermi liquid parameters can be deduced. For sodium $B_0 = -0.18$ and $B_1 = 0.05$.

Notice that D^* may be positive or negative, depending on the direction of $\boldsymbol{q}$ with respect to the dc field. In the next chapter we shall find that the spinwave spectrum in a ferromagnetic insulator always curves upward. The fact that the spectrum may curve downward in the metal

is due to the "repulsion" of these spin-wave modes by the higher-lying orbital modes, whose $\boldsymbol{q} = 0$ frequencies we have derived in (7.45).

Since the Platzmann-Wolff solution (7.46) is valid only for small $\boldsymbol{q}$ near ω_s, it does not include the single-particle excitations illustrated in Fig. 7.1. In the presence of interactions these modes are shifted upward by a self-energy that is proportional to $N_\downarrow - N_\uparrow$. Therefore the spectrum has the form shown in Fig. 7.7. When the spin waves enter the single-particle band, they can decay into single-particle excitations. In this region they become critically damped.

10.8.5 7.5 Local Moments in Metals

It is interesting to consider the case when a "potentially" magnetic ion is dissolved in a metal. We say "potentially" because in some cases the magnetic response of the impurity is destroyed by the conduction electrons, a situation referred to as the Kondo effect which was discussed in Chap. 3. For our present purposes we shall assume that the local moment exists and that its coupling to the conduction electrons may be described by the exchange interaction

$$\mathcal{H} = -J\boldsymbol{S} \cdot s \tag{7.55}$$

As a result of this interaction the magnetization associated with the local impurity will relax with a rate

$$\frac{1}{T_{ie}} = \frac{\pi}{\hbar} J^2 N \left(\epsilon_F\right)^2 k_B T \tag{7.56}$$

(This result is analogous to the relaxation of nuclear spins in a metal through their hyperfine coupling with conduction electrons. In that case J is replaced

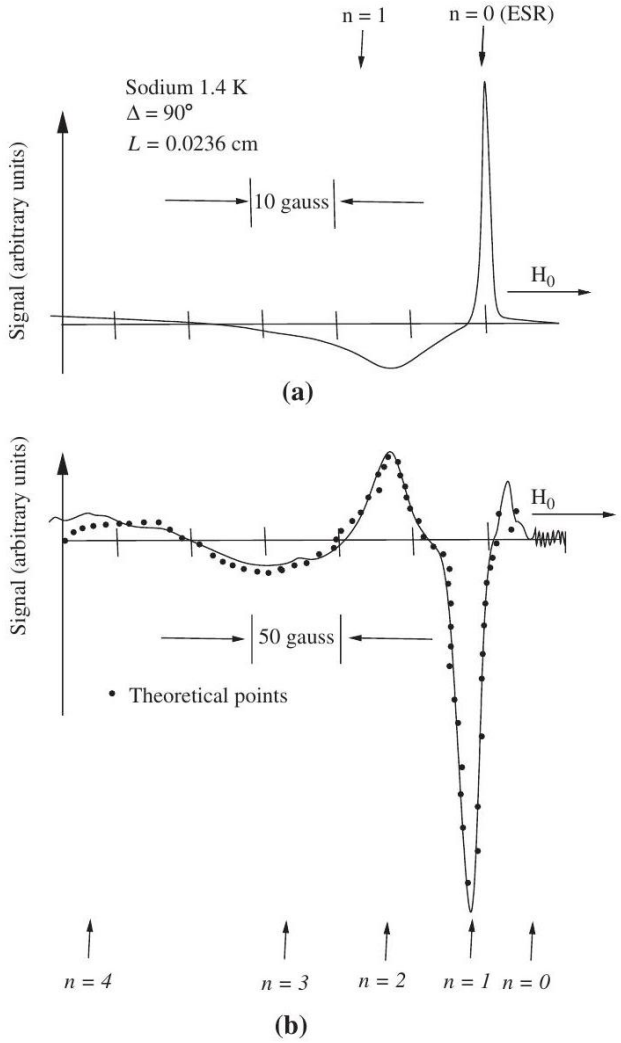


Fig. 7.6. Typical spin wave signals as a function of applied *dc* magnetic field ($H_{0 \approx 3,250 G$). (a) The $n = 0$ mode (usual conduction-electron spin resonance) and the $n = 1$ mode are clearly shown. (b) The gain and field sweep have been increased to display the first four spinwave modes beyond the CESR. The theoretical points have been obtained from the susceptibility in (7.46) [146]}

by the hyperfine constant A, and we refer to the result as the Korringa relaxation rate.) The impurity spins also relax the conduction-electron magnetization with the rate

$$\frac{1}{T_{ei}} = \frac{2\pi}{\hbar} J^2 N\left(\epsilon_F\right) \frac{S(S+1)}{3} c, \tag{7.57}$$

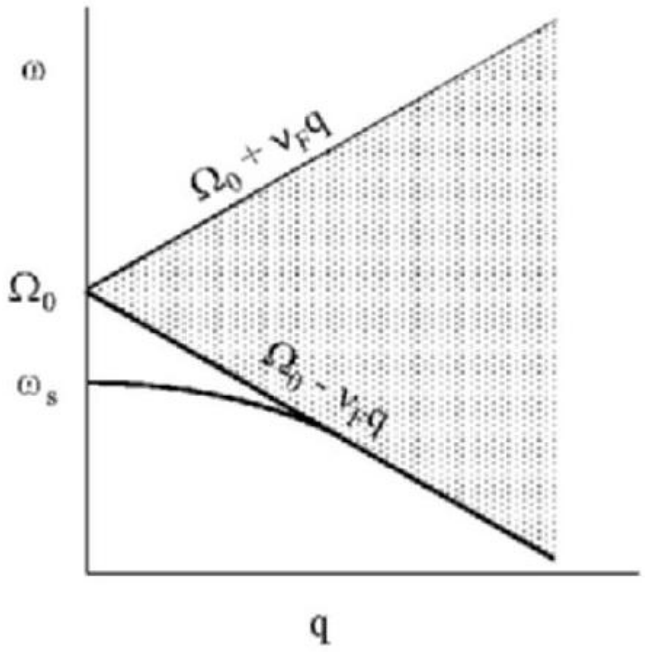


Fig. 7.7. Magnetic excitation spectrum of an interacting electron system

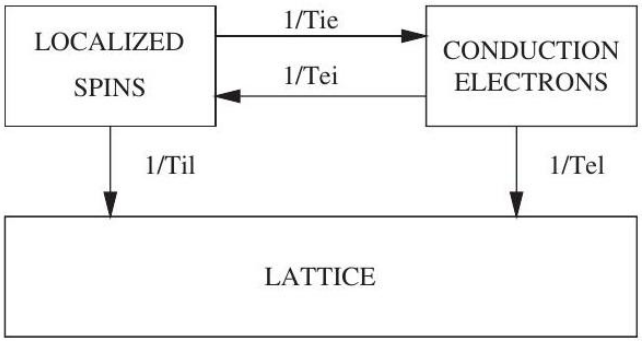


Fig. 7.8. Schematic representation of the energy resevoirs associated with a localized spin impurity in a metal

where c is the impurity concentration. Notice that

$$\frac{T_{ie}}{T_{ei}} = \frac{\chi_i}{\chi_e}, \tag{7.58}$$

where χ_i is the Curie susceptibility and χ_e the Pauli susceptibility. This relation is actually very general, being a consequence of the principle of detailed balance. Considering that both spin systems can also relax directly to the lattice we have the situation illustrated in Fig. 7.8. The magnetizations are described by a pair of coupled Bloch-like equations [147],

$$\begin{aligned}
\frac{d\boldsymbol{M}_e}{dt} &= \gamma_e \boldsymbol{M}_e \times (\boldsymbol{H} + \lambda \boldsymbol{M}_i) - \frac{\boldsymbol{M}_e}{T_{ei}} + \frac{\boldsymbol{M}_i}{T_{ie}} - \frac{\left(\boldsymbol{M}_e - \boldsymbol{M}_e^0\right)}{T_{el}} \\
\frac{d\boldsymbol{M}_i}{dt} &= \gamma_i \boldsymbol{M}_i \times (\boldsymbol{H} + \lambda \boldsymbol{M}_e) - \frac{\boldsymbol{M}_i}{T_{ie}} + \frac{\boldsymbol{M}_e}{T_{ei}} - \frac{\left(\boldsymbol{M}_i - \boldsymbol{M}_i^0\right)}{T_{il}}
\end{aligned} \tag{7.59}$$

Here λ is an effective field parameter which is proportional to the exchange J. For transition-metal impurities, such as Mn in Cu the relaxation rates $1/T_{ei}$ and $1/T_{ie}$ are much larger than

the relaxation rate to the lattice. This
corresponds to a "bottleneck" condition. If the coupling J is strong, then the two magnetizations remain parallel to one another and the transverse components of (7.59) become

$$\frac{d}{dt}\begin{pmatrix} M_e^+ \\ M_i^+ \end{pmatrix} = \begin{pmatrix} -i\gamma_e H - \frac{1}{T_{ei}} & \frac{1}{T_{ie}} \\ \frac{1}{T_{ei}} & -i\gamma_i H - \frac{1}{T_{ie}} \end{pmatrix} \begin{pmatrix} M_e^+ \\ M_i^+ \end{pmatrix}. \tag{7.60}$$

Diagonalizing this equation we find that the effective g value for the strongly coupled case is

$$g_{\text{eff}} = \frac{g_i\chi_i + g_e\chi_e}{\chi_i + \chi_e}. \tag{7.61}$$

At low impurity concentrations or high temperatures χ_i is small and the observed g value should be closed to that of the pure metal. As one increases the impurity concentration or goes to lower temperatures, g_{eff} should shift towards a value characteristic of the impurity. The transmission electron spin resonance data [148] on Cu Mn shown in Fig. 7.9 has these features.

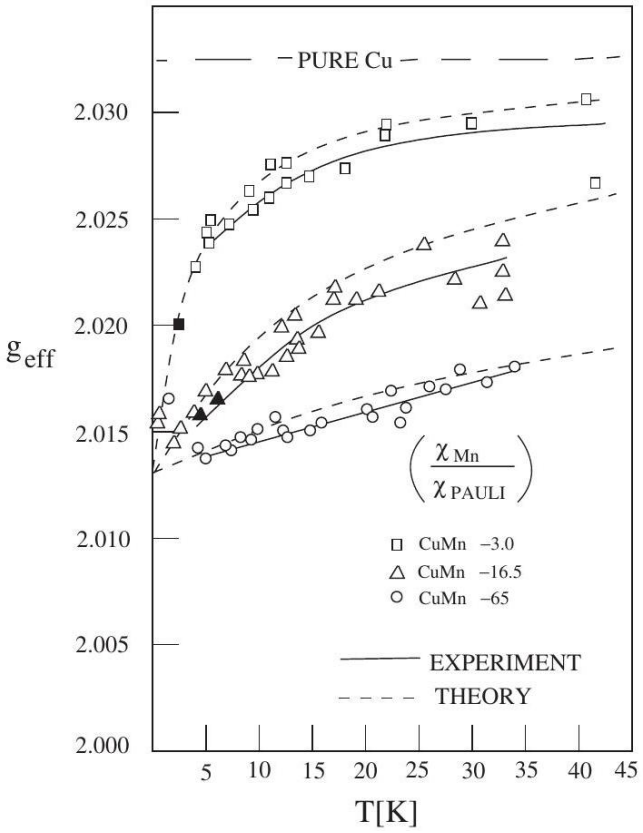


Fig. 7.9. Plot of the g value obtained from transmission resonance versus temperature for three $C u \mathrm{Mn
$ samples. The theoretical curves were obtained from (7.61) using $g_{\text{Cu}} = 2.033$ and $g_{\text{Mn}} = 2.013$ and the susceptibility ratios shown. A susceptibility ratio of 3 corresponds to 13 ppm of Mn in Cu [148]}

10.8.6 7.6 Faraday Effect

The Faraday effect refers to the rotation of the plane of polarization of a linearly polarized electromagnetic wave as it propagates through a magnetic or magnetized material. Assuming for the moment that the propagation through such a medium is adequately characterized by

the susceptibility (or permeability) tensor, Maxwell's equations tell us that the rotation per unit length of the medium is given by

$$\frac{\Delta\theta}{\Delta z} = \frac{2\pi\omega}{\eta}\chi''_{xy}$$

In the case of a Kramers ground state we found that when $\omega \gg \omega_0, \chi''_{xy}$ decreases as $1/\omega$ and $\Delta\theta/\Delta z$ reduces to

$$\frac{\Delta\theta}{\Delta z} = \frac{\omega_M}{2c/n}$$

where n is the real part of the complex index of refraction in the absence of a static magnetization and $\omega_M = 4\pi\gamma M$. For a magnetic material with a saturation magnetization of $4\pi M = 2000$G and a dielectric constant of 13, this gives a rotation of approximately 140°/cm. Such rotation forms the basis for various nonreciprocal microwave devices.

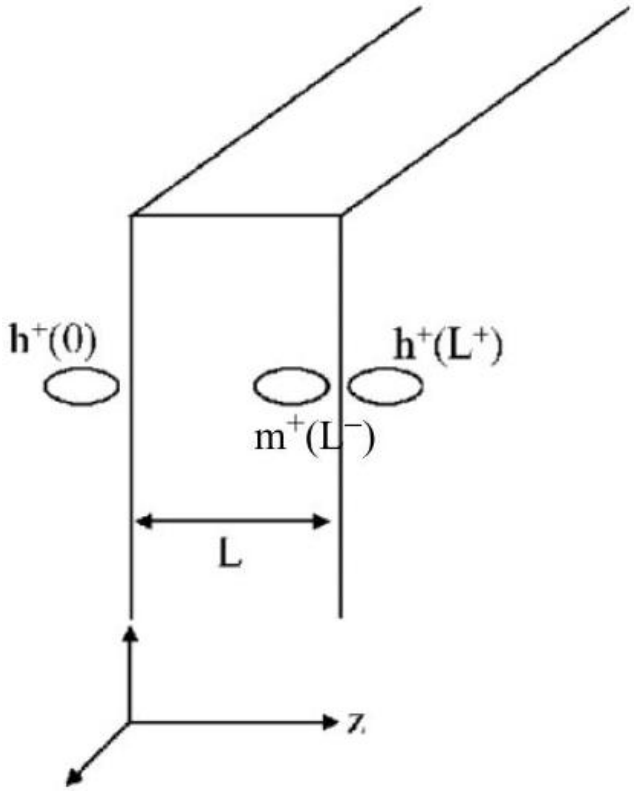
As one goes beyond microwave frequencies, one must also consider the time-varying electric field which accompanies the "probing" magnetic field and brings the dielectric, or conductivity, tensor into play. The Faraday rotation associated with this "polarization" response is

$$\frac{\Delta\theta}{\Delta z} = -\frac{2\pi}{cn}\sigma'_{xy} = \frac{2\pi\omega}{cn}\varepsilon''_{xy}$$

Symmetry arguments, similar to those used in deriving the Onsager relation (1.89), tell us that ε''_{xy} is a linear function of M_z. Physically, this dependence of the dielectric function on the magnetization arises through the spin-orbit interaction. The important point, however, is that the polarization-induced rotation increases with frequency while the magnetic contribution is independent of frequency. Thus there will be a frequency beyond which the polarization contribution dominates. This typically occurs in the near infrared. For this reason, one generally finds magnetooptical phenomena described entirely in terms of the conductivity. Pershan [149], in fact, has argued that so long as the electromagnetic wavelength is less that the sample dimensions, one can always describe the wave properties in terms of an effective dielectric function, setting the magnetic susceptibility to zero. However, as Pershan has also noted, one must pay careful attention to the boundary conditions when applying such a decription. On the other hand, there may be situations in which the microscopic behavior of the Faraday rotation may be more conveniently understood, or calculated, in terms of magnetic dipole matrix elements.

Problems

7.1. The transmission resonances described in this chapter offer a good example of the nonlocal susceptibility introduced in Sect. 1.3. Suppose we wish to calculate the magnetic field at the back face of a metal film excited by a field at the front face (see figure on page 236). It can be shown that the magnetic field just outside the back face is proportional to the magnetization at the back surface, i.e., $h^+(L^+) \sim m^+(L^-)$. The proportionality involves the surface impedance. Use (1.58) and (1.59) to determine $m_+(L^-, \omega)$. Assume $\chi(\boldsymbol{q}, \omega)$ has the form (7.46) in the noninteracting limit, i.e., $B_n = 0$, and that $h^+(z', \omega)$ is a delta function i.e., $h^+(z', \omega) \sim \delta(z')$. Since $q_z = n\pi/L$ the sum in (1.59) can be written as a sum over n. Using the identity



$$\sum_{n=-\infty}^{\infty} \frac{(-1)^n}{A^2 - \pi^2 n^2} = \frac{\cot A/2 - \cot A}{A}$$

obtain an expression for $m^+(L^-, \omega)$ and, hence, $h^+(L^+, \omega)$. Plot $h^+(L^+, \omega)$ versus $(\omega_s - \omega)/\omega_s$ for $\sqrt{DT_2}/L = 6$ and 1.4.
7.2. Using the Fermi "Golden Rule" calculate the relaxation rates (7.56) and (7.57).

10.9 The Dynamic Susceptibility of Strongly Interacting Systems

In the previous chapter we have seen that in nonferromagnetic metals the presence of interactions leads to magnetic excitations of a wavelike nature called spin waves. In this chapter we shall investigate these excitations in magnetically ordered systems.

10.9.1 8.1 Broken Symmetry

At one time spin waves were only thought to be associated with local moment systems. But they have a far more fundamental basis. In the paramagnetic state the system is rotationally invariant. In the ferromagnetically ordered state this symmetry is broken-there is a uniform magnetization $\boldsymbol{M}$ in a particular direction. In principle, this direction is arbitrary, but in practice it is determined by anisotropy. In Chap. 4 we argued that the energy associated with a gradual variation in the magnetization has the form

$$\frac{\delta E}{V} = \frac{A}{M_0^2} |\nabla \boldsymbol{M}|^2$$

This energy implies that the magnetization in any macroscopically small region is acted on by an exchange torque coming from an effective field

$$H_{\text{eff}} = -\frac{2A}{M_0^2} \nabla^2 \boldsymbol{M}$$

M responds to this effective field just as it does to a real magnetic field. Therefore,

$$\frac{d\boldsymbol{M}}{dt} = -\gamma \boldsymbol{M} \times \boldsymbol{H}_{\text{eff}} \tag{8.1}$$

Thus, a transverse fluctuation in the magnetization with a wave vector $\boldsymbol{q}$ has a frequency

$$\omega_{\boldsymbol{q}} = \gamma \frac{2A}{M_0} q^2.$$

Such modes are a direct consequence of having broken the continuous rotational symmetry and are referred to as "Goldstone modes". A system characterized by an Ising Hamiltonian does not possess such a continuous symmetry and, indeed, there are no spin-wave-like excitations in the ordered state. The behavior of these modes at shorter wavelengths and the form of the exchange stiffness parameter depend upon the microscopic details which we shall consider in this chapter.

10.9.2 8.2 Insulators

Let us begin by considering a lattice of spins whose interactions may be described by the Heisenberg exchange interaction (2.89). Suppose we apply a uniform static field which serves to define a z-axis. We now wish to determine how this system responds to the time- and space-dependent field $\boldsymbol{H}_1 \cos(\boldsymbol{q} \cdot \boldsymbol{r}) \cos \omega t$. If this field is in the x direction, the total Hamiltonian becomes

$$\mathcal{H} = -\sum_i \sum_{j \neq i} J_{ij} S_i \cdot S_j + g\mu_B H_0 \sum_i S_i^z + g\mu_B H_1 \sum_i S_i^x \cos\left(\boldsymbol{q} \cdot \boldsymbol{R}_i\right) \cos \omega t. \tag{8.2}$$

Introducing the Fourier components defined by (4.4), we obtain

$$\begin{aligned} \mathcal{H} = &-\sum_{\boldsymbol{q}'} J\left(-\boldsymbol{q}'\right) S\left(\boldsymbol{q}'\right) \cdot S\left(-\boldsymbol{q}'\right) + g\mu_B H_0 S_z(0) + \frac{1}{2} g\mu_B H_1 \\ &\times \left[S_x(\boldsymbol{q}) + S_x(-\boldsymbol{q})\right] \cos \omega t. \end{aligned} \tag{8.3}$$

Since this Hamiltonian is space dependent as well as time dependent, the appropriate equation of motion, analogous to (5.4), is

$$\frac{d\boldsymbol{M}(\boldsymbol{q})}{dt} = -\frac{i}{\hbar V} \langle [\mathcal{M}(\boldsymbol{q}), \mathcal{H}] \rangle. \tag{8.4}$$

Since we are expressing the spin in units of $\hbar$, the commutation relations have the form

$$\left[S_x(\boldsymbol{q}), S_y\left(\boldsymbol{q}'\right)\right] = i S_z\left(\boldsymbol{q} + \boldsymbol{q}'\right). \tag{8.5}$$

The commutator involving the x component of the spin with the Hamiltonian then becomes

$$\begin{aligned} \left[S_x(\boldsymbol{q}), \mathcal{H}\right] = &-i \sum_{\boldsymbol{q}'} J\left(\boldsymbol{q}'\right) \left[S_z\left(\boldsymbol{q} + \boldsymbol{q}'\right) S_y\left(-\boldsymbol{q}'\right) + S_y\left(\boldsymbol{q}'\right) S_z\left(\boldsymbol{q} - \boldsymbol{q}'\right)\right. \\ &\left. - S_y\left(\boldsymbol{q} + \boldsymbol{q}'\right) S_z\left(-\boldsymbol{q}'\right) - S_z\left(\boldsymbol{q}'\right) S_y\left(\boldsymbol{q} - \boldsymbol{q}'\right)\right] - i g\mu_B H_0 S_y(\boldsymbol{q}). \end{aligned} \tag{8.6}$$

In the random-phase approximation each Fourier component is assumed to be independent. This implies that $\left\langle S_z\left(\boldsymbol{q} + \boldsymbol{q}'\right) S_y\left(-\boldsymbol{q}'\right)\right\rangle = \left\langle S_z\left(\boldsymbol{q} + \boldsymbol{q}'\right)\right\rangle \left\langle S_y\left(-\boldsymbol{q}'\right)\right\rangle$. Furthermore, in the low-temperature region we may linearize (8.4) by making the approximation

$$\langle S_z(\boldsymbol{q}')\rangle = NS\Delta(\boldsymbol{q}'). \tag{8.7}$$

With these approximations, (8.4) becomes

$$\frac{dM_x(\boldsymbol{q})}{dt} = -\left\{\gamma H_0 + \frac{2NS}{\hbar}[J(0) - J(\boldsymbol{q})]\right\} M_y(\boldsymbol{q}), \tag{8.8}$$

where $\gamma = g\mu_b/\hbar$ is the gyromagnetic ratio, which we take as positive for electrons. This equation is analogous to (6.7a), and the equation for $M_y(\boldsymbol{q})$ is analogous to (6.7b). Therefore all the arguments leading to the susceptibility (6.18) also apply in this case. Thus we find that the imaginary part of the susceptibility of a ferromagnet at low temperatures has the form

$$\chi''_{xx}(\boldsymbol{q},\omega) = \frac{\pi g^2\mu_B^2 NS}{2\hbar}\{\delta[\omega - \omega(\boldsymbol{q})] - \delta[\omega + \omega(\boldsymbol{q})]\}, \tag{8.9}$$

where

$$\omega(\boldsymbol{q}) = \gamma H_0 + \frac{2NS}{\hbar}[J(0) - J(\boldsymbol{q})]. \tag{8.10}$$

These are the spin-wave modes. Classically, these excitations correspond to a precession of the spin system in which the phase of one spin relative to another is determined by the wave vector $\boldsymbol{q}$. An example of these modes is shown in Fig. 8.1. The $\boldsymbol{q} = 0$, or uniform-precession, mode corresponds to ferromagnetic resonance. Quantum mechanically, a spin wave consists of a single spin flip propagating throughout the lattice. Since the spins are quantized, these excitations are also quantized. Consequently, they are often referred to as magnons. Because these modes are collective, in the sense that they involve

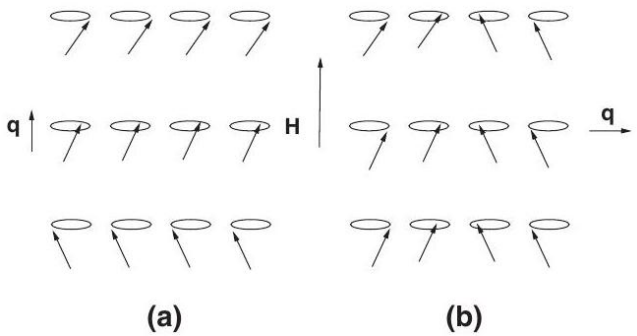


Fig. 8.1. Schematic representations of a spin wave (a) propagating parallel to the applied magnetic field, and (b) propagating perpendicular to this field

all the spins in the lattice, the excitation energy varies continuously from $\gamma\hbar H_0$ up to the exchange energy. If the only excitations were single-particle excitations, the energy required to produce a spin deviation would be of the order of the exchange energy.

For a simple cubic lattice with lattice parameter a, and a nearest-neighbor exchange J, the long wavelength limit of $J(0) - J(\boldsymbol{q})$ is Ja^2q^2/N. Therefore the exchange stiffness introduced at the beginning of this chapter is

$$A = S^2a^2J/v,$$

where v is the atomic volume, V/N.

Combining (4.16), (8.10), we see that the ratio of the maximum spinwave frequency to $k_BT_C/\hbar$ for a simple cubic lattice is $6/(S+1)$. Therefore the high-temperature approximation used in writing (4.48) is valid only for $T/T_C > 6/(S+1)$.

8.2.1 Spin-Wave Theory

Our low-temperature approximation (8.7) has led to a set of modes that, it turns out, behave just like a system of independent harmonic oscillators. In 1940 Holstein and Primakoff introduced a very useful technique which formalizes this harmonic-oscillator behavior. Their approach was based on the expansion of spin operators in terms of the creation and annihilation operators of the harmonic oscillator. For example, consider the effect of S_i^- acting on the spin state $|S, M_S\rangle$

$$S_i^- |S, M_S\rangle = \sqrt{(S + M_S)(S - M_S + 1)}\, |S, M_S - 1\rangle . \tag{8.11}$$

If we introduce the spin-deviation operator $\hat{n}_i = S - S_i^z$, with the quantum numbers $n_i = S - M_S$, this relation becomes

$$S_i^- |n_i\rangle = \sqrt{2S}\sqrt{n_i + 1}\sqrt{1 - \frac{n_i}{2S}}\, |n_i + 1\rangle . \tag{8.12}$$

The harmonic-oscillator operator $a^\dagger$ has the property (as a consequence of its commutation relations) that

$$a^\dagger |n\rangle = \sqrt{n+1}|n+1\rangle. \tag{8.13}$$

Therefore Holstein and Primakoff made the associations

$$\begin{aligned} S_i^- &= \sqrt{2S} a_i^\dagger f_i(S), \\ S_i^+ &= \sqrt{2S} f_i(S) a_i, \\ S_i^z &= -S + a_i^\dagger a_i, \end{aligned} \tag{8.14}$$

where $f_i(S) \equiv \sqrt{1 - \left(a_i^\dagger a_i / 2S\right)}$. Notice that this factor makes the transformation nonlinear. The point is that the eigenvalues of a spin system consist of a finite number $(2S+1)$ of equally spaced values while those of the harmonic oscillator consist of an infinite number of equally spaced levels. It is an easy matter to show that the representations (8.14) satisfy the required commutation relations.

As we have seen, in dealing with coupled systems it is convenient to introduce Fourier components. This suggests that we expand the spin-deviation operators according to

$$a_i = \frac{1}{\sqrt{N}} \sum_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i} \boldsymbol{a}_{\boldsymbol{k}} \tag{8.15}$$

$$a_i^\dagger = \frac{1}{\sqrt{N}} \sum_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i} \boldsymbol{a}_{\boldsymbol{k}}^\dagger \tag{8.16}$$

We shall refer to $a_{\boldsymbol{k}}$ and $a_{\boldsymbol{k}}^\dagger$ as magnon annihilation and magnon creation operators, respectively. The spin operators may be written in terms of these operators as

$$S_i^+ = \sqrt{\frac{2S}{N}} \sum_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot R_i} a_{\boldsymbol{k}} - \frac{1}{N\sqrt{8NS}} \sum_{\boldsymbol{k},\boldsymbol{k}',\boldsymbol{k}''} e^{-i(\boldsymbol{k}+\boldsymbol{k}'-\boldsymbol{k}'')\cdot \boldsymbol{r}_i} a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}'} a_{\boldsymbol{k}''} + \ldots, \tag{8.17}$$

$$S_i^- = \sqrt{\frac{2S}{N}} \sum_{\boldsymbol{k}} e^{-i\boldsymbol{k}\cdot R_i} a_{\boldsymbol{k}}^\dagger - \frac{1}{N\sqrt{8NS}} \sum_{\boldsymbol{k},\boldsymbol{k}',\boldsymbol{k}''} e^{-i(\boldsymbol{k}+\boldsymbol{k}'-\boldsymbol{k}'')\cdot R_i} a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}'}^\dagger a_{\boldsymbol{k}''} + \ldots, \tag{8.18}$$

$$S_i^z = -S + \frac{1}{N} \sum_{\boldsymbol{k},\boldsymbol{k}'} e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot R_i} a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}'}. \tag{8.19}$$

The Hamiltonian for a system of independent harmonic oscillators has the form $\sum \hbar\omega_{\boldsymbol{k}} a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}}$. Therefore the objective of the Holstein-Primakoff approach is to determine the quadratic part of the Hamiltonian resulting from substitutions (8.17)-(8.19). Let us consider this procedure for a ferromagnet (other systems are treated in [150]).

The Zeeman and exchange interactions were considered above. Let us add to these the dipole-dipole interaction (2.50). The total Hamiltonian is then

$$\mathcal{H} = \mathcal{H}_Z + \mathcal{H}_{\text{ex}} + \mathcal{H}_{\text{dip}}. \tag{8.20}$$

The low-temperature or spin-wave approximation (8.7) corresponds to taking $f_i(S) = 1$. This is because at very low temperatures the number of spin waves
excited is much less than the total number of spins. Therefore the average spin deviation n_i is much less than 1 . In terms of the magnon operators, the Zeeman interaction becomes

$$\mathcal{H}_z = \gamma H_0 \sum_i S_i^z = -\gamma H_0 \sum_i \left(S - \frac{1}{N} \sum_{\boldsymbol{k},\boldsymbol{k}'} e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot R_i} a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}'} \right). \tag{8.21}$$

If the lattice has translational invariance, then

$$\sum_i e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot \boldsymbol{r}_i} = N\Delta\left(\boldsymbol{k}-\boldsymbol{k}'\right). \tag{8.22}$$

Therefore (8.21) reduces to

$$\mathcal{H}_Z = -\gamma H_0 NS + \gamma H_0 \sum_k a_k^\dagger a_k. \tag{8.23}$$

Similarly, if the exchange coupling between an ion i and those ions at positions $\boldsymbol{\delta}$ relative to i has the value J, then the exchange interaction is

$$\mathcal{H}_{\text{ex}} = -J\sum_{i,\delta}\left[\frac{1}{2}S_i^+S_{i+\delta}^- + \frac{1}{2}S_i^-S_{i+\delta}^+ + S_i^zS_{i+\delta}^z\right] \tag{8.24}$$

$$\begin{aligned}\mathcal{H}_{\text{ex}} = -J\sum_{i,\delta}\Bigg[&\frac{S}{N}\sum_{\boldsymbol{k},\boldsymbol{k}'}e^{i\boldsymbol{k}'\cdot(\boldsymbol{r}_i+\boldsymbol{\delta})}a_{\boldsymbol{k}}a_{\boldsymbol{k}'}^\dagger \\ &+\frac{S}{N}\sum_{\boldsymbol{k},\boldsymbol{k}'}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_i}e^{i\boldsymbol{k}'\cdot(\boldsymbol{r}+\boldsymbol{\delta})}a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}'} \\ &+\left(S-\frac{1}{N}\sum_{\boldsymbol{k},\boldsymbol{k}'}e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}_i}a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}'}\right) \\ &\times\left(S-\frac{1}{N}\sum_{\boldsymbol{k},\boldsymbol{k}'}e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot(\boldsymbol{R}_i+\boldsymbol{\delta})}a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}'}\right)\Bigg].\end{aligned} \tag{8.25}$$

In keeping with our spirit of setting $f_i(S) = 1$, we also neglect the fourth-order terms in this expansion. Thus, if z is the number of interacting neighbors,

$$\begin{aligned}\mathcal{H}_{\text{ex}} = &-JS\sum_{\boldsymbol{k}}\sum_{\boldsymbol{\delta}}e^{-i\boldsymbol{k}\cdot\boldsymbol{\delta}}a_{\boldsymbol{k}}a_{\boldsymbol{k}}^\dagger - JS\sum_{\boldsymbol{k}}\sum_{\boldsymbol{\delta}}e^{i\boldsymbol{k}\cdot\boldsymbol{\delta}}a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}} \\ &-NzJS^2 + 2zJS\sum_{\boldsymbol{k}}a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}}.\end{aligned} \tag{8.26}$$

We define

$$\gamma_{\boldsymbol{k}} = \frac{1}{z}\sum_{\delta}e^{i\boldsymbol{k}\cdot\boldsymbol{\delta}}.$$

For example, with nearest-neighbor interactions in a simple cubic crystal having a lattice parameter a,

$$\gamma_{\boldsymbol{k}} = \frac{1}{3}\left(\cos k_x a + \cos k_y a + \cos k_z a\right).$$

Notice that

$$\sum_k \gamma_k = 0$$

Thus

$$\mathcal{H}_{\text{ex}} = -NzJS^2 + 2zJS\sum_{\boldsymbol{k}}\left(1-\gamma_{\boldsymbol{k}}\right)a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}} \tag{8.27}$$

Let us now consider the dipole-dipole interaction, whose treatment is more tedious.

$$\mathcal{H}_{\text{dip}} = \frac{1}{2} g^2 \mu_B^2 \sum_{i,j} \left[\frac{S_i \cdot S_j}{r_{ij}^3} - \frac{3\left(S_i \cdot \boldsymbol{r}_{ij}\right)\left(S_j \cdot \boldsymbol{r}_{ij}\right)}{r_{ij}^5} \right] \tag{8.28}$$

$$\begin{aligned} \mathcal{H}_{\text{dip}} =& \frac{1}{2} g^2 \mu_B^2 \sum_{i,j} \left\{ \frac{S_i^+ S_j^-}{2 r_{ij}^3} + \frac{S_i^- S_j^+}{2 r_{ij}^3} + \frac{S_i^z S_j^z}{r_{ij}^3} \right. \\ & \left. - \frac{3}{r_{ij}^5} \left[\left(\frac{1}{2} S_i^+ r_{ij}^- + \frac{1}{2} S_i^- r_{ij}^+ + S_i^z z_{ij} \right) \left(\frac{1}{2} S_j^+ r_{ij}^- + \frac{1}{2} S_j^- r_{ij}^+ + S_j^z z_{ij} \right) \right] \right\} \end{aligned} \tag{8.29}$$

$$\begin{aligned} \mathcal{H}_{\text{dip}} \approx& \frac{1}{2} g^2 \mu_B^2 \sum_{i,j} \left[S_i^+ S_j^- \frac{1}{2 r_{ij}^3} \left(1 - \frac{3}{2} \frac{r_{ij}^+ r_{ij}^-}{r_{ij}^2} \right) + S_i^- S_j^+ \frac{1}{2 r_{ij}^3} \left(1 - \frac{3}{2} \frac{r_{ij}^+ r_{ij}^-}{r_{ij}^2} \right) \right. \\ & \left. + S_i^z S_j^z \frac{1}{r_{ij}^3} \left(1 - 3 \frac{z_{ij}^2}{r_{ij}^2} \right) - \frac{3}{4} S_i^+ S_j^+ \frac{\left(r_{ij}^-\right)^2}{r_{ij}^5} - \frac{3}{4} S_i^- S_j^- \frac{\left(r_{ij}^+\right)^2}{r_{ij}^5} \right] \end{aligned} \tag{8.30}$$

where $r_{ij}^{\pm} = x_{ij} \pm i y_{ij}$.

The approximation symbol in (8.30) indicates that we have kept only those terms which will lead to quadratic magnon terms. In addition to neglecting three- and four-magnon terms, we have also neglected linear terms. The presence of such terms is an indication that our assumed ground-state spin configuration is not correct. To find the correct configuration we assume that the ground-state spin orientation at site i has some direction, characterized by the angles θ_i and φ_i with respect to the crystallographic axes. Setting the coefficients of the one-magnon terms to 0 determines these angles. In the dipolar
case this is a small effect, and we shall neglect it. However, in certain canted spin systems care must be exercised in choosing the proper configuration. Introducing the magnon operators, we obtain

$$\begin{aligned} \mathcal{H}_{\text{dip}} =& \frac{1}{2} g^2 \mu_B^2 \sum_{\boldsymbol{k},\boldsymbol{k}'} \sum_{i,j} \left[\frac{S}{N r_{ij}^3} \left(1 - \frac{3}{2} \frac{r_{ij}^+ r_{ij}^-}{r_{ij}^2} \right) a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}'} e^{i\boldsymbol{k}\cdot\boldsymbol{r}_{ij}} e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}_j} \right. \\ & + \frac{S}{N r_{ij}^3} \left(1 - \frac{3}{2} \frac{r_{ij}^+ r_{ij}^-}{r_{ij}^2} \right) a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}'} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_{ij}} e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}_j} + \frac{S^2}{r_{ij}^3} \left(1 - 3 \frac{z_{ij}^2}{r_{ij}^2} \right) \\ & - \frac{2S}{N r_{ij}^3} \left(1 - 3 \frac{z_{ij}^2}{r_{ij}^2} \right) a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}'} e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}_j} \\ & - \frac{3S}{2N} \frac{\left(r_{ij}^-\right)^2}{r_{ij}^5} a_{\boldsymbol{k}} a_{\boldsymbol{k}'} e^{i\boldsymbol{k}\cdot\boldsymbol{r}_{ij}} e^{i(\boldsymbol{k}+\boldsymbol{k}')\cdot\boldsymbol{R}_j} \\ & \left. - \frac{3S}{2N} \frac{\left(r_{ij}^+\right)^2}{r_{ij}^5} a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}'}^\dagger e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_{ij}} e^{-i(\boldsymbol{k}+\boldsymbol{k}')\cdot\boldsymbol{R}_j} \right] \end{aligned} \tag{8.31}$$

To evaluate the spatial sums we neglect boundary effects, so that the sums may be carried out independently; that is,

$$\sum_{\boldsymbol{R}_j} \sum_{\boldsymbol{R}_i} \to \sum_{\boldsymbol{R}_j} \sum_{r = \boldsymbol{R}_i - \boldsymbol{R}_j} \tag{8.32}$$

The dipole Hamiltonian then reduces to

$$\begin{aligned}\mathcal{H}_{\text{dip}} =&\frac{1}{2}g^2\mu_B^2\left\{N^2S^2\sum_{\boldsymbol{r}}\frac{1}{r^3}\left(1-\frac{3z^2}{\boldsymbol{r}^2}\right)+S\sum_{\boldsymbol{r}}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\frac{1}{r^3}\left(1-\frac{3}{2}\frac{r^+r^-}{r^2}\right)\right.\\&-\left[2S\sum_{\boldsymbol{r}}\frac{1}{r^3}\left(1-\frac{3z^2}{r^2}\right)-2S\sum_{\boldsymbol{r}}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\frac{1}{r^3}\left(1-\frac{3}{2}\frac{r^+r^-}{r^2}\right)\right]a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}}\\&\left.-\left[\frac{3}{2}S\sum_{\boldsymbol{r}}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\frac{(r^-)^2}{r^5}\right]a_{\boldsymbol{k}}a_{-\boldsymbol{k}}-\left[\frac{3}{2}S\sum_{\boldsymbol{r}}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\frac{(r^+)^2}{r^5}\right]a_{\boldsymbol{k}}^\dagger a_{-\boldsymbol{k}}^\dagger\right\}.\end{aligned}\tag{8.33}$$

We now break up the sum over $\boldsymbol{r}$ into a sum within a sphere of radius a plus an integral over the rest of the sample,

$$\sum_r=\sum\nolimits'+\frac{N}{V}\int dr\tag{8.34}$$

The reason for doing this is that if $a\ll 2\pi/k$, the magnon wavelength, then the sum for a cubic crystal is 0 .

Now consider the remaining integral associated with the first term in the Hamiltonian,

$$\begin{aligned}\frac{N}{V}\int\left(1-\frac{3z^2}{r^2}\right)d\boldsymbol{r}&=\frac{N}{V}\int\nabla\cdot\left(\frac{\boldsymbol{z}}{r^3}\right)d\boldsymbol{r}\\&=\frac{N}{V}\int_{\text{sphere}}\frac{\boldsymbol{z}\cdot d\boldsymbol{s}}{r^3}-\frac{N}{V}\int_{\substack{\text{sample}\\\text{surface}}}\frac{\boldsymbol{z}\cdot d\boldsymbol{S}}{r^3}\\&=\frac{N}{V}\left(\frac{4\pi}{3}-4\pi N_{\boldsymbol{z}}\right)\end{aligned}\tag{8.35}$$

where the demagnetization factor

$$N_z\equiv\frac{1}{4\pi}\int_{\substack{\text{sample}\\\text{surface}}}\frac{\boldsymbol{z}\cdot d\boldsymbol{S}}{r^3}$$

gives rise to the demagnetization field and the term $4\pi/3$ gives rise to the Lorentz field. Now consider the integral

$$\frac{N}{V}\int e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\frac{1}{r^3}\left(1-\frac{3z^2}{r^2}\right)d\boldsymbol{r}\tag{8.36}$$

which involves a plane-wave factor. Expanding the plane wave in spherical harmonics and recognizing that $1-3\left(z^2/r^2\right)$ is a spherical harmonic, we have

$$\begin{aligned}&-\frac{N}{V}\int\sum_{l,m}4\pi(i)^l j_l(kr)Y_l^{m*}(\Omega)Y_l^m\left(\Omega'\right)\frac{2}{r^3}\sqrt{\frac{4\pi}{5}}Y_2^0(\Omega)r^2drd\Omega\\&=-\frac{N}{V}4\pi\left[\frac{j_l(\boldsymbol{kr})}{\boldsymbol{kr}}\right]_a^{\text{surface}}\left(3\cos^2\theta_{\boldsymbol{k}}-1\right)\end{aligned}\tag{8.37}$$

where $\theta_{\boldsymbol{k}}$ is the angle that $\boldsymbol{k}$ makes with the z-axis. If the sample dimension d is large, then $kd\gg 1$ and $j_l(kd)/dk\to 0$. If $ka\ll 1$, then $j_l(ka)/ka\to\frac{1}{3}$. Therefore (8.36) becomes

$$\frac{N}{V}\left(\frac{8\pi}{3}-4\pi\sin^2\theta_k\right)\tag{8.38}$$

The condition $kd \gg 1$ gives a zero result because the magnetization on the surface of the sample alternates rapidly, averaging to 0 . This condition does not hold for small wave vectors. In particular, let us consider the $\boldsymbol{k} = 0$ mode. We shall return later to a discussion of the nonzero, but small, wave vectors. For the case $\boldsymbol{k} = 0$ we treat those terms involving $r^{\pm}$as we did those involving z^2. This leads to transverse demagnetizing factors N_x and N_y. After we have carried out all the integrals, the dipole-dipole Hamiltonian becomes

$$\begin{aligned}\mathcal{H}_{\text{dip}} = \sum_{\boldsymbol{k}} \Bigg[& \frac{1}{2} MV \left(4\pi N_z M - \frac{4\pi}{3} M \right) + \left\{ \begin{array}{l} \frac{1}{2}\omega_M \left(N_x + N_y \right) - N_z \omega_M \\ \frac{1}{2}\omega_M \sin^2 \theta_{\boldsymbol{k}} - N_z \omega_M \end{array} \right\} a_{\boldsymbol{k}}^{\dagger} a_{\boldsymbol{k}} \\ & + \left\{ \begin{array}{l} \frac{1}{4}\omega_M \left(N_x - N_y \right) \\ \frac{1}{4}\omega_M \sin^2 \theta_{\boldsymbol{k}} \exp\left(-2\varphi_{\boldsymbol{k}} \right) \end{array} \right\} a_{\boldsymbol{k}} a_{-\boldsymbol{k}} \\ & + \left\{ \begin{array}{l} \frac{1}{4}\omega_M \left(N_x - N_y \right) \\ \frac{1}{4}\omega_M \sin^2 \theta_{\boldsymbol{k}} \exp\left(i2\varphi_{\boldsymbol{k}} \right) \end{array} \right\} a_{\boldsymbol{k}}^{\dagger} a_{-\boldsymbol{k}}^{\dagger} \Bigg] \end{aligned} \tag{8.39}$$

where $M = Ng\mu_B S/V$ and $\omega_M = 4\pi\gamma M$. The upper terms in the braces apply to the $\boldsymbol{k} = 0$ mode and the lower terms apply to the $\boldsymbol{k} \neq 0$ modes. The total Hamiltonian may now be written as

$$\mathcal{H} = E_0 + \sum_k \left(A_k a_k^{\dagger} a_k + B_k a_k a_{-k} + B_k^* a_k^{\dagger} a_{-k}^{\dagger} \right) \tag{8.40}$$

where $A_{\boldsymbol{k}}$ and $B_{\boldsymbol{k}}$ are determined from (8.23), (8.27), and (8.39).
This is not yet the desired form, since the dipole-dipole interaction has produced a coupling between the $+\boldsymbol{k}$ and $-\boldsymbol{k}$ spin waves. This means that the plane waves are not the correct normal modes. To find the correct modes we must diagonalize $\mathcal{H}$. Since the Hamiltonian (8.40) is the product of operators, its diagonalization is slightly different from that of a product of c-numbers. In particular, suppose we write the Hamiltonian as

$$\mathcal{H} = x^{\dagger} H x \tag{8.41}$$

where x is a column vector whose components are the operators entering the Hamiltonian and H is a c-number matrix. For example, if we neglect the zero-point term, we may write the Hamiltonian (8.40) as

$$\mathcal{H} = \frac{1}{2} \sum_k x_k^{\dagger} H_k x_k \tag{8.42}$$

where

$$x_{\boldsymbol{k}} = \begin{bmatrix} a_{\boldsymbol{k}} \\ a_{-\boldsymbol{k}}^{\dagger} \end{bmatrix} \quad \text{and} \quad H_{\boldsymbol{k}} = \begin{bmatrix} A_{\boldsymbol{k}} & 2B_{\boldsymbol{k}}^* \\ 2B_{\boldsymbol{k}} & A_{-\boldsymbol{k}} \end{bmatrix}. \tag{8.43}$$

The operator nature of x may be specified by the commutator

$$\left[x, x^{\dagger} \right] \equiv x \left(x^{*T} \right) - \left(x^* x^T \right)^T \equiv g \tag{8.44}$$

Note that when the elements of matrices are operators, the transpose of a product of such matrices is not equal to the product of the transposes in reverse order.

It is generally desired that the new modes that diagonalize the Hamiltonian have the same commutation relations as the original modes; that is, the transformation must be canonical. If the new modes are written as

$$x = Sx' \tag{8.45}$$

then the condition that the transformation S diagonalizes the Hamiltonian and also preserves the commutations relations is [151]

$$HS = g^{-1} S g \Omega_H, \tag{8.46}$$

where Ω_H is the eigenvalue matrix. Solving this eigenvalue problem for our case leads to the so called Bogolubov transformation

$$\begin{bmatrix} a_k \\ a_{-k}^\dagger \end{bmatrix} = \begin{bmatrix} u_k & v_k^* \\ v_k & u_k \end{bmatrix} \begin{bmatrix} \alpha_k \\ \alpha_{-k}^\dagger \end{bmatrix}. \tag{8.47}$$

where

$$u_{\boldsymbol{k}} = \sqrt{\frac{A_{\boldsymbol{k}} + \omega_{\boldsymbol{k}}}{2\omega_{\boldsymbol{k}}}}$$
$$v_{\boldsymbol{k}} = \sqrt{\frac{A_{\boldsymbol{k}} - \omega_{\boldsymbol{k}}}{2\omega_{\boldsymbol{k}}}} \exp\left(i\varphi_{\boldsymbol{k}}\right)$$

In terms of these new modes, the Hamiltonian has the desired form

$$\mathcal{H} = \sum_k \hbar \omega_k \alpha_k^\dagger \alpha_k, \tag{8.48}$$

where

$$\begin{aligned} \omega_{\boldsymbol{k}} &= \sqrt{A_{\boldsymbol{k}}^2 - 4\left|B_{\boldsymbol{k}}\right|^2}, \\ &= \begin{cases} \sqrt{\left(\gamma H_i + N_x \omega_N\right)\left(\gamma H_i + N_y \omega_M\right)} & \boldsymbol{k} = 0, \\ \sqrt{\left[\gamma H_i + 2zJS\left(1 - \gamma_{\boldsymbol{k}}\right)\right]\left[\gamma H_i + 2zJS\left(1 - \gamma_{\boldsymbol{k}}\right) + \omega_M \sin^2\theta_{\boldsymbol{k}}\right]}, & \boldsymbol{k} \neq 0, \end{cases} \end{aligned}$$

and H_i is the internal field $H_0 - 4\pi N_z M$. These modes are plotted in Fig. 8.2 as a function of wave vector. Notice that as a result of the demagnetization fields the ferromagnetic resonance frequency depends on the shape of the sample. This result was first derived by Kittel [152].

The precession associated with the modes a_k is circular and that associated with the α_k is, in general, elliptical. Owing to the nature of the coefficient B_k, the uniform precession is elliptical when the transverse demagnetization factors are different. We also find that the ellipticity associated with a spinwave mode depends on the direction of propagation with respect to the dc field. The most elliptical spin waves are those which propagate at right angles to the dc field.

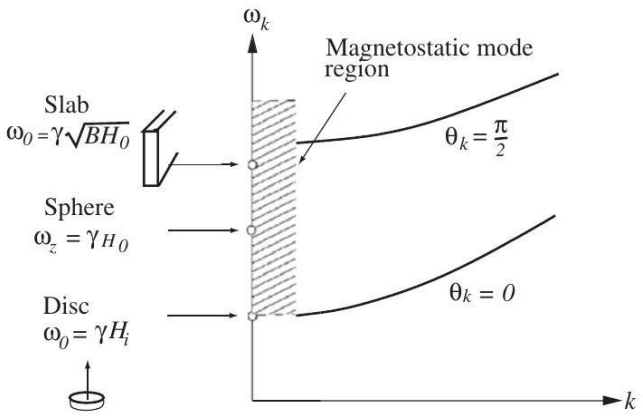


Fig. 8.2. Spin-wave frequencies as a function of their wave vector

8.2.2 Coherent Magnon State

The reader might notice that since the representations for $S_i^+$and S_i^-involve odd powers of the magnon creation and annihilation operators, then S_i^x and S_i^y will have zero expectation value in a pure magnon state, i.e.,

$$\langle n_{\boldsymbol{k}}| \, S_i^{\mu}(t) \, |n_{\boldsymbol{k}}\rangle = 0$$

How does this relate to the classical picture shown, for example, in Fig. 8.1? This situation is analogous to that in optics where the number of photons and the phase are conjugate variables. If a state is characterized by a given number of photons the phase is undetermined. In our magnetic case, to obtain the classical picutre where the phase is well-defined, we need a superposition of magnon states. In particular, the "coherent" magnon state, $|c_{\boldsymbol{k}}\rangle$, is defined as the eigenstate of the magnon annihilation operator, i.e.,

$$\alpha_{\boldsymbol{k}} \, |c_{\boldsymbol{k}}\rangle = c_{\boldsymbol{k}} \, |c_{\boldsymbol{k}}\rangle \tag{8.50}$$

where $c_{\boldsymbol{k}}$ is a complex number having a phase $\beta_{\boldsymbol{k}}$ and an amplitude given by the average number of magnons in the state $|c_{\boldsymbol{k}}\rangle$,

$$|c_{\boldsymbol{k}}|^2 = \langle c_{\boldsymbol{k}}| \, n_{\boldsymbol{k}} \, |c_{\boldsymbol{k}}\rangle$$

By analogy with the photon case, the coherent state is given by [153]

$$|c_{\boldsymbol{k}}\rangle = e^{-\frac{1}{2}|c_k|^2} \sum_{n_k=0}^{\infty} \frac{(c_k)^{n_k}}{(n_k!)^{1/2}} \, |n_k\rangle \tag{8.51}$$

The average spin components in this state are

$$\langle c_k| \, S_i^x(t) \, |c_k\rangle = S_k \cos\left(\boldsymbol{k}\cdot\boldsymbol{r}_i - \omega_k t + \beta_k\right) \tag{8.52a}$$
$$\langle c_k| \, S_i^y(t) \, |c_k\rangle = \epsilon_k S_k \sin\left(\boldsymbol{k}\cdot\boldsymbol{r}_i - \omega_k t + \beta_k\right), \tag{8.52b}$$

where

$$S_k = \left(\frac{2S}{N}\right)^{1/2} (u_k - v_k) \, |c_k| \tag{8.53}$$

and

$$\epsilon_{\boldsymbol{k}} = \frac{u_k + v_k}{u_k - v_k} \tag{8.54}$$

is the ellipticity. Equations (8.52a + b) are the familiar spin precession solutions.

8.2.3 Magnetostatic Modes

Let us now consider those modes which, although not uniform, have a spatial variation comparable to the sample dimensions. We use the fact that at such wavelengths the exchange interaction is negligible. Therefore the problem is a classical one. Furthermore, since $kc \gg \omega$, these modes are essentially static. Thus they are described by the magnetostatic Maxwell equations

$$\nabla \times \boldsymbol{H} = 0 \tag{8.55}$$
$$\nabla \cdot \boldsymbol{H} = -4\pi \nabla \cdot \boldsymbol{M} \tag{8.56}$$

plus the constitutive torque equation

$$\frac{d\boldsymbol{M}}{dt} = -\gamma \boldsymbol{M} \times \boldsymbol{H}. \tag{8.57}$$

The solution to (8.57) under the assumption of small transverse fields is

$$\boldsymbol{B} = \begin{bmatrix} \mu & -i\kappa & 0 \\ i\kappa & \mu & 0 \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{H}, \tag{8.58}$$

where

$$\mu = 1 + \frac{4\pi\gamma^2 M_0 H_i}{\gamma^2 H_i^2 - \omega^2} \tag{8.59}$$

and

$$\kappa = \frac{4\pi\gamma M_0 \omega}{\gamma^2 H_i^2 - \omega^2}. \tag{8.60}$$

If we introduce a magnetic scalar potential ϕ, defined by $\boldsymbol{H} = \boldsymbol{\nabla}\phi$, Maxwell's equations become

$$\begin{aligned} \mu\left(\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2}\right) + \frac{\partial^2\phi}{\partial z^2} = 0 \quad &(\text{ inside }) \\ \nabla^2\phi = 0 \quad &(\text{ outside }). \end{aligned} \tag{8.61}$$

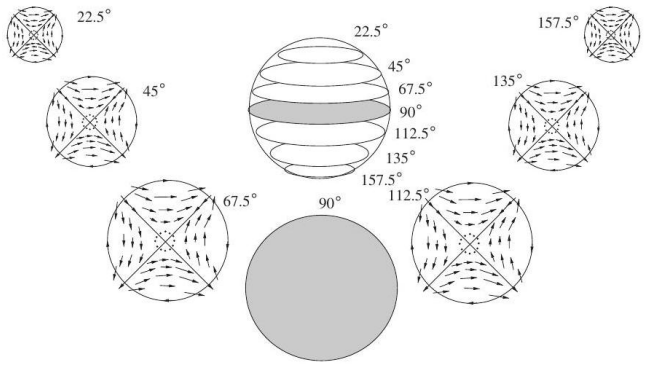


Fig. 8.3. The $(4, 3, 0)$ Walker mode of a sphere showing the instantaneous positions of the radio-frequency magnetization vectors in the places transverse to the dc magnetic field. This illustrates the indexing scheme and the concept of nonuniform precession [154]

Solving these equations for an ellipsoidal sample gives the so-called magnetostatic, or Walker, modes [154]. These modes are described by a set of indices (n, m, r) instead of (k_x, k_y, k_z). For example, the ($4, 3, 0$) mode is shown in Fig. 8.3. Since these modes have an inhomogeneous magnetization, they may be excited by an inhomogenous microwave field. In fact, if the sample is very large or if it is mounted on a dielectric post, the resulting inhomogeneity in the field may be enough to excite magnetostatic modes. These modes were discovered in this way [155]. A typical absorption spectrum is shown in Fig. 8.4. An interesting feature of Fig. 8.4a is that the magnetostatic resonances are not broadened by the inhomogeneous magnetic field. In a paramagnet such a field produces an inhomogeneously broadened line whose

width is approximately equal to the field variation. In the ferromagnetic case, however, the magnetostatic modes are the eigensolutions to (8.61). Thus, even if μ given by (8.59), has a spatial dependence through H_i the modes may not be as symmetrical as shown in Fig. 8.3 but will still be infinitely narrow. We might describe this as "dipolar narrowing".

8.2.4 Solitons

The theory of magnetostatic modes for ferromagnetic thin films was developed by Damon and Eshbach [157]. For the coordinate system shown in Figs. 8.5 and 8.6 they found that the modes for an unbounded film consisted of propagating waves of the form

$$m_{x,y}(z) \sim e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} \tag{8.62}$$

where the wavevector $\boldsymbol{k}$ lies in the plane of the film. They found that for angles of propagation, φ, less than a critical angle given by

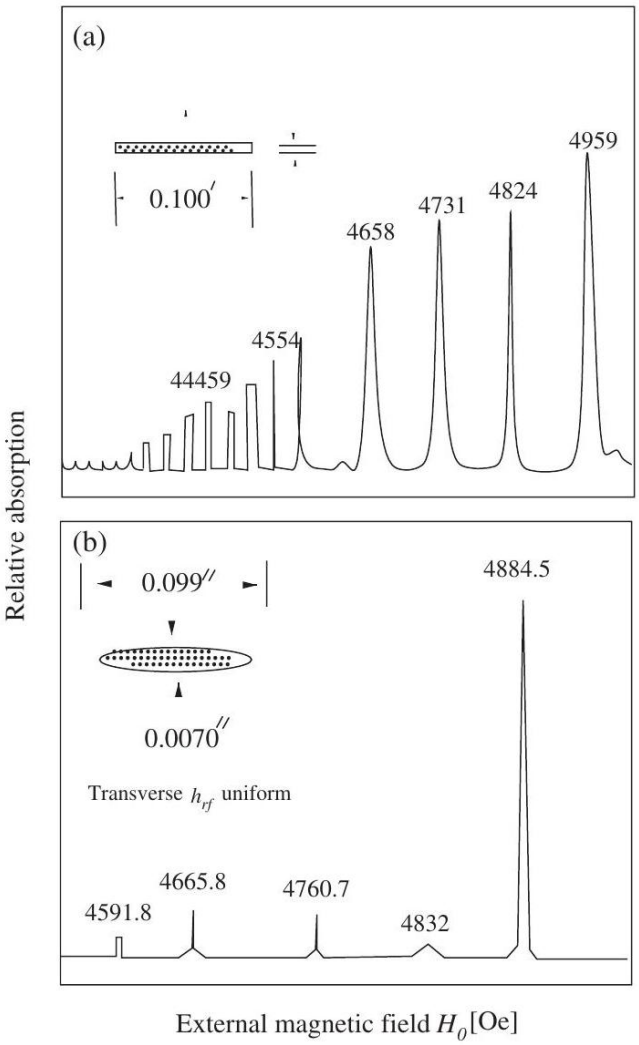


Fig. 8.4. Absorption spectrum of a yttrium-iron garnet disk. The sharp edges in (a) produce nonuniform fields which excite many modes. Rounding these edges (b) suppresses the modes [156]

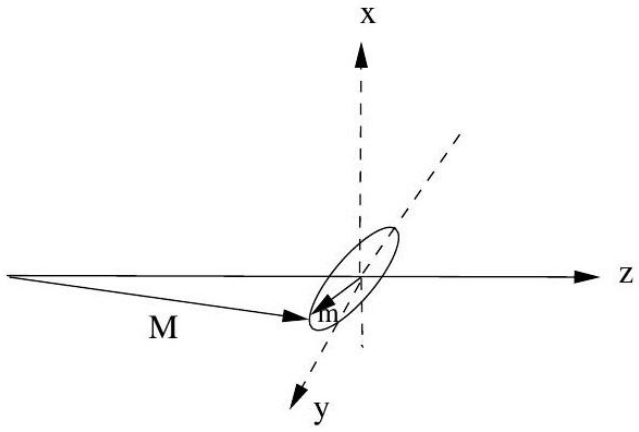


Fig. 8.5. Elliptical precession of the magnetization in a thin film

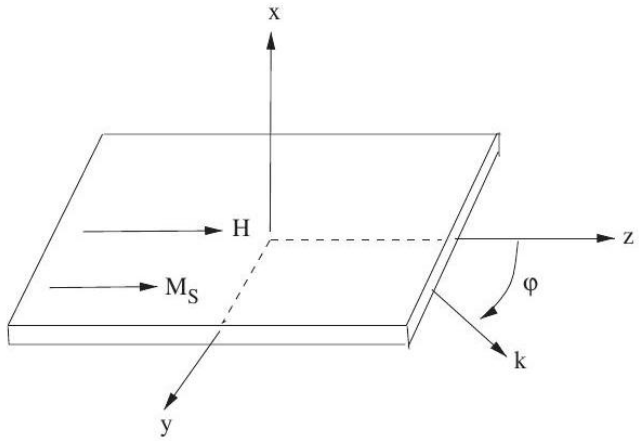


Fig. 8.6. Coordinate system used to define the spinwave modes of a thin film

$$\varphi_c = \tan^{-1} \sqrt{\frac{H_0}{4\pi M_s}} \tag{8.63}$$

the frequencies, ω, of these modes falls within the band of bulk spinwave frequencies, i.e., between $\omega = \gamma H_0$ and $\omega = \gamma\sqrt{H_0\,(H_0 + 4\pi M_s)}$. When φ exceeds φ_c then surface waves exist with frequencies above the bulk spinwave manifold. These modes are indicated in Fig. 8.7. The terminology "forward" or "backward" volume wave indicates the sign of the group velocity, $v_g = \partial\omega_k/\partial k$. A magnetostatic backward volume wave (MSBVW) has a negative group velocity. This means that as a wave packet of sinusoidal waves moves forward the sinusoidal wave moves backward. (Much like the stroboscopic effect that makes the stagecoach wheels in Western movies appear to rotate backwards!)

The surface waves are "nonreciprocal" in the sense that when φ has one sign the mode is localized on one surface. When φ has the opposite sign the localization is on the other surface. Also, the exponential decay away from the surface scales with the propagation wave vector k. Thus, as k increases the response becomes more and more localized at the surface.

The existence of surface waves is of interest because of the possible application for signal processing. These waves may be excited by exposing the surface of a magnetic film to pulses of microwave radiation generated, for example, by a transmission line. Such a transducer generates a wave packet. This packet is described by an envelope function $u(z,t)$, i.e., the signal is given by

$$m_y(z,t) = M_S u(z,t) e^{-i(kz-\omega t)}. \tag{8.64}$$

The equation of motion for the envelope function can be obtained from the magnetostatic equations which take the form [158]

$$i\left[\frac{\partial u}{\partial t} + v_g \frac{\partial u}{\partial z} + \eta u\right] + \frac{1}{2}\omega_k'' \frac{\partial^2 u}{\partial z^2} - N|u|^2 u = 0. \tag{8.65}$$

v_g is the group velocity; η is a damping parameter; $\omega_k'' = \partial^2\omega_k/\partial k^2$ and represents dispersion; N represents the effect of nonlinearity, $N = \partial\omega_k/\partial|u|^2$.

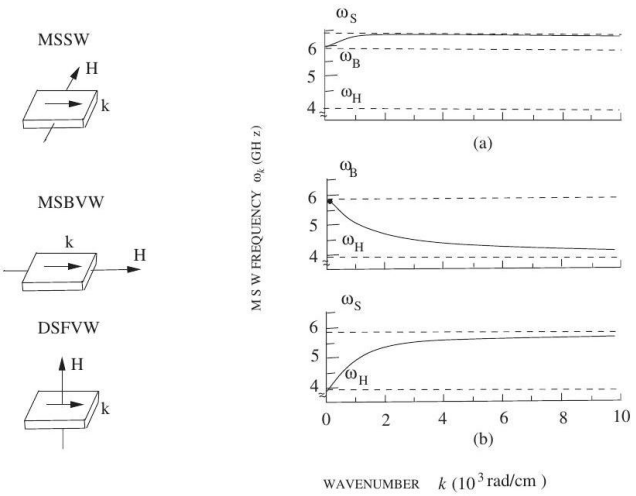


Fig. 8.7. Dispersion diagrams of magnetostatic wave frequency $\omega_{k $ vs wave number k for the three basic thin-film magnetostatic wave configurations, (a) magnetostatic surface waves (MSSW), (b) magnetostatic-backward-volume waves (MSBVW), and magnetostatic-forward-volume waves (MSFVW). The relative orientations of the net static field $\boldsymbol{H}$ and the wave vector $\boldsymbol{k}$ to each other and to the film for each configuration are shown by the diagrams to the left of the dispersion diagrams. The horizontal dashed lines and the frequencies ω_S, ω_B, and ω_H denote band limits for the different configurations. The curves are actual calculated curves from magnetostatic wave theory for a $7.2-\mu$m-thick yttrium iron garnet film with $H = 1407$Oe [158]}

Dispersion tends to broaden the packet while nonlinearity tends to steepen it. If $N\omega_k'' < 0$ there are solutions to this equation corresponding to propagation of pulses that preserve their shape. Such a pulse is called a soliton. The conditions for the observation of solitons are rather restrictive. However, solitons have been observed in films of yttrium iron garnet (YIG) [158]. The film was excited with a microwave pulse having a width of several ns and detected by a stripline antenna a few mm away. As the input power (of the order of a watt) is increased the nonlinear term in (8.65) increases and when the condition for soliton formation is reached the output power increases rapidly and may also show several pulses corresponding to multi-soliton generation.

8.2.5 Thermal Magnon Effects

Fortunately the density of spin-wave modes is much greater than that of the magnetostatic modes, so that most of the properties of magnetic materials may be understood by considering

only the simpler spin-waves modes. The importance of spin waves in determining the low-temperature properties of ferromagnets was first recognized by Bloch [159]. Let us first consider the saturation magnetization. In the spin-wave region this is given by

$$M(T) = -\frac{g\mu_B}{V}\sum_i \langle S_i^z\rangle = \frac{g\mu_B}{V}\sum_i (S - \langle n_i\rangle) = M(0) - \frac{g\mu_B}{V}\sum_{\boldsymbol{k}} \langle n_{\boldsymbol{k}}\rangle \,. \tag{8.66}$$

In principle, there is no limit to the number of magnons that may be in any given mode as long as the system does not become "saturated" as described by the factor $f_i(S)$ in (8.14). Their thermal occupation number is given by the Bose-Einstein distribution as

$$\langle n_{\boldsymbol{k}}\rangle = 1/\left[\exp\left(\hbar\omega_{\boldsymbol{k}}/k_BT\right) - 1\right] \tag{8.67}$$

Since most of the modes lie in the exchange region of the spectrum, we shall neglect dipolar effects. Thus $a_{\boldsymbol{k}}$ is the correct mode, and we do not have to transform to $\alpha_{\boldsymbol{k}}$. If we also use the small k expansion of (8.27) and neglect the external field, then

$$\hbar\omega_{\boldsymbol{k}} = Dk^2 \tag{8.68}$$

where, for example, D is $2JSa^2$ for a simple cubic lattice. The sum over $\boldsymbol{k}$ is now replaced by an integral, and we have

$$M(0) - M(T) = -\frac{g\mu_B}{V}\frac{4\pi V}{(2\pi)^3}\int_0^\infty \frac{k^2dk}{\exp\left(Dk^2/k_BT\right) - 1} \tag{8.69}$$

The integral may be written in the dimensionless form

$$\int_0^\infty \frac{k^2dk}{\exp\left(Dk^2/k_BT\right) - 1} = \frac{1}{2}\left(\frac{k_BT}{D}\right)^{3/2}\int_0^\infty \frac{x^{1/2}dx}{e^x - 1} \tag{8.70}$$

The evaluation of this integral is simplified by using the expansion

$$\frac{1}{e^x - 1} = \sum_{r=1}^{\infty} e^{-rx} \tag{8.71}$$

The resulting magnetization is

$$M(0) - M(T) = \zeta\left(\frac{3}{2}\right)\frac{g\mu_B}{M(0)}\left(\frac{k_B}{4\pi D}\right)^{3/2} T^{3/2} \tag{8.72}$$

where ζ is the Riemann zeta function.
This is referred to as the Bloch $T^{3/2}$ law. The fact that it accurately describes the magnetization at low temperatures, as shown in Fig. 8.8, is evidence of the importance of such collective modes. If we keep those terms of order k^4, k^6, etc., in the dispersion relation, they lead to contributions to the magnetization that vary as $T^{5/2}, T^{7/2}$, etc. If only single-particle excitations were important, the low-temperature magnetization would vary exponentially, with an activation energy of the order of the exchange energy.

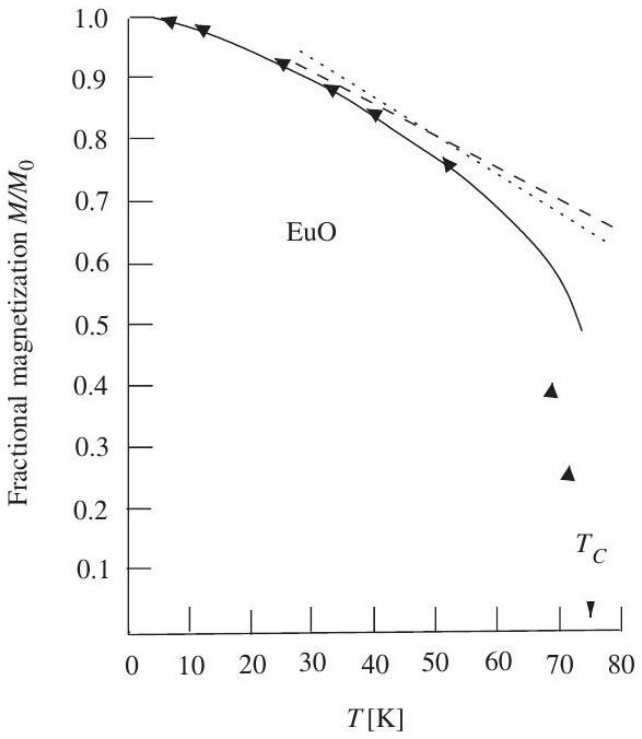


Fig. 8.8. The temperature dependence of the magnetization of EuO . The broken curve corresponds to noninteracting spin waves; the dotted curve corresponds to the series expansion (8.71) leading to terms of the form $T^{3 / 2}$, $T^{5 / 2}$, etc. The solid curve, calculated from spin-wave theory, includes the effects of dynamical interactions between pairs of magnons and subsequent renormalization of the spin-wave energies used in evaluating the magnon populations. Experimental points: ($\boldsymbol{\Delta}$) observations corrected to infinite applied field; (∇) observations corrected to zero applied field [160]}

The spin-wave contribution to the specific heat is given by

$$C_V = \frac{\partial}{\partial T} \sum_{\boldsymbol{k}} \hbar \omega_{\boldsymbol{k}} \langle n_{\boldsymbol{k}} \rangle . \tag{8.73}$$

With the same approximations employed in obtaining the magnetization, (8.73) gives

$$C_V = \frac{15}{4} \zeta \left(\frac{5}{2} \right) k_B \left(\frac{k_B}{4\pi D} \right)^{3/2} T^{3/2}. \tag{8.74}$$

8.2.6 Nonlinear Processes

So far our treatment has been entirely linear. That is, we have restricted ourselves to those terms in the Hamiltonian that involve quadratic products of magnon operators and may therefore be transformed into a set of non-interacting harmonic oscillators. The presence of higher-order terms corresponds to coupling between these oscillators. This coupling has important consequences for many properties of ordered systems.

Renormalization. The lowest-order nonlinear terms arising from the isotropic exchange interaction are the four-magnon terms

$$\mathcal{H}_{\text{ex}}^{(4)} = \frac{1}{2} \sum_{\boldsymbol{k},\boldsymbol{k}',\boldsymbol{k}''} \left[J(\boldsymbol{k}) + J\left(\boldsymbol{k} + \boldsymbol{k}' - \boldsymbol{k}''\right) - 2J\left(\boldsymbol{k} - \boldsymbol{k}''\right) \right] a_{\boldsymbol{k}}^{\dagger} a_{\boldsymbol{k}'}^{\dagger} a_{\boldsymbol{k}''} a_{\boldsymbol{k}+\boldsymbol{k}'+\boldsymbol{k}''}. \tag{8.75}$$

The effect of such spin-wave interactions on the magnetization was first analyzed by Dyson [161], who found that, in addition to the terms discussed above, there is a contribution to $M(0)-M(T)$ which is proportional to T^4. Although it is experimentally difficult to distinguish this contribution from $T^{3/2}$ and $T^{5/2}$, it has important theoretical consequences in that any alternative approach to computing $M(T)$ must reproduce Dyson's result at low temperatures.

A quantity that is more directly affected by the nonlinear terms of (8.75) is the magnon frequency itself. Let us consider only the exchange interaction. In the linear approximation the magnon frequency as given by (8.49) is independent of temperature. Applying the random-phase approximation and considering only nearest-neighbor interactions, we may write (8.75) as

$$-\left(2zJS^2N\right)^{-1}\sum_{\boldsymbol{k}}\omega_{\boldsymbol{k}}\left\langle n_{\boldsymbol{k}}\right\rangle\sum_{\boldsymbol{k}'}\omega_{\boldsymbol{k}'}\left\langle n_{\boldsymbol{k}'}\right\rangle. \tag{8.76}$$

If we add this expression to the linear part $\sum\omega_{\boldsymbol{k}}\left\langle n_{\boldsymbol{k}}\right\rangle$ and minimize the free energy with respect to $\left\langle n_{\boldsymbol{k}}\right\rangle$, we obtain the same form as (8.67), but with the renormalized frequency

$$\omega_{\boldsymbol{k}}(T)=\omega_{\boldsymbol{k}}\left[1-\left(zJS^2N\right)^{-1}\sum_{\boldsymbol{k}'}\omega_{\boldsymbol{k}'}\left\langle n_{\boldsymbol{k}'}\right\rangle\right]. \tag{8.77}$$

We recognize the sum in this expression as the total energy in the spin-wave system. As the temperature increases this energy also increases, causing a decrease in the spin-wave frequencies. This decrease has been observed by neutron scattering, which we shall discuss in Chap. 10.

Relaxation. Nonlinearity also plays an imporant role in spin-wave relaxation phenomena. Consider, for example, the lowest-order nonlinear terms arising from the dipolar interaction. These are the three-magnon terms

$$\mathcal{H}_{\text{dip}}^{(3)}=\sum_{\substack{\boldsymbol{k},\boldsymbol{k}'\\ \boldsymbol{k}\neq 0}}\left(g_{\boldsymbol{k}}\alpha_{\boldsymbol{k}}\alpha_{\boldsymbol{k}'}\alpha_{\boldsymbol{k}+\boldsymbol{k}'}^{\dagger}+g_{\boldsymbol{k}}^{*}\alpha_{\boldsymbol{k}}^{\dagger}\alpha_{\boldsymbol{k}'}^{\dagger}\alpha_{\boldsymbol{k}+\boldsymbol{k}'}\right), \tag{8.78}$$

where

$$g_{\boldsymbol{k}}=-\frac{4\pi}{V}\sqrt{2NS}g^2\mu_B^2\cos\theta_{\boldsymbol{k}}\sin\theta_{\boldsymbol{k}}\exp\left(i\phi_{\boldsymbol{k}}\right). \tag{8.79}$$

Since $\alpha_{\boldsymbol{k}}^{\dagger}$ corresponds to the creation of a magnon with wave vector $\boldsymbol{k}$ and $\alpha_{\boldsymbol{k}}$ corresponds to the annihilation of a magnon with wave vector $\boldsymbol{k}$, the

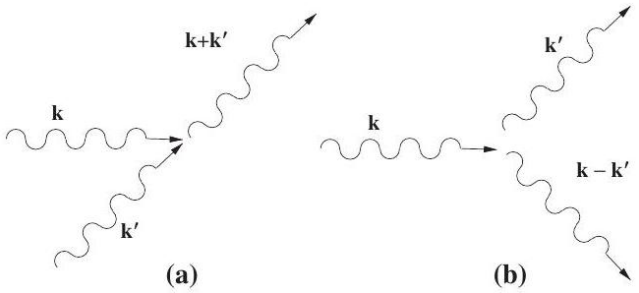


Fig. 8.9. Representation of (a) three-magnon confluence process and (b) threemagnon splitting process

processes involved in (8.78) may be represented schematically as shown in Fig. 8.9. Figure 8.9a shows the confluence of a magnon $\boldsymbol{k}$ with a magnon $\boldsymbol{k}'$ to produce a third magnon $\boldsymbol{k}+\boldsymbol{k}'$. This type of process constitutes a relaxation channel for the magnon $\boldsymbol{k}$ (or $\boldsymbol{k}'$).

Since we are dealing with a ferromagnet, the exchange-narrowing effect, which was discussed in Chap. 5, with be quite pronounced. The origin of this narrowing has to do with the fact that when the frequency spectrum of the effective field is very broad, only a narrow band of nearly degenerate modes influences the relaxation. Under this condition the relaxation frequency may be calculated by time-dependent perturbation theory. This is a particularly convenient way of computing magnon relaxation frequencies. As an illustration of this approach consider the three-magnon process of Fig. 8.9. The assumption that the magnon occupation number, if disturbed, relaxes exponentially to its equilibrium value $\langle n_{\boldsymbol{k}} \rangle$ enables us to define a relaxation frequency $\eta_{\boldsymbol{k}}$ by the rate equation

$$\frac{dn_{\boldsymbol{k}}}{dt} = -\eta_{\boldsymbol{k}} \left(n_{\boldsymbol{k}} - \langle n_{\boldsymbol{k}} \rangle \right). \tag{8.80}$$

The rate of change of the number of magnons in mode $\boldsymbol{k}$ is also given by

$$\frac{dn_k}{dt} = W_{n_k \to n_k+1} - W_{n_k \to n_k-1}, \tag{8.81}$$

where W is the transition probability per unit time. For the three-magnon interaction (8.68), this is given by [162]

$$\begin{aligned} W_{n_{\boldsymbol{k}} \to n_{\boldsymbol{k}}+1} =& \frac{2\pi}{\hbar^2} \sum_{\boldsymbol{k}'} \left| g^*_{\boldsymbol{k}} + g^*_{\boldsymbol{k}'} \right|^2 (n_{\boldsymbol{k}} + 1)(n_{\boldsymbol{k}'} + 1) \\ & \times n_{\boldsymbol{k}+\boldsymbol{k}'} \delta \left(\omega_{\boldsymbol{k}} + \omega_{\boldsymbol{k}'} - \omega_{\boldsymbol{k}+\boldsymbol{k}'} \right). \end{aligned} \tag{8.82}$$

The result is that this three-magnon confluence process gives a relaxation frequency which, for small wave vectors, is proportional to the magnitude of the wave vector and is linear in temperature. These features, as well as the order of magnitude, are found to agree well with experimental measurements of $\eta_{\boldsymbol{k}}$ in the ferromagnetic insulator yttrium-iron garnet. This process is not found
to describe the relaxation of the $\boldsymbol{k} = 0$, or ferromagnetic resonance, mode. It appears that surface and volume inhomogeneities provide the dominant relaxation channel for this mode [163].

Parametric Excitation

Suhl Instabilities. Since the $\boldsymbol{k} \neq 0$ magnons have such short wavelengths, they cannot be directly excited by an electromagnetic field. However, there are a number of mechanisms by which they may be indirectly excited. Some of these mechanisms involve the nonlinearities discussed above. For example, consider the three-magnon terms in (8.68), in which $\boldsymbol{k}' = -\boldsymbol{k}$, that is,

$$\sum_{\boldsymbol{k}} \left(g_{\boldsymbol{k}} \alpha_{\boldsymbol{k}} \alpha_{-\boldsymbol{k}} \alpha_0^\dagger + g^*_{\boldsymbol{k}} \alpha^\dagger_{\boldsymbol{k}} \alpha^\dagger_{-\boldsymbol{k}} \alpha_0 \right)$$

These terms describe the relaxation of the $\boldsymbol{k} = 0$, or uniform-precession, mode by decay into two magnons with equal and opposite wave vectors. Since energy must be conserved in this process, the resultant magnons will each have a frequency of $\omega_0/2$, as illustrated in Fig. 8.10.

The essential feature of this relaxation process is that it is nonlinear. That is, the rate at which the uniform-precession mode decays in this fashion depends on the occupation number $\langle n_{\boldsymbol{k}} \rangle$ of the final-state magnons, as seen for example in (8.82). This occupation number, however, depends on the magnon relaxation rate $\eta_{\boldsymbol{k}}$, as well as on the amplitude of the uniform precession $\langle n_0 \rangle$ which is driving it. There is a critical value of $\langle n_0 \rangle$ at which the rate of increase in the number of magnons in mode $\boldsymbol{k}$ exceeds their rate of decrease through relaxation. At this point the number of magnons $\langle n_{\boldsymbol{k}} \rangle$ increases abruptly, producing a corresponding increase in the relaxation rate of the uniform precession. Since the amplitude of the uniform precession $\langle n_0 \rangle$ is proportional to the amplitude of the driving microwave field H_1, this process appears as a saturation phenomenon. It turns out, however, that the field at which this

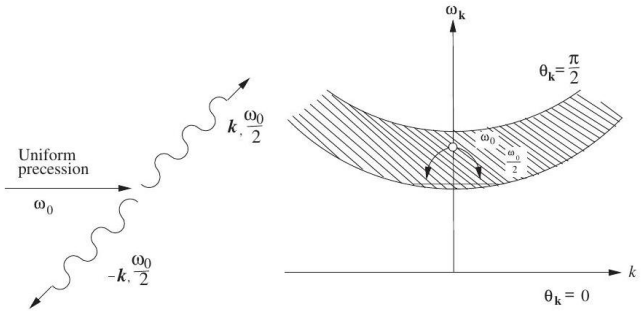


Fig. 8.10. Representation of the first-order Suhl instability in which uniform precession magnons split into ($\boldsymbol{k
,-\boldsymbol{k}$) pairs}

saturation happens is lower than that associated with the longitudinal relaxation time T_1 (see Sect. 6.4.1). The observation of this "premature" saturation in ferrites by Bloembergen and Damon in 1952 opened an era of exploration of nonlinear phenomena in microwave ferrites [164]. The explanation of this saturation was given by Suhl and is referred to as the first-order Suhl instability [165]. There is also a second-order Suhl instability which arises from terms in (8.65) of the form $\alpha^\dagger_{\boldsymbol{k}} \alpha^\dagger_{-\boldsymbol{k}} \alpha_0 \alpha_0$, in which two uniform precession magnons decay into a ($\boldsymbol{k}, -\boldsymbol{k}$) magnon pair.

Parallel Pumping. Since all these processes involve the magnon relaxation rate $\eta_{\boldsymbol{k}}$, they serve to measure this quantity. However, the onset of these processes is sometimes difficult to determine experimentally. A more accurate determination of $\eta_{\boldsymbol{k}}$ is obtained by means of the so-called parallel-pumping instability. In parallel pumping, as the name implies, a microwave magnetic field $H_1 \cos \omega t$ is applied parallel to the dc magnetic field. From (8.19) we find that the interaction Hamiltonian is

$$\mathcal{H}_1 = -g\mu_B H_1 \cos \omega t \sum_{\boldsymbol{k}} a^\dagger_{\boldsymbol{k}} a_{\boldsymbol{k}}. \tag{8.83}$$

We saw that because of the dipolar interaction the $a_{\boldsymbol{k}}$ are not the normal modes. When (8.47) is used to express the interaction in terms of the normal modes $\alpha_{\boldsymbol{k}}$ we obtain terms of the form

$$g\mu_B H_1 \cos \omega t \sum_{\boldsymbol{k}} \frac{|B_{\boldsymbol{k}}|}{\omega_{\boldsymbol{k}}} \exp\left(i2\varphi_{\boldsymbol{k}}\right) \alpha^\dagger_{\boldsymbol{k}} \alpha^\dagger_{-\boldsymbol{k}} \tag{8.84}$$

These terms correspond to the creation of $(\boldsymbol{k}, -\boldsymbol{k})$ magnon pairs. Again, since this is a nonlinear process, the rate at which energy flows into these magnons depends on their occupation number $\langle n_{\boldsymbol{k}} \rangle$. This, in turn, depends upon the amplitude of the driving field, as well as on the

relaxation rate $\eta_{\boldsymbol{k}}$. When the driving rate exceeds the relaxation rate, the magnon occupation number becomes very large. This results in an increased in absorption. The threshold for this process is given by

$$(H_1)_{\text{crit}} = \min\left(\frac{2\omega\eta_{\boldsymbol{k}}}{\gamma\omega_M \sin^2\theta_{\boldsymbol{k}}}\right), \tag{8.85}$$

where min means that we use that value of $\boldsymbol{k}$ giving the minimum.
Notice that those magnons which propagate perpendicular to the dc field, and hence have the greatest ellipticity, are the first to become unstable. Since there is essentially no absorption below this threshold, its onset provides us with an accurate means of determining the spin-wave relaxation frequency $\eta_{\boldsymbol{k}}$.

8.2.7 Chaos

If the driving fields in the first- or second-order Suhl instabilities or in the parallel-pumping instability are increased beyond their critical values a variety
of time dependent phenomena appear. Consider, for example, the secondorder Suhl process in which the ferromagnetic resonance is excited to a large amplitude. When the microwave field $\boldsymbol{H}_1$ reaches the critical value

$$(H_1)_{\text{crit}} = \frac{\eta_0}{\gamma}\left(\frac{2\eta_{\boldsymbol{k}}}{\omega_M}\right)^{1/2} \tag{8.86}$$

the amplitude of the spinwave mode with frequency $\omega_{\boldsymbol{k}} = \omega$ and $\theta_{\boldsymbol{k}} = 0$ begins to grow. This results in the "premature" saturation of the amplitude of the uniform precession. Unexplained so-called auto-oscillations had also been reported at high power levels. To study these phenomena Gibson and Jeffries [166] used the experimental arrangement shown in Fig. 8.11 in which a pickup coil monitors changes in the transverse magnetization of a Ga-YIG sample that is being excited by an rf field H_1 at a frequency of 1.3 GHz . When the rf field reaches a value several times $(H_1)_{\text{crit}}$ the voltage across the pickup coil develops auto-oscillations at a frequency of 16 kHz . As $\boldsymbol{H}_1$ is increased further sucessive period doublings occur as shown in Fig. 8.12. These results are part of a broad range of phenomena referred to as "chaos". Virtually all physical systems show chaotic behavior, and a large body of literature exists on chaotic dynamics. Nonlinearity is an essential part of chaos. Thus the description of a system showing chaotic behavior is generally based on a set of differential equations containing nonlinearity. For example, in 1963 Lorenz developed a set of three differential equations to describe thermal convection. The variables are related to the flow velocity and the temperature. The solution consists of

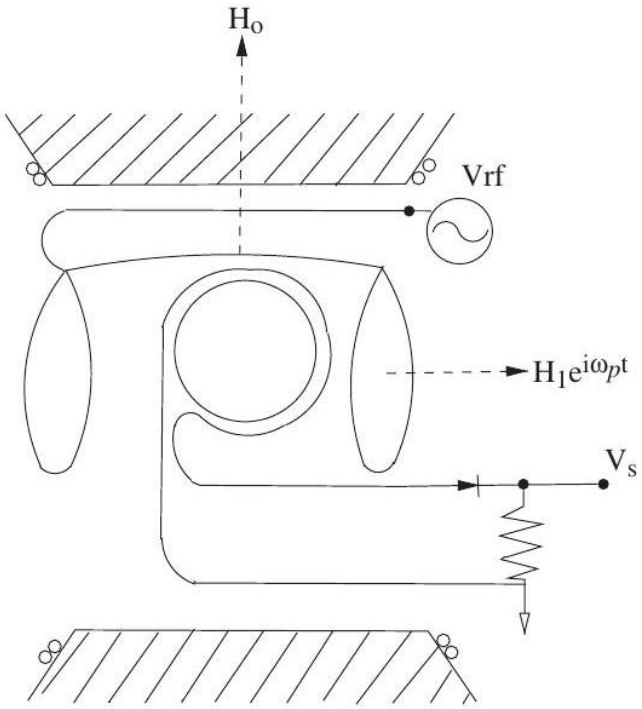


Fig. 8.11. Experimental arrangements used by Gibson and Jeffries to study autooscillations and chaos in the second-order Suhl processes showing a Ga-YIG sphere subjected to an rf field $h_{1

$ at $f_p = 1.3$GHz and a dc field $H_{dc} \approx 460$ Oe. A pickup coil generates a signal V_s proportional to the time derivative of the transverse magnetization of the sample [166]}

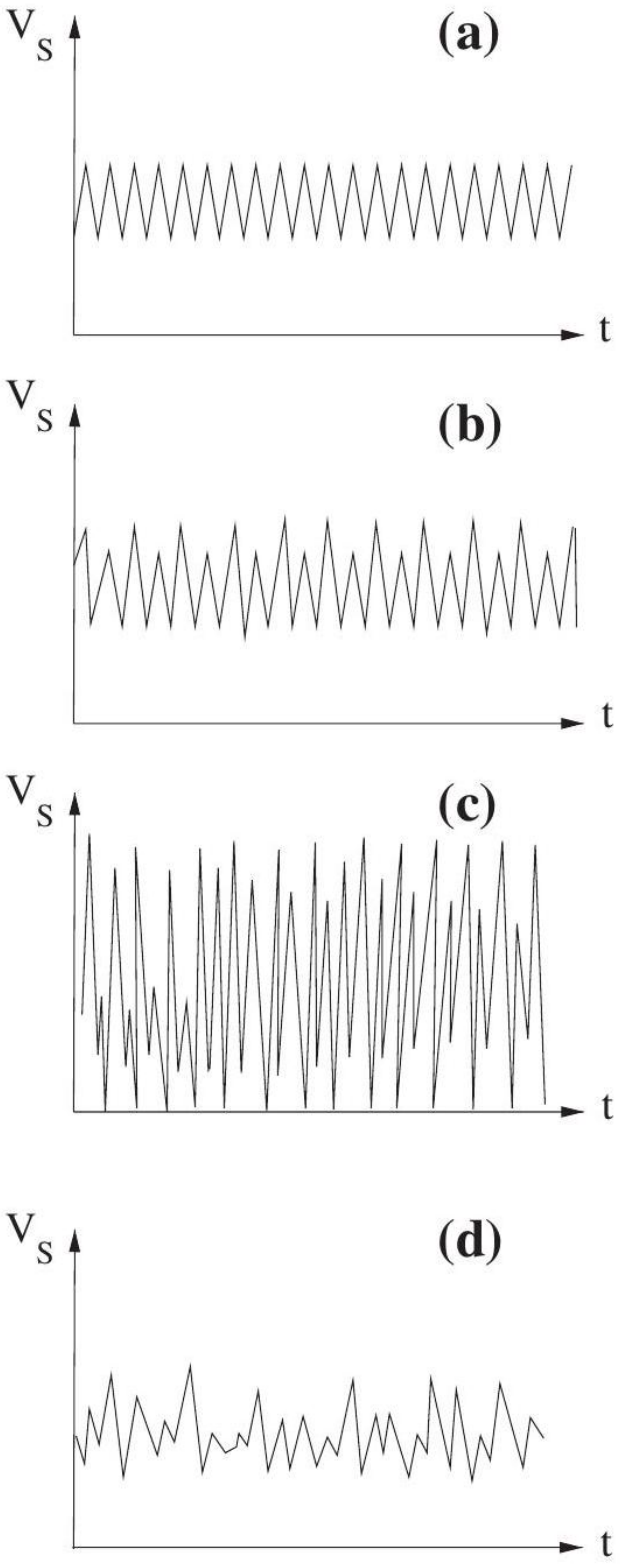


Fig. 8.12. Real-time signals, $V_{s
(t)$, for auto-oscillations: (a) auto-oscillation frequency $f_o \approx$ 16kHz; (b) at higher h, bifurcation of $f_o/2$; (c) bifurcation to $f_o/4$; (d) chaos [166]}

knowing how these variables behave with time and how they vary relative to one another. What Lorenz found was that there is a rich variety of solutions depending upon the parameters that enter the differential equations. In some cases a variable may relax to a stationary value, or "fixed point". For other parameters the solution may jump between stationary values. If one

plots the solutions in a space whose coordinates are the variables themselves then the allowed solutions will, in general, map out a trajectory as a function of time
called the "attractor". In some cases the attractor spans a continuum in this phase space and we call it a "strange attractor". In this case the motion is chaotic.

In our magnetic case the nonlinearity arises from the higher-order terms in the Hamiltonian of the form (8.75) and (8.78). Since the Gibson-Jeffries experiment involves the second-order Suhl process the appropriate nonlinear terms are the 4 -magnon terms

$$\mathcal{H}^{(4)'} = \hbar \sum_{\boldsymbol{k}} \frac{1}{2} S_{0\boldsymbol{k}} \alpha_0^\dagger \alpha_0^\dagger \alpha_{\boldsymbol{k}} \alpha_{-\boldsymbol{k}} + \text{ c.c.} \tag{8.87}$$

Removing the high frequency part of the magnon amplitude with the transformation,

$$\tilde{a}_{\boldsymbol{k}} = \langle \alpha_{\boldsymbol{k}} \rangle \, e^{i\chi_{\boldsymbol{k}}} e^{i(\omega/2)t}, \tag{8.88}$$

where $\chi_{\boldsymbol{k}}$ is a real phase, the equations for these amplitudes become

$$\begin{aligned}
\frac{d\tilde{\alpha}_0}{dt} =& -\left(\eta_0 + i\Delta\omega_0\right)\tilde{\alpha}_0 - i\left(S_{00}\tilde{\alpha}_0^2 + \sum_{\boldsymbol{k}\neq 0} S_{\boldsymbol{k}0}\tilde{\alpha}_{\boldsymbol{k}}^2\right)\tilde{\alpha}_0^* \\
& - i\gamma \sum_{\boldsymbol{k}\neq 0} T_{\boldsymbol{k}_0} \left|\tilde{\alpha}_{\boldsymbol{k}}\right|^2 \tilde{\alpha}_{\boldsymbol{k}} - i\gamma \left(\frac{SN}{2}\right)^{\frac{1}{2}} h \\
\frac{d\tilde{\alpha}_{\boldsymbol{k}}}{dt} =& -\left(\eta_{\boldsymbol{k}} + i\Delta\omega_{\boldsymbol{k}}\right)\tilde{\alpha}_{\boldsymbol{k}} - i\left(S_{\boldsymbol{k}0}\tilde{\alpha}_0^2 + S_{\boldsymbol{k}}\tilde{\alpha}_{\boldsymbol{k}}^2\right)\tilde{\alpha}_{\boldsymbol{k}}^* - i2T_{\boldsymbol{k}0}\left|\tilde{\alpha}_0\right|^2 \tilde{\alpha}_{\boldsymbol{k}}
\end{aligned} \tag{8.89}$$

Rezende et al. [167] have analyzed these equations assuming the $\theta_{\boldsymbol{k}} = 0$ magnon provides the dominant contribution. Figure 8.13 shows the magnon populations n_0 and $n_{\boldsymbol{k}}$ for this "two-mode" approximation. This figure clearly shows the successive period doubling bifurcations.

10.9.3 8.2.8 Optical Processes

Infrared Absorption. There are other mechanisms for exciting magnons which do not involve nonlinearities [168]. For example, consider the antiferromagnet MnF_2. The ground state of the Mn^{2+} ion is an orbital singlet. However, since $S = \frac{5}{2}$, there are six spin states, which in the magnetically ordered crystal are split by the exchange interaction. Those ions which occupy, say, the A sublattice have $M_S = -\frac{5}{2}$ as their ground state, while the ground state for those ions on the B sublattice is $M_S = +\frac{5}{2}$.$A(\boldsymbol{k}, -\boldsymbol{k})$ magnon pair may be excited by the process illustrated in Fig. 8.14. First, the applied radiation induces an electric dipole transition to a high-lying state of one of the Mn^{2+} ions. Since this is a virtual process, it need not conserve energy. However, it must conserve parity, which means that the high-lying state must be opposite

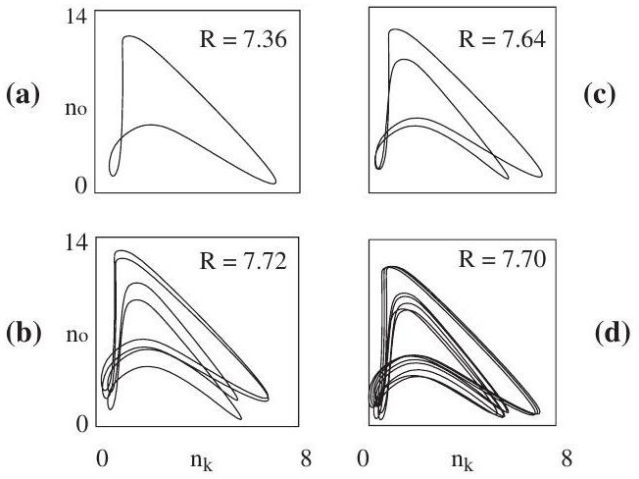


Fig. 8.13. Limit cycles projected onto the $n_{0
-n_{k}$ plane for different values of $R = h/h_c$ where h_c is the critical value for the mode with the lowest threshold [167]}

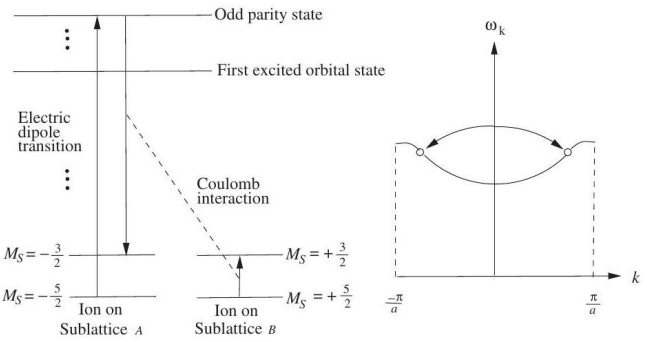


Fig. 8.14. Representation of the mechanism for the excitation to two magnons

in parity from the ground state. Then, as a result of the Coulomb interaction between this ion and a neighboring ion on the opposite sublattice, the excited ion drops down to its first excited spin state ($M_S = -\frac{3}{2}$), while the second ion is simultaneously excited up to its first excited spin state ($M_S = +\frac{3}{2}$). Thus we end up with two spin deviations. Since any pair of ions may be excited, the final state will be a linear combination of such pair excitations. This results in the excitation of two magnons. The absorption coefficient for this process is proportional to the density of final states. Therefore, since the magnon density of states is largest at the Brillouin zone boundary, the absorption is strongest at a frequency approximately twice that of a Brillouin zone magnon, corresponding to an infrared frequency. However, the detailed shape of the absorption is complicated by the interaction between the two output magnons.

Spin-wave Sidebands. In addition to this infrared two-magnon absorption, there is an optical exciton-magnon absorption [169]. In this process the ion on
sublattice A in Fig. 8.14 ends up in its first excited orbital state rather than its first excited spin state. This orbital excitation then propagates throughout the lattice, forming what is known as a Frenkel exciton. The absorption frequency for this process is approximately that of the exciton plus a magnon. In MnF_2 the first excited orbital state lies in the green part of the optical spectrum. The absorption of a single Mn^{2+} ion into this state appears as a very weak magnetic-dipole transition. The exciton-magnon absorption appears as a stronger electric-dipole sideband on the high-frequency side of this "no-magnon" line.

Light Scattering. Magnons may also be excited as the result of light scattering [170]. One mechanism for such scattering is illustrated in Fig. 8.15. The incident photon is virtually absorbed in an electric dipole process. The spin-orbit interaction then produces a spin flip in the excited state and a second photon is virtually emitted leaving the system with a spin excitation. Macroscopically this gives rise to a dielectric function which depends upon the magnetization or spin density,

$$\epsilon^{\alpha\beta} = \epsilon_0 \delta_{\alpha\beta} + \delta\epsilon^{\alpha\beta}(S(\boldsymbol{r}, t)), \tag{8.90}$$

where $\alpha, \beta = (x, y, z)$. The spin-dependent part may be expanded in powers of the spin density

$$\delta\epsilon^{\alpha\beta} = \sum_{\gamma} K_{\alpha\beta\gamma} S_{\gamma}(\boldsymbol{r}, t) + \ldots, \tag{8.91}$$

for a cubic material $K_{\alpha\beta\gamma} = \pm K$ if all the indices are different, and zero otherwise. For the mechanism described in Fig. 8.15K would be proportional to

$$\frac{\lambda \langle 0^* | \beta | 2 \rangle \langle 2 | L_{\gamma} | 1 \rangle \langle 1 | \alpha | 0 \rangle}{(E_1 - E_0)(E_2 - E_0)}. \tag{8.92}$$

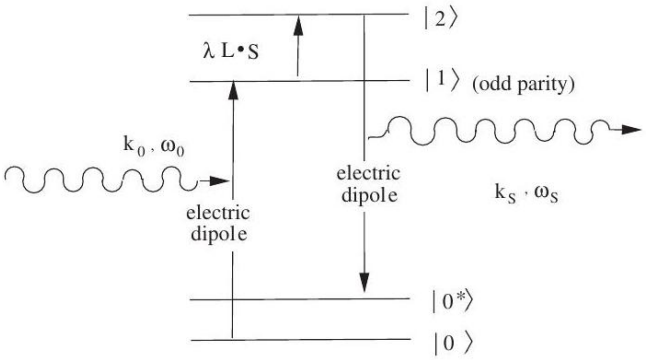


Fig. 8.15. Schematic representation of a perturbation process contributing to magnon scattering. The electric dipole transitions are "virtual" in the sense that they do not conserve energy, i.e., $E_{1
-E_{0} \neq h \omega_{0}$

The static contributions to $\delta\epsilon$ lead to magnetooptic effects while the fluctuating contributions are responsible for the light scattering. Thus, in a cubic material, at least, magnetooptical effects and light scattering are governed by the same parameter K.

To see how the intensity of the scattered light depends upon the magnetic fluctuations consider an incident wave

$$\mathcal{E}^{\alpha}_{\text{inc}}(\boldsymbol{r}, t) = \mathcal{E}^{\alpha}_{0} \exp\left[i\left(\boldsymbol{k}_0 \cdot \boldsymbol{r} - \omega_0 t\right)\right]. \tag{8.93}$$

Inside the medium the total electric field will consist of this incident wave plus the scattered wave,

$$\mathcal{E}^{\alpha}(\boldsymbol{r}, t) = \mathcal{E}^{\alpha}_{\text{inc}}(\boldsymbol{r}, t) + \mathcal{E}^{\alpha}_{\text{scat}}(\boldsymbol{r}, t). \tag{8.94}$$

which is related to the displacement vector by

$$D^{\alpha}(\boldsymbol{r}, t) = \epsilon^{\alpha\beta} \mathcal{E}^{\beta}(\boldsymbol{r}, t). \tag{8.95}$$

Assuming $\delta\epsilon^{\alpha\beta}$ and $\mathcal{E}^{\alpha}_{\rm scat}$ are small, a linearization of Maxwell's equations leads to a driven wave equation for $\mathcal{E}^{\alpha}_{\rm scat}$ where the driving, or source, term involves the product $\delta\epsilon^{\alpha\beta}(\boldsymbol{r},t)\mathcal{E}^{\beta}_{\rm inc}(\boldsymbol{r},t)$. The field at the detector is obtained by multiplying this source term by the so-called Green's function for the wave equation, a function that describes how a disturbance propagates, and integrating over the sample volume and over all time,

$$\mathcal{E}^{\alpha}_{\rm scat}(\boldsymbol{r},t) \sim \left(\frac{\omega_0}{c}\right)^2 \int d\boldsymbol{r}' \int dt G_{\alpha\beta}(\boldsymbol{r},\boldsymbol{r}';t,t')\,\delta\epsilon^{\beta\gamma}(\boldsymbol{r}',t')\,\mathcal{E}^{\gamma}_{\rm inc}(\boldsymbol{r}',t') \tag{8.96}$$

The scattered intensity is found by squaring this field and then performing an ensemble average $\langle\ldots\rangle$ over the fluctuations in $\delta\epsilon$. This results in the correlation function

$$\left\langle \delta\epsilon^{\beta'\gamma'}(\boldsymbol{r},t)\delta\epsilon^{\beta\gamma}(\boldsymbol{r}',t')\right\rangle . \tag{8.97}$$

For a bulk geometry this is a function only of $\boldsymbol{r}-\boldsymbol{r}'$ and $t-t'$. Introducing the time and spatial dependencies in G and $\mathcal{E}_{\rm inc}$ one finds, for example, that the intensity of light scattered with a polarization z, when the incident polarization is x, is

$$\int d\boldsymbol{r} \int dt e^{i(\boldsymbol{q}\cdot\boldsymbol{r}-\Omega t)} \langle S_y(\boldsymbol{r},t)S_y(0,0)\rangle , \tag{8.98}$$

where $\boldsymbol{q}$ is the momentum transfer, $\boldsymbol{q}=\boldsymbol{k}_s-\boldsymbol{k}_0$, and Ω is the excitation energy $\Omega=\omega_s-\omega_0$. The maximum momentum transfer occurs for back scattering,

$$q_{\rm max} = 2\left|\boldsymbol{k}_0\right| = \frac{2n\omega_0}{c}, \tag{8.99}$$

which, for typical optical frequencies is of the order of 3×10^5 cm^{-1}. This is to be compared with the Brillouin zone wave vector, $\boldsymbol{q}_{BZ}\sim 3\times10^8$ cm^{-1}.

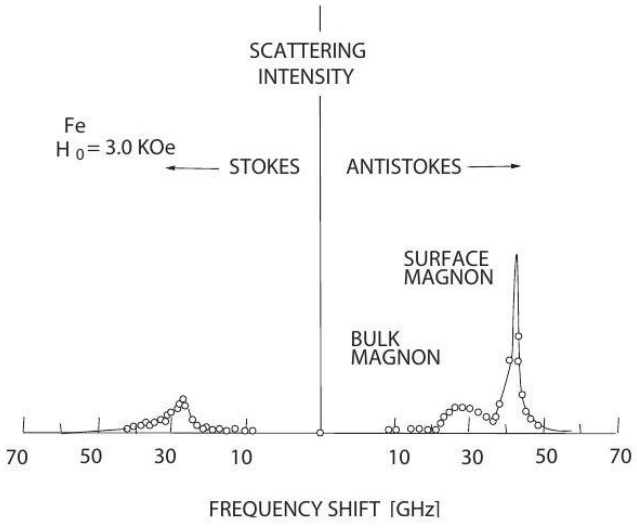


Fig. 8.16. Brillouin spectrum of light scattered from magnons in iron

Thus light scattering probes only the very long-wavelength excitations. If the frequencies exceed 1 cm^{-1} we speak of Raman scattering, while scattering at smaller frequencies is referred to as Brillouin scattering, the distinction being largely based on experimental technique.

If we expand the correlation function $\langle S_y(\boldsymbol{r},t)S_y(0,0)\rangle$ in terms of the spin-wave amplitudes introduced above, we obtain two terms: $\left\langle a_{\boldsymbol{q}}a^{\dagger}_{\boldsymbol{q}}\right\rangle = n_{\boldsymbol{q}}+1$ and $\left\langle a^{\dagger}_{\boldsymbol{q}}a_{\boldsymbol{q}}\right\rangle = n_{\boldsymbol{q}}$. The first corresponds to the creation of a spin wave. The scattered light appears at a frequency $\omega_0-\Omega_{\boldsymbol{q}}$ with an

intensity proportional to $n_{\boldsymbol{q}}+1$ and is referred to as the Stokes line. The second term gives an anti-Stokes line at the frequency $\omega_0+\omega_{\boldsymbol{q}}$.

In Fig. 8.16 we show the result for light scattering from iron. Since the skin depth is relatively small, the intensity of the surface mode is relatively strong. Notice that it only appears on one side (anti-Stokes side in this case) of the exciting frequency. This is a direct consequence of the nonreciprocal nature of surface magnons which we pointed out above.

10.9.4 8.3 High Temperatures

We have seen how effectively the spin-wave concept describes the low-temperature properties of magnetically ordered systems. However, near the transition temperature, and particularly above it, the lifetime of the spin waves becomes too short for them to be a useful concept. At very high temperatures the relaxation-function approach developed in Chap. 5 is applicable. There we found that a strong exchange modulation of the magnetic dipolar field led to a long-time spin-spin correlation function which has the form

$$\langle \mathcal{M}_x(t)\mathcal{M}_x\rangle_{T\to\infty} = \frac{1}{3}NS(S+1)g^2\mu_B^2 \exp\left(i\omega_0 t\right)\exp\left[-\left(2M_2/\omega_{\text{ex}}\right)t\right] \tag{8.100}$$

where M_2 is the Van Vleck second moment and ω_{ex} characterizes the rate at which the exchange interaction modulates the dipolar field. This result was
derived under the assumption that the exchange energy was small in comparison with the Zeeman energy. When the exchange energy is very large, as in the case we are now considering, the off-diagonal elements of the dipole-dipole interaction become important, with the result that the exponent $2M_2/\omega_{\text{ex}}$ in (8.100) is larger by a factor of about $\frac{10}{3}$ (the factor is exactly $\frac{10}{3}$ for a polycrystalline sample with cubic symmetry).

The theory of relaxation phenomena at finite temperatures is a complex subject. Consequently, we shall restrict ourselves to a few qualitative remarks that give some indication of the problems involved (most of this discussion is based on [171]). Let us begin by considering the temperature dependence of the second moment. In the finite-temperature region this is

$$M_2 = -\frac{\left\langle \left[\mathcal{H}'_{\text{dip}}, \mathcal{M}_x\right]^2\right\rangle}{\hbar^2\left\langle \mathcal{M}_x^2\right\rangle} \tag{8.101}$$

where $\mathcal{H}'_{\text{dip}}$ is given by (5.48). After computing this commutator and squaring the result, we are left with a numerator involving terms such as $\left\langle S_i^z S_j^y S_m^z S_n^y\right\rangle$. Making the random-phase approximation and using the high-temperature form of the fluctuation-dissipation theorem,

$$\left\langle \mathcal{M}_x^2\right\rangle = 2\pi k_B T\chi(\boldsymbol{q}=0,\omega=0)$$

we obtain

$$M_2 = \frac{9g^2\mu_B^2\sum_{\boldsymbol{q}}|F(\boldsymbol{q})|^2\left\langle S_z(\boldsymbol{q})S_z(-\boldsymbol{q})\right\rangle^2}{2\pi\hbar^2 k_B T\chi} \tag{8.102}$$

where

$$F(\boldsymbol{q}) = Ng^2\mu_B^2\sum_{\boldsymbol{r}}\frac{3\cos^2\theta-1}{r^3}e^{i\boldsymbol{q}\cdot\boldsymbol{r}} \tag{8.103}$$

From (4.45) we have

$$\langle S_z(\boldsymbol{q})S_z(-\boldsymbol{q})\rangle = \frac{k_B TV}{g^2\mu_B^2}\chi(\boldsymbol{q},\omega=0) \tag{8.104}$$

Converting the sum over $\boldsymbol{q}$ in (8.102) into an integral, we find that for a ferromagnet the second moment varies as $(T - T_c)^{1/2}$. If the linewidth were described by the second moment alone, this would indicate that the paramagnetic resonance line should narrow as the temperature approaches the Curie point. For an antiferromagnet, however, the uniform static susceptibility does not diverge, with the result that the second moment increases as $(T - T_N)^{-1/2}$ where T_N is the Neél temperature.

This is not the complete story, for the exchange frequency is also temperature dependent. The exchange frequency was defined in (6.71) as

$$\langle \Delta\omega(\tau)\Delta\omega \rangle = M_2 \left(1 - \omega_{\text{ex}}^2 \tau^2\right) \tag{8.105}$$

With the same approximations used in obtaining (8.102), this may also be written as

$$\langle \Delta\omega(\tau)\Delta\omega \rangle = \sum_{\boldsymbol{q}} |F(\boldsymbol{q})|^2 \left\langle S_z(\boldsymbol{q},\tau) S_z(-\boldsymbol{q}) \right\rangle^2 . \tag{8.106}$$

Therefore the exchange frequency is given by

$$\omega_{\text{ex}}^2 = -\frac{1}{M_2} \sum_{\boldsymbol{q}} |F(\boldsymbol{q})|^2 \left\langle S_z(\boldsymbol{q}) S_z(-\boldsymbol{q}) \right\rangle \frac{d^2}{d\tau^2} \left\langle S_z(\boldsymbol{q},\tau) S_z(-\boldsymbol{q}) \right\rangle . \tag{8.107}$$

We see that this expression involves the second derivative of the correlation function. Without evaluating this quantity it is possible to say something very important about (8.107), namely, that its numerator does not diverge or go to 0 as the temperature approaches the critical temperature.

To see this let us rewrite the derivative of the correlation function in terms of individual spins as

$$\left\langle \left(d^2/d\tau^2\right) S_i^z(\tau) S_j^z \right\rangle e^{i\boldsymbol{q}\cdot\boldsymbol{r}_{ij}}$$

We now apply a thermodynamic result derived by Kubo [172] which says that this derivative is proportional to $\left\langle \left[\left[\mathcal{H}_0, S_i^z \right], S_j^z \right] \right\rangle$. The consequence of these commutators is that the derivative of the correlation function is nonzero only for spins very close to one another. This, in turn, implies that the correlation function in the numerator of (8.107) involving the second derivative shows no drastic temperature dependence as the critical temperature is approached, since only short-range static correlations are involved, and these do not change suddenly near the critical point. These considerations lead us to the result that the square of the exchange frequency also goes to 0 as $(T - T_c)^{1/2}$. Therefore, if we assume that the method of moments provides an adequate description of a ferromagnet in its paramagnetic region, we are led to the conclusion that the resonance linewidth remains exchange narrowed as the temperature decreases toward the Curie point, decreasing as $(T - T_c)^{1/4}$. Evidence of such behavior is indicated for the ferromagnet $CrBr_3$ in Fig. 8.17.

The above arguments assume zero applied field. In a finite field the divergences are masked, and the simple power-law dependencies are not expected to hold in the immediate vicinity of T_c. Since resonance linewidths are measured in finite fields of appreciable magnitude, we may expect considerable departure from these simple relations. In particular, the ferromagnetic-resonance linewidth does not go to 0 at the critical temperature but in typical situations decreases by only about an order of magnitude from its high-temperature value (explicit calculations of the moments, taking the finite field into account, are reported in [173]). In an antiferromagnet the concept of exchange narrowing breaks down. Our simple ideas predict that the linewidth should diverage as $(T - T_N)^{1/4}$. The experimental data on MnF_2, shown in Fig. 8.18 indicate that the zero-field linewidth does, in fact, increase as $(T - T_N)^{-3/8}$.

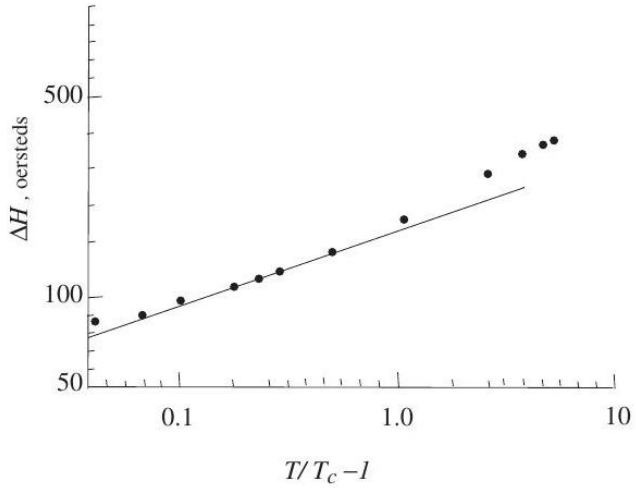


Fig. 8.17. Linewidth as a function of reduced temperature in $\mathrm{CrBr _{3}$. The solid line is proportional to $[(T/T_c) - 1]^{1/4}$ [171]}

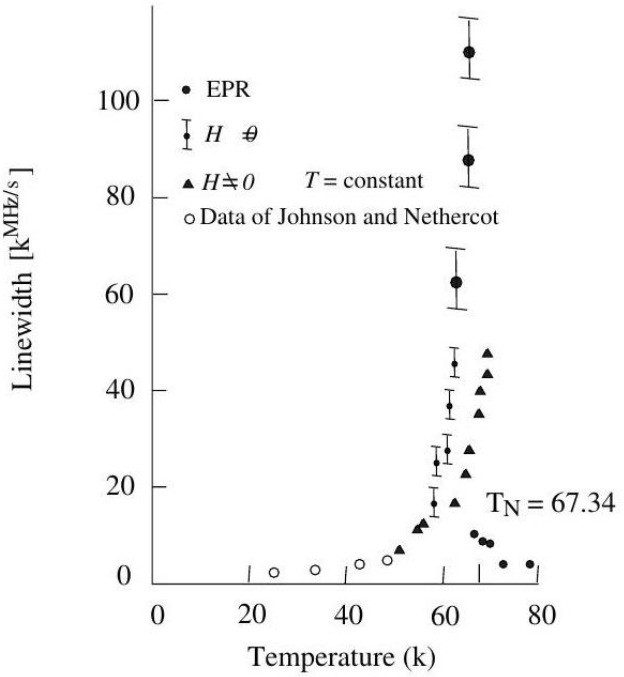


Fig. 8.18. Linewidth as a function of temperatures in $\mathrm{MnF _{2}$ [174]}

10.9.5 8.4 Micromagnetics

The magnetostatic modes discussed above have spatial variations that are large compared with interatomic spacings. Therefore, we did not include exchange. On the other hand, spinwaves are capable of describing very short wavelength excitations in which exchange plays an important role. However, these modes are small amplitude excitations relative to the saturated ferromagnetic (or antiferromagnetic) state. How do we describe large changes in the magnetization, such as a complete magnetic reversal? Micromagnetic modelling is particularly useful for such cases. The principles of micromagnetics were first formulated by Landau and Lifshitz

[175] and later developed into a self-consistent theory by Brown [176]. This theory has been applied to
many situations, particularly by Zhu [177]. In this theory the magnetic system is described by the magnetization at a position i by the vector $\boldsymbol{M}_i = M_S \boldsymbol{m}_i$ where M_S is the saturation magnetization. The dynamics of the magnetization are governed by equations of the form (8.1) at every point i. Since the fields $\boldsymbol{H}_i$ arise from all the other points this leads to coupled, nonlinear differential equations for the magnetization at every point in the system. In general, analytic solutions do not exist. Early work was confined to one-dimensional models and was often based on linear approximations. The introduction of modern computers have made it possible to numerically solve realistic micromagnetic problems.

The first step in setting up a micromagnetic calculation is to discretize the magnetic system with a three-dimensional mesh. The energy density is then calculated for each discrete point. The exchange energy density, for example, has the continuum form introduced at the beginning of this chapter,

$$E_{\text{ex}}(\boldsymbol{r}) = 2A \left[\left(\frac{\partial \boldsymbol{m}}{\partial x} \right)^2 + \left(\frac{\partial \boldsymbol{m}}{\partial y} \right)^2 + \left(\frac{\partial \boldsymbol{m}}{\partial z} \right)^2 \right]. \tag{8.108}$$

In discrete form this becomes

$$E_{\text{ex}}\left(\boldsymbol{r}_i\right) = -\frac{2A}{M_S^2 a^2} \boldsymbol{M}_i \cdot \sum_{nn} \boldsymbol{M}_j, \tag{8.109}$$

where a is the distance between neighboring sites in the mesh. Energy densities for the dipole-dipole (magnetostatic) interaction, anisotropy, and the Zeeman interaction are similarly introduced giving a total energy density, $E_{\text{tot}}\ (\boldsymbol{r}_i)$. The effective magnetic field acting on the magnetization at the i^{th} site is

$$\boldsymbol{H}_i = -\frac{\partial E_{\text{tot}}\left(\boldsymbol{r}_i\right)}{\partial \boldsymbol{M}_i}. \tag{8.110}$$

Introducing reduced units $\tau = \gamma M_S t$ and $\boldsymbol{h}_i = \boldsymbol{H}_i / M_S$ the Landau-Lifshitz equation takes the form

$$\frac{d\boldsymbol{m}_i}{dt} = -\gamma \boldsymbol{m}_i \times \boldsymbol{h}_i. \tag{8.111}$$

Not only did modern computers make micromagnetic calculations practical, but the development of the magnetic force microscope (MFM) made it possible to validate this theoretical approach. This instrument is described in the next section.

8.4.1 Magnetic Force Microscope

In 1986 Binning, Quate, and Gerber [178] showed how the scanning tunneling microscope could be employed to measure extremely small forces. It was quickly realized that this "atomic force microscope" (AFM) could be modified
to image magnetostatic fields emanating from a magnetic sample. The magnetic force microscope consists of a miniature cantilever beam clamped at one end that acts as a force-sensing spring (Fig. 8.19). A tip is integrated into the single-crystal silicon cantilever by etching techniques and then coated with a magnetic material. The deflection or frequency of the cantilever is monitored by reflecting a laser beam off the back of the cantilever and sensing the beam with a split photodiode detector. The tip is scanned across the sample by means of a piezoelectric scanner.

Row A of Fig. 8.20 shows an MFM image of a 30 nm thick permalloy film. This film was prepared by sputtering permalloy onto a NiO substrate and then patterning it into a 20×10 nm rectangle by lithography. The antiferromagnetic NiO produces exchange coupling which stabilizes the domain structure during the MFM measurement.

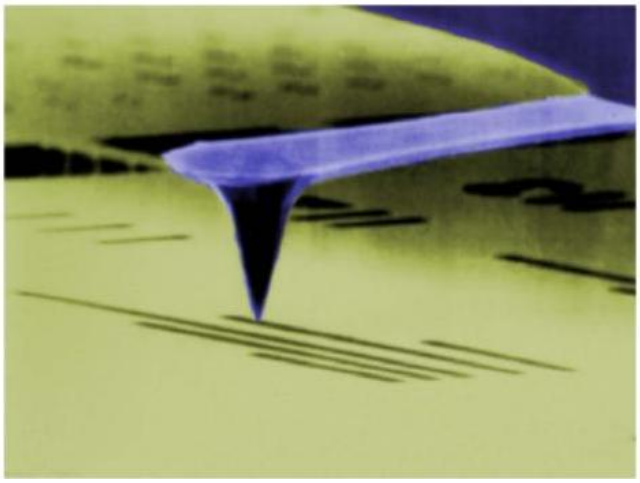


Fig. 8.19. Microphotograph of cantilevered probe tip used in a magnetic force microscope (C.F. Quate)

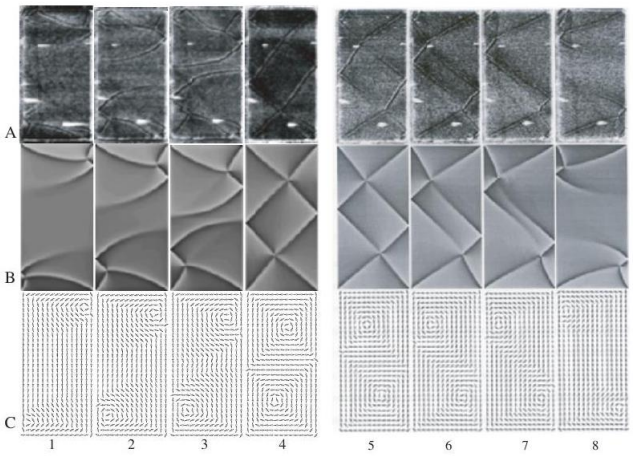


Fig. 8.20. The magnetization switching process of a permalloy thin film. Row A is the magnetic force microscope images. Row B is the calculated magnetic charge density from micromagnetic simulation. Row C is the simulated magnetization patterns [179]

The external magnetic field is initially applied downward along the long axis of the rectangular film. Under this applied field, the thin film is in a quasisingle domain state, as shown in state A1 in Fig. 8.20. There are small curved edge domains at both ends of the film because the applied field is not strong enough to drive them completely out of the film. The existence of these curved edge domains also help to minimize the magnetostatic energy that would otherwise arise from the magnetic poles on the film surfaces. The applied field is then gradually reduced to zero and then increased in the opposite direction to reverse the magnetization of the thin film. The state A8 in Fig. 8.20 shows the fully reversed state. The domain configuration in this state is symmetric to the initial state around the center line of the film (along the short axis). The states A2 to A8 in Fig. 8.20 show the subsequent domain images observed along

the reversal process between state A1 and A8. In A2 and A3, the curved domain walls that extend from the two ends to the center of the thin film move towards the center and close to each other. In A4, a diamond shaped domain is formed at the center of the film, resulting in a flux-closure multi-domain structure in the permalloy thin film. With further change of the external field, the domains that are parallel to the direction of the applied field expand at the expense of the domains that are antiparallel to the applied field. This is clearly seen in the states A5, A6 and A7 in Fig. 8.20. As a result, the diamond shaped domain in the center narrows. The two edges of this center domain eventually cross each other, move away from the center of the thin film and away from each other, and reaches the final reversed state A8.

States B1 to B8 in Fig. 8.20 are the domain images obtained from the micromagnetic simulation. These are gray scale images which map the magnetic charge density calculated from the simulated magnetization distribution of the permalloy thin film. Since the MFM detects the magnetic field gradient which results from the magnetic charges on the permalloy thin film, the MFM images closely resemble the magnetic charge density distribution. Comparing the MFM images (A1 to A8) and the simulated domain configurations (B1 to B8) in Fig. 8.20, we can clearly see that excellent agreement between them. This strongly validates the ability of the micromagnetic model to simulate the domain configurations during the magnetic reversal process.

The vector plots C1 to C8 in Fig. 8.20 show the corresponding magnetization distribution of the thin film, obtained directly from the micromagnetic simulation. The cell size is $10 \times 10 \times 12.5$ nm. The arrows in these plots indicate the directions of the magnetization. The ability to give direct information on the magnetization distribution is one of the powerful features of micromagnetic simulation. This enables us to better understand the reversal process. As we can see from C1, the two edge domains observed in the initial state are two vortices. The center part of the thin film is uniformly magnetized downward. The two vortices are formed at the two ends in order to reduce the free magnetic charges otherwise would occur at the two end surfaces. As the applied field changes, the top vortex moves toward the left, while the bottom vortex moves towards the right. As a result, the net magnetization in
the downwards direction decreases. This is clearly seen from C 1 to C 4 . In C 4 a near-perfect flux closure structure is formed, with very little net magnetization in the downwards direction. Further increase of the applied field in the upward direction continues to drive the motions of the two vortices. In C5 the net magnetization has been reversed to the upward direction. This is followed by C6 and C7 where the vortices are driven further to the side. Eventually the center part of the magnetization is reversed completely to the upwards direction, and the thin film reaches the reversed state C8.

8.4.2 Phenomenological Damping

In our preface to this section we said one of the features of micromagnetic modeling was its applicability to time-dependent phenomena. Figure 8.21 shows the micromagnetic response based on (8.111) of a permalloy film to a field of 250 Oe suddenly applied in the reverse direction. At $t = 0$ the Zeeman energy density is $M_S H$. At 1.5 ns the magnetization has predominantly reversed, giving a Zeeman energy density $-M_S H$. The variations in the magnetization shown in Fig. 8.21b indicate that this change in Zeeman energy density has gone into exchange energy density. That is, the energy has spread throughout the spin-wave spectrum shown in Fig. 8.2.

Since (8.111) describes a precession of the magnetization $\boldsymbol{m}_i$ about the local field $\boldsymbol{h}_i$ the spins in Fig. 8.21b are precessing about their local fields. However, we know that the magnetization eventually relaxes into the direction of $\boldsymbol{h}_i$. It does this by transferring energy to the lattice. This relaxation effect is incorporated by adding a phenomenological damping term. Landau and Lifshitz did this by introducing a term that can be written in the form [175]

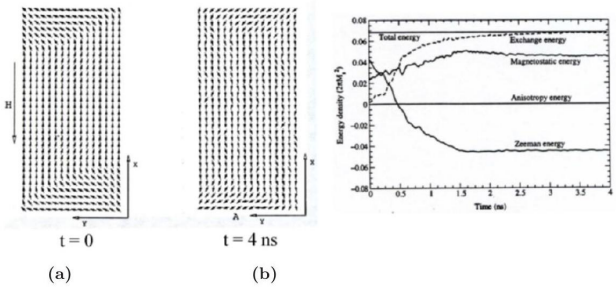


Fig. 8.21. Micromagnetic description of the magnetization reversal in a permalloy film including only precessional terms [180]

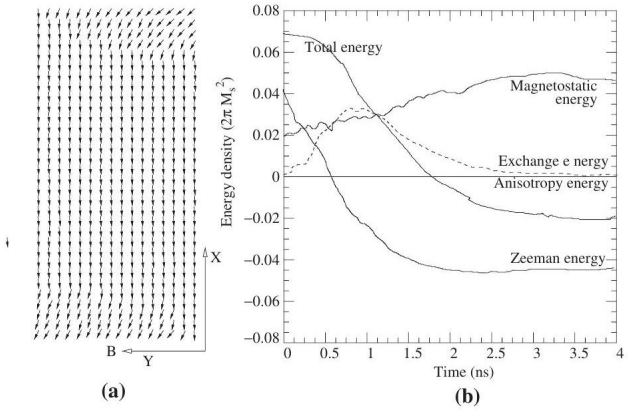


Fig. 8.22. Micromagnetic description of the magnetization reversal in a permalloy film including phenomenological damping [180]

$$-\frac{\lambda}{M^2}\boldsymbol{M}\times\boldsymbol{M}\times\boldsymbol{H},$$

which is perpendicular to the plane of $\boldsymbol{M}\times\boldsymbol{H}$. In 1955 Gilbert developed a damping term based on a variational approach [181] which has the form

$$-\frac{\alpha}{M}\boldsymbol{M}\times\frac{\partial\boldsymbol{M}}{\partial t}.$$

The Landau-Lifshitz damping term may be shown to be identical to Gilbert's if one modifies the gyromagnetic term in the Landau-Lifshitz equation to have the coefficient $\gamma/\left(1+\alpha^2\right)$. That is, in order to describe Gilbert's damping in terms of a triple cross-product the gyromagnetic ratio must become a function of the damping. The resulting Landau-Lifshitz-Gilbert equation becomes

$$\frac{d\boldsymbol{m}_i}{d\tau}=\gamma^*\boldsymbol{m}_i\times\boldsymbol{h}_i-\alpha\boldsymbol{m}_i\times\left(\boldsymbol{m}_i\times\boldsymbol{h}_i\right), \tag{8.112}$$

where $\gamma^*=\gamma/\left(1+\alpha^2\right)$. If the micromagnetic calculation for the permalloy rectangle is now carried out using (8.112) with a damping parameter $\alpha=0.005$ we obtain the results in Fig. 8.22. We see that the exchange energy eventually decreases to zero. The "total" energy is not conserved because it does not include the energy that has been transferred to the lattice.

10.9.6 8.5 Metals

Let us now consider the response of a ferromagnetic metal to a time-dependent field. The Fermi liquid formulation, which we used in the last chapter to study the response of nonmagnetic metals, is not rigorous inthis case. This is because

Fermi liquid theory rests on the assumption that the quasiparticle excitation spectrum lies close to the Fermi surface. In a ferromagnet this may not be the case. For example, when the spin of a quasiparticle at the Fermi surface is reversed, the resulting quasiparticle state will not, in general, lie close to the Fermi surface and will therefore have a finite decay time. Therefore, we shall consider the microscopic model discussed in Chaps. 4 and 5 which was characterized by the delta-function interaction. In Chap. 5 we obtained the generalized susceptibility,

$$\chi_{-+}(\boldsymbol{q},\omega)=\frac{2\mu_B^2\Gamma(\boldsymbol{q},\omega)}{1-2I\Gamma(\boldsymbol{q},\omega)}, \tag{8.113}$$

where

$$\Gamma(\boldsymbol{q},\omega)=\frac{1}{V}\sum_{\boldsymbol{k}}\frac{n_{\boldsymbol{k}\uparrow}-n_{\boldsymbol{k}+\boldsymbol{q}\downarrow}}{\hbar\omega-(\tilde{\epsilon}_{\boldsymbol{k}\uparrow}-\tilde{\epsilon}_{\boldsymbol{k}+\boldsymbol{q}\downarrow})}. \tag{8.114}$$

Adding the constant I to the energy does not change the energy difference in $\Gamma(\boldsymbol{q},\omega)$ and enables us to write

$$\tilde{\epsilon}=\epsilon_{\boldsymbol{k}}+\frac{nI}{2}-\sigma\frac{nI}{2}\zeta, \tag{8.115}$$

where $\zeta=(n_\uparrow-n_\downarrow)/n$ is the relative magnetization and $\sigma=\pm1$. For parabolic energy bands $\Gamma(\boldsymbol{q},\omega)$ may be evaluated exactly for $T=0$:

$$\begin{aligned}\Gamma(\boldsymbol{q},\omega)=-\frac{N(\epsilon_F)}{4\tilde{q}}\Bigg\{&A(\tilde{q}-\tilde{\omega}/\tilde{q})+\frac{1}{2}\left[A^2+(\tilde{q}-\tilde{\omega}/\tilde{q})^2\right]\ln\frac{\tilde{q}-\tilde{\omega}/\tilde{q}+A}{\tilde{q}-\tilde{\omega}/\tilde{q}-A}\\&+B(\tilde{q}+\tilde{\omega}/\tilde{q})+\frac{1}{2}\left[B^2-(\tilde{q}+\tilde{\omega}/\tilde{q})^2\right]\ln\frac{\tilde{q}+\tilde{\omega}/\tilde{q}+B}{\tilde{q}+\tilde{\omega}/\tilde{q}-B}\Bigg\}\end{aligned} \tag{8.116}$$

where

$$\begin{aligned}&\tilde{q}=a/2q_F,\quad \tilde{\omega}=(\omega-nI\zeta)/4\epsilon_F,\\&A=(1+\zeta)^{1/3},\quad B=(1-\zeta)^{1/3}.\end{aligned}$$

Since $\ln z=\ln r+i\theta$ where $z=re^{i\theta}$, the appearance of the logarithms in (8.116) gives rise for $\zeta<1$ to four regions in the $\omega-q$ plane. Three of these are characterized by $\Gamma''(\boldsymbol{q},\omega)\neq0$ corresponding to single-particle spin-flip, or "Stoner", excitations. These three regions are indicated by the shaded portions of Fig. 8.23. In the fourth region $\Gamma''(\boldsymbol{q},\omega)=0$. Therefore, since

$$\chi''_{-+}(\boldsymbol{q},0\omega)=\frac{2\mu_B^2\Gamma''(\boldsymbol{q},\omega)}{\left[1-2I\Gamma'(\boldsymbol{q},\omega)\right]^2+4I^2\Gamma''(\boldsymbol{q},\omega)^2} \tag{8.117}$$

in the limit $\Gamma''(\boldsymbol{q},\omega)\to0$,

$$\chi''(\boldsymbol{q},\omega)\to\frac{\mu_B^2}{I}\delta\left[1-2I\Gamma'(\boldsymbol{q},\omega\right]. \tag{8.118}$$

From the property of the delta function,

$$\delta[f(\omega)] = \delta(\omega - \omega_q) \left|\partial f/\partial\omega\right|^{-1}_{\omega_q},$$

where $\omega_{\boldsymbol{q}}$ is defined by $f(\omega_{\boldsymbol{q}}) = 0$, we see that the imaginary part of the susceptibility contains poles at $\omega = \omega_{\boldsymbol{q}}$. These correspond to collective excitations, or spin waves.

To find the dispersion relation for these spin waves we expand $\Gamma'(\boldsymbol{q}, \omega)$ for small q and ω and set the result equal to $1/2I$. The result is [182]

$$\begin{aligned}\hbar\omega(q) =& \frac{1}{N_\uparrow - N_\downarrow} \\ &\times \sum_{\boldsymbol{k}} \left[\frac{1}{2}\left(n_{\boldsymbol{k}\uparrow} + n_{\boldsymbol{k}\downarrow}\right)\left(\boldsymbol{q}\cdot\frac{\partial}{\partial\boldsymbol{k}}\right)^2 \epsilon_{\boldsymbol{k}} - \frac{N\left(n_{\boldsymbol{k}\uparrow} - n_{\boldsymbol{k}\downarrow}\right)}{I\left(N_\uparrow - N_\downarrow\right)}\left(\boldsymbol{q}\cdot\frac{\partial\epsilon_{\boldsymbol{k}}}{\partial\boldsymbol{k}}\right)^2\right]\end{aligned} \tag{8.119}$$

This has the same quadratic dependence upon q that we found for the spinwave spectrum of a ferromagnetic insulator. The first term in (8.119) corresponds to the additional kinetic energy that the electrons develop as they change their spin directions to keep up with the change in the macroscopic spin direction of the spin wave. It can be shown [183] that this is always positive and vanishes for a filled band. The second term corresponds to a reduction in this kinetic energy due to a tilting of the electron spin out of the plane in which the macroscopic magnetization varies.

From Fig. 8.23 we see that for wave vectors for which undamped spin waves exist there will be contributions to $\chi''(\boldsymbol{q}, \omega)$ from the nonzero values of $\Gamma''(\boldsymbol{q}, \omega)$ corresponding to virtual Stoner excitations. If we imagine a vertical cut through the $\omega - q$ plane at increasingly larger wave vectors, the susceptibility behaves [184] as shown in Fig. 8.24. The spin waves becomes heavily damped as they approach and merge into the Stoner excitation region. This limits their experimental identification to relatively small wave vectors.

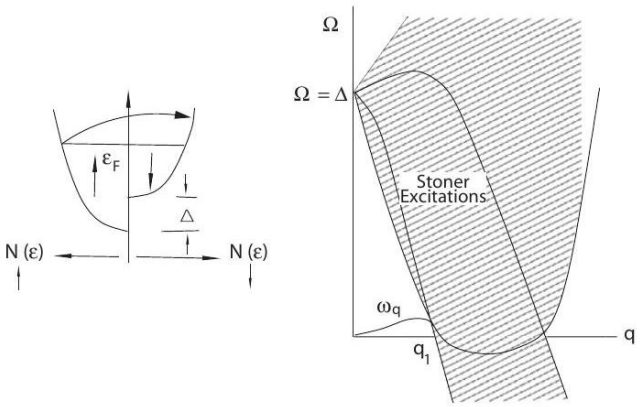


Fig. 8.23. Regions of \$\Omega-\boldsymbol{q}
\$ space showing Stoner excitations and spin waves. Here $\Omega = \omega/4\epsilon_F$ and $\Delta = nI\zeta/4\epsilon_F$ is the Stoner gap}

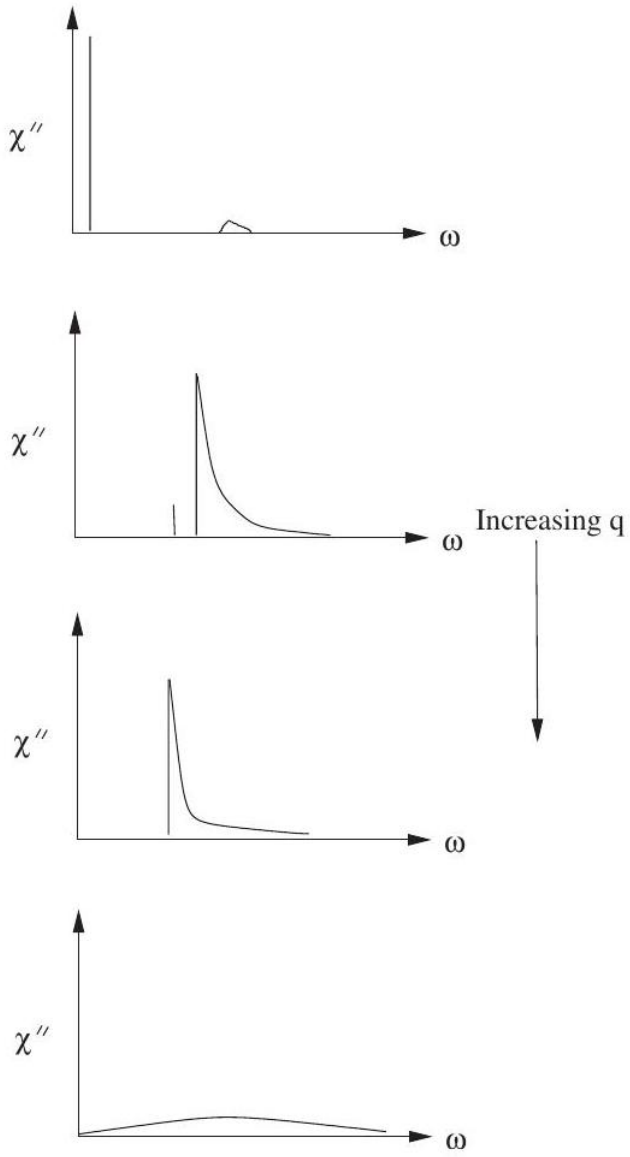


Fig. 8.24. Behavior of the spectral function $\chi^{\prime \prime}(q, \omega)$ as $\boldsymbol{q}$ increases toward $\boldsymbol{q}_1$, in Fig. 8.23}

In most ferromagnetic metals the low-temperature magnetization is found to follow the Bloch $T^{3/2}$ law. This is usually taken as evidence of the importance of the spin-wave modes. However, in weak itinerant ferromagnets, such as $ZrZn_2$, the Stoner excitations may dominate, producing a T^2 contribution to the magnetization.

The first direct evidence of spin waves in metals was made by Lowde, who observed the inelastic scattering of neutrons from iron [185]. Figure 8.25 shows the spin-wave dispersion relations for iron and nickel. In addition to confirming the quadratic dependence on q, these results show an interesting aspect of these spin waves, namely, their existence well above the Curie temperature.

The existence of spin waves above T_c has been confirmed by large scale computer simulation [188]. These simulations indicate that the spin waves will be propagating, as opposed to dissipative, when their wavelengths are smaller than the correlation length which measures the short range order. For iron at $T = 1.1T_c$ the correlation length is approximately twice the bcc lattice parameter, i.e., $2a$. Therefore we only have propagating spin waves for wavevectors $k \geq 0.77$Å.

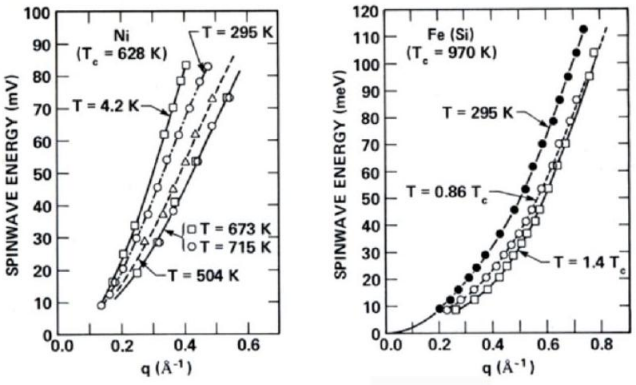


Fig. 8.25. Spin-wave dispersion relations in iron [186] and nickel [187]. The iron sample was alloyed with silicon to bring the Curie temperature down to a convenient value

Seavey and Tannenwald observed the microwave excitation of standing spin waves in thin films of permalloy [189], and light scattering has also proven useful in studying spin waves in metals [190].

Problems

8.1.
(a) Express $|\boldsymbol{M}|$ and M_z in terms of the number of magnons, $n_{\boldsymbol{k}}$. How many magnons are involved in changes of these quantities?
(b) Which of the following magnon scattering processes (Fig. 8.26) represent microscopic mechanisms for the phenomenological damping parameter α ?

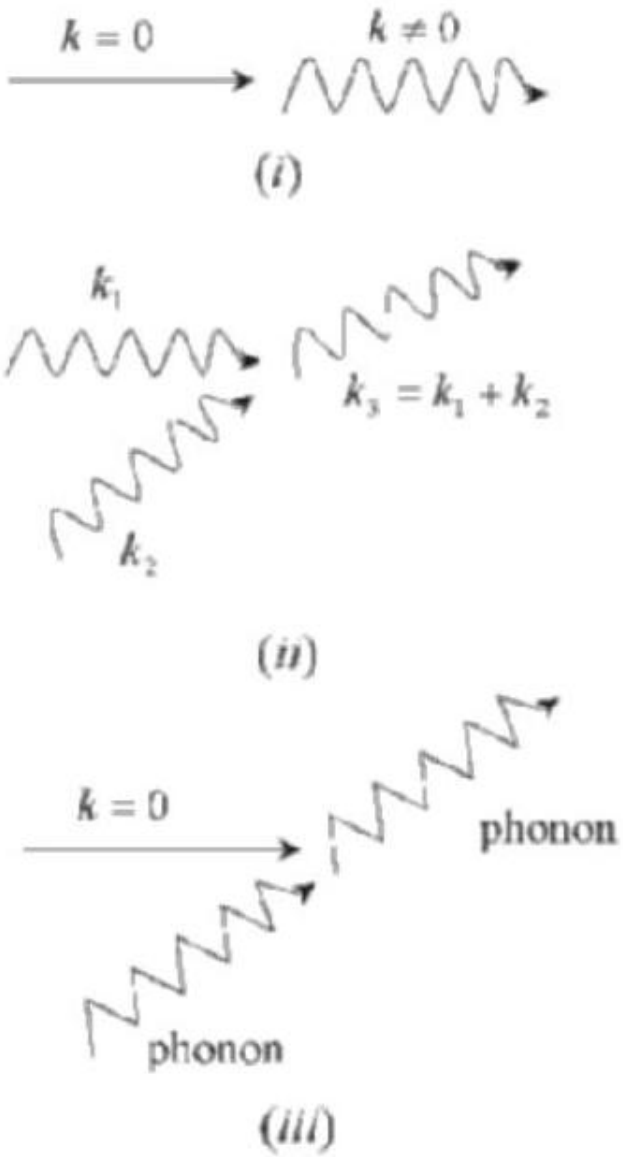


Fig. 8.26. Relaxation processes that may, or may not, be described by " α "

8.2. Consider 1-magnon scattering described by the Hamiltonian

$$\mathcal{H} = \sum_k g_{0k} \left(a_k^+ a_0 + a_0^+ a_k \right).$$

Using perturbation theory calculate the relaxation rate of the uniform precession mode and determine the frequency dependence of the FMR linewidth.
8.3. From the kinetic theory of gases the thermal conductivity is given by

$$\kappa = \frac{1}{3} C \eta \Lambda$$

where C is the heat capacity per unit volume, v is the particle velocity, and Λ the mean free path. Calculate the thermal conductivity of a "magnon gas" where $C(\boldsymbol{k}) = \frac{\partial}{\partial T} \left(n_k \varepsilon_k \right), v$ is the group velocity $v = \frac{\partial \varepsilon_k}{\partial_k}$, and Λ is independent of $\boldsymbol{k}$. Use the same approximations in obtaining the heat capacity (8.74).

10.10 Thin Film Systems

The development of thin film deposition systems, such as molecular beam epitaxy and sputtering during the 1970s and 1980s and pulsed laser deposition in the 1990s made it possible to synthesize novel materials systems which exhibited new phenomena. Quantum wells, for example, led to the quantum Hall effect.

Similarly, thin magnetic films led to the giant magnetoresistance (GMR) effect. In this chapter we shall explore the concepts associated with magnetic multilayer structures. We shall begin with a single interface and then consider magnetic layers separated by thin conducting and insulating layers.

10.10.1 9.1 Interfaces

9.1.1 Exchange Bias

Just as there is exchange between the sublattices of an antiferromagnet or ferrimagnet, we expect there to be exchange across the interface between two dissimilar magnetic materials. Consider, for example, the interface between the [111] plane of antiferromagnetic CoO and Co metal. The [111] plane of CoO corresponds to one of the sublattices. Therefore, in the antiferromagnetic state the spins on this plane are all parallel. Suppose this antiferromagnet is in proximity with a ferromagnet at a temperature T such that $T_N < T < T_c$ and is then cooled below T_N. This aligns the spins on the surfaces of the antiferromagnet with the ferromagnet. Below T_N the anisotropy of the antiferromagnet "locks in" the spin configuration. The ferromagnet now "sees" a unidirectional exchange from the antiferromagnet. Figure 9.1 shows the hysteresis loop of Co particles coated with CoO . The loop is shifted by 2 Oe . This is referred to as exchange bias. This is a very important phenomenon because it may be used to bias ferromagnetic films for a variety of applications.

To describe this effect let us assume that the exchange across the interface between the antiferromagnetic spins and the ferromagnetic spins may be expressed in terms of an exchange energy density (per unit area) as $J\cos(\alpha - \beta)$,

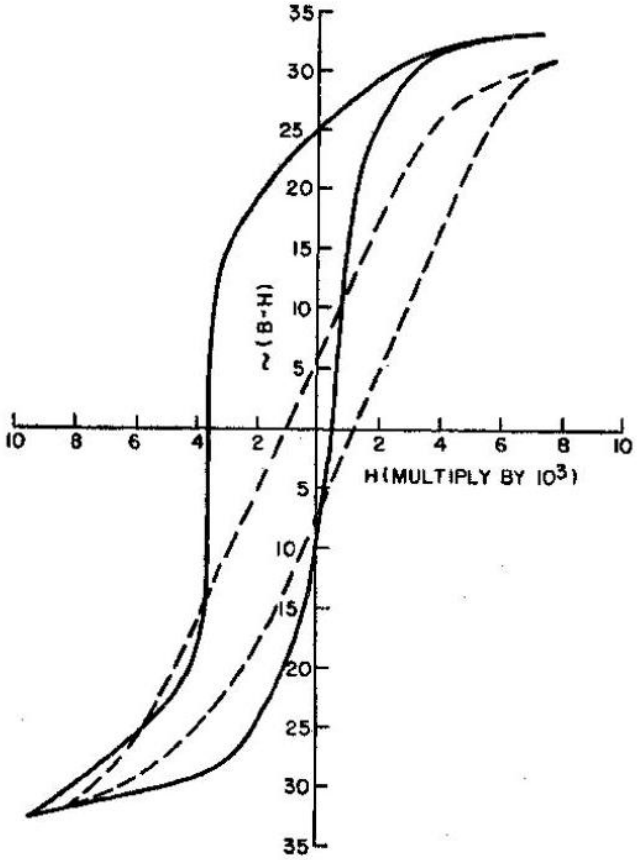


Fig. 9.1. Hysteresis loops of fine oxide-coated particles of cobalt taken at $77^{\circ}{\circ}\mathrm{K}$. The dashed lines show the hysteresis loop when the material is cooled in the absence of a magnetic field. The solid lines show the hysteresis loop when the material is cooled in a saturating magnetic field [191]}

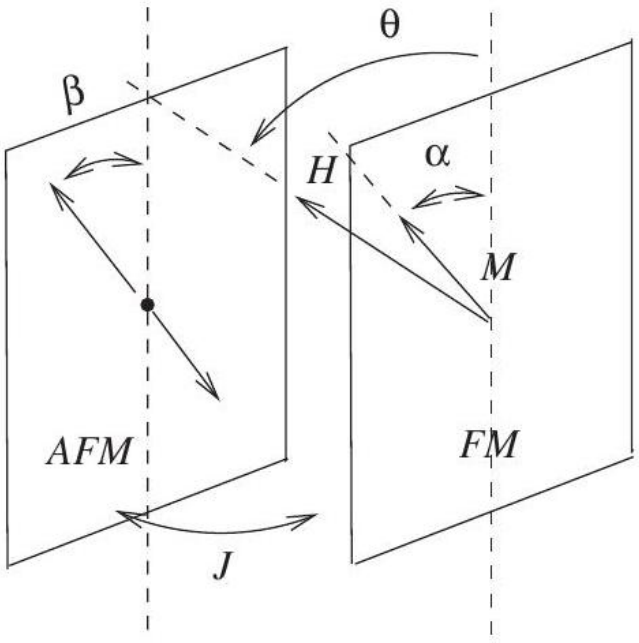


Fig. 9.2. Definition of angles used in describing the coupling between an antiferromagnet (AFM) and a ferromagnet (FM)

where the angle α is the direction of the magnetization and the angle β is the direction of the antiferromagnetic axis (see Fig. 9.2). If an external field H is applied at an angle θ the Zeeman energy density is $-HMt_F\cos(\theta-\alpha\ )$, where t_F is the thickness of the ferromagnet. If we add an anisotropy term to the ferromagnet then the problem of determining the direction of the magnetization of the ferromagnet is identical to the Stoner-Wohlfarth model discussed in Sect. 4.6, except the hysteresis loop is now shifted by an amount

$$H_E = \frac{J}{Mt_F} \tag{9.1}$$

While this simple model seems to describe this effect, there are quantitative discrepancies. First of all, we would expect the interface exchange to be comparable to the exchange in the materials on either side. However, the observed values of H_E lead to values of J which are several orders of magnitude too small. Secondly, H_E is found to depend upon the thickness of the antiferromagnet, t_{AF}. In particular, H_E vanishes for antiferromagnetic thicknesses less than a critical value.

The second observation suggests that perhaps the assumption that the antiferromagnetic spin structure remains rigid is not valid. Antiferromagnets do exhibit domains. These domains are separated by walls very similar to those found in ferromagnets. The energy per unit wall area of an antiferromagnetic wall is given by

$$\sigma_w = \sqrt{A_{AF}K_{AF}} \tag{9.2}$$

where A_{AF} and K_{AF} are the exchange and anisotropy parameters in the antiferromagnet, respectively. Let us assume that as the magnetization of the ferromagnet is pulled around in the plane of the film by an external field this induces a twist in the antiferromagnetic through an angle β. The energy density is then $\sigma_w(1-\cos\beta)$. The total energy density is now

$$\varepsilon = -HMt_F\cos(\theta-\alpha) + J\cos(\beta-\alpha) + \sigma_w(1-\cos\beta) \tag{9.3}$$

The equilibrium equations are

$$\frac{\partial\varepsilon}{\partial\alpha} = 0, \quad \frac{\partial\varepsilon}{\partial\beta} = 0$$

In a reverse field ($\theta=\pi$) where the magnetization has rotated half way around ($\alpha=\frac{\pi}{2}$) these equations become

$$\begin{aligned} -HMt_F + J\cos\beta &= 0 \\ -J\cos\beta + \sigma_w\sin\beta &= 0 \end{aligned} \tag{9.5}$$

Solving for $\cos\beta$ gives

$$\cos\beta = \sqrt{\frac{\sigma_w^2}{J^2+\sigma_w^2}} \tag{9.6}$$

The value of the exchange deduced from a hysteresis loop in this case is

$$J_{h\ell} = Mt_F H_e \tag{9.7}$$

where H_e is the field required to drive the magnetization to zero,

$$H_e = \frac{J}{Mt_F}\cos\beta \tag{9.8}$$

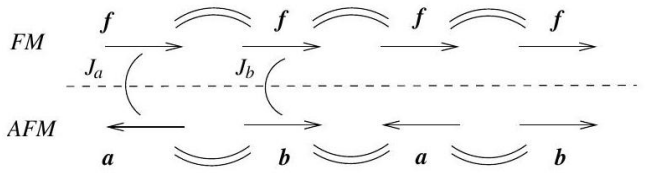


Fig. 9.3. Exchange interactions at the interface between an antiferromagnetic (AFM) and a ferromagnet (FM)

Therefore

$$J_{h\ell} = \frac{J\sigma_w}{\sqrt{J^2 + \sigma_w^2}} \tag{9.9}$$

Thus, if the antiferromagnet is allowed to twist, the exchange deduced from a hysteresis loop measurement is renormalized by the wall energy. In particular, the loop can never be shifted by more than σ_w, no matter how large the direct exchange J. There are other mechanisms that can lead to a reduced exchange bias. Malozemoff [192], for example, has pointed out that interface roughness will produce a random magnetic field which, according to Imry and Ma [193], will lead to domains in the antiferromagnet at the interface.

There is independent evidence for the role of interface roughness in exchange bias. Namely, the linewidth of the ferromagnetic resonance (FMR) associated with a ferromagnetic/antiferromagnetic (FM/AFM) bilayer is an order of magnitude larger than the linewidth without exchange bias and increases with decreasing thickness of the FM film. Rezende et al. [194] argue that this is due to two-magnon scattering in which the uniform mode (the FMR) is scattered into a $\boldsymbol{k} \neq 0$ mode by the fluctuations in the exchange at the FM/AFM interface. This is analogous to the origin of the FMR linewidth in a YIG sphere due to magnon scattering induced by the dipole fields produced by surface roughness [163] (see, also Problem 8.2).

9.1.2 Biquadratic Exchange

Neutron scattering studies of superlattices composed of antiferromagnetic CoO and ferrimagnetic Fe_3O_4 show that when the Fe_3O_4 is magnetized the CoO spins align in domains with spins perpendicular to the net Fe_3O_4 moment. This is attributed to an exchange energy proportional to $\cos^2\theta$, where θ is the angle between the Fe_3O_4 moment and the AFM sublattice.

To understand the origin of this phenomenon let us consider a FM/AFM interface as shown in Fig. 9.3. Due to the surface roughness both compensated and uncompensated interfaces will have a surface energy consisting of exchange coupling between the FM spins, $\tilde{f}$, and the two AFM sublattice spins $\boldsymbol{a}$ and $\boldsymbol{b}$, i.e., $\boldsymbol{a} \cdot \boldsymbol{f}$. Let J_a and J_b be the exchange between a FM spin and an AFM spin at sublattice a or b, respectively. The interface energy then becomes

$$\mathcal{H}_{\text{int}} = -J_a \boldsymbol{a} \cdot \boldsymbol{f} - J_b \boldsymbol{b} \cdot \boldsymbol{f}. \tag{9.10}$$

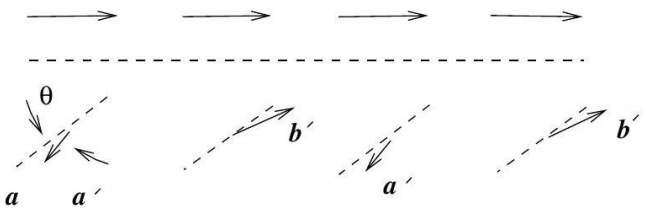


Fig. 9.4. Interface between a FM and AFM in which the AFM has fluctuated into its spin-flop state

It is convenient to introduce the sum and difference vectors,

$$\ell = a + b, \quad t = a - b.$$

The interface energy then becomes

$$\mathcal{H}_{\text{int}} = -\left(\frac{J_a + J_b}{2}\right) \boldsymbol{\ell} \cdot \boldsymbol{f} - \left(\frac{J_a - J_b}{2}\right) \boldsymbol{t} \cdot \boldsymbol{f}. \tag{9.11}$$

If the interface is uncompensated; for example if $\boldsymbol{b} = -\boldsymbol{a}$ (i.e., $\boldsymbol{\ell} = 0$) but $J_a \neq J_b$, then

$$\mathcal{H}_{\text{int}} = -\left(J_a - J_b\right) \boldsymbol{a} \cdot \boldsymbol{f}. \tag{9.12}$$

If the interface is compensated, i.e., $\boldsymbol{b} = -\boldsymbol{a}$ and $J_a = J_b$, then $\mathcal{H}_{\text{int}} = 0$. However, the AFM may fluctuate into the spin-flop state (Fig. 9.4).

This "twist" penetrates into the AFM and has an energy associated with it equal to the wall energy, σ_w. Now $\boldsymbol{b}' \neq -\boldsymbol{a}'$ and the change in energy according to second order perturbation theory is

$$\frac{\left(\frac{J_a+J_b}{2}\right)^2 \left(\ell' \cdot \boldsymbol{f}\right)^2}{\sigma_w} = \frac{\left(J_a + J_b\right)^2 \theta^2}{\sigma_w} (\boldsymbol{a} \cdot \boldsymbol{f})^2, \tag{9.13}$$

where θ is the canting angle. Thus, if the AFM surface is compensated, one of the "sublattices will be frustrated", i.e., points in a direction opposite to that preferred by the neighboring FM. This leads to the biquadratic exchange favoring a perpendicular alignment.

10.10.2 9.2 Trilayers

Very interesting behavior is found in magnetic systems consisting of two magnetic films separated by a nonmagnetic layer. In anticipation of fabricating magnetic multilayers theoretical analyses predicted a variety of magnetic structures. Majkrzek et al. [195] and Grunberg et al. [196] independently demonstrated exchange coupling between magnetic films separated by nonmagnetic spacers. Majkrzek et al. studied Gd_nY_m superlattices using neutron diffraction while Grunberg et al. studied light scattering from spinwaves in

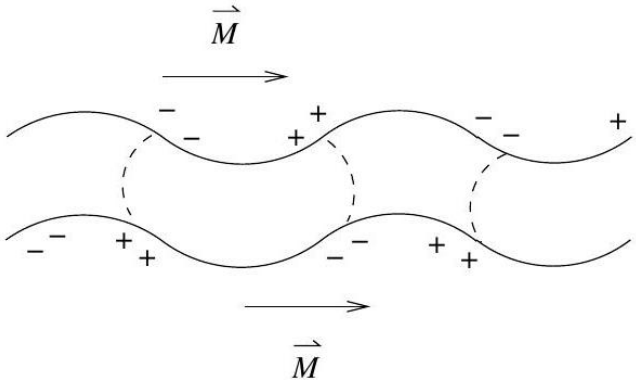


Fig. 9.5. Magnetostatic coupling across a nonmagnetic spacer layer associated with magnetic poles generated by surface roughness. (This is referred to as "orange peel" coupling after the texture of the surface of an orange)

$Fe-Cr-Fe$ and $Fe-Au-Fe$ trilayers. Both groups observed antiferromagnetic coupling at certain spacer thicknesses.

There are several mechanisms for coupling to occur across the spacer layer. If the spacer is too thin the most obvious coupling is through "pin holes", or regions in which the magnetic material connect through the spacer. This can be avoided by growing high quality films.

Néel also pointed out that surface roughness will develop magnetic poles that lead to magnetostatic coupling across the spacer (see Fig. 9.5). Since the roughness tends to propogate through the structure, the topology on both sides of the spacer will generally have the same "phase". This leads to ferromagnetic coupling. The observed antiferromagnetic coupling is believed to be intrinsically associated with the metallic spacer. In the next two sections we shall describe two models that give antiferromagnetic coupling.

9.2.1 The RKKY Ineraction

As we saw in Sect.2.2, there is a contact hyperfine interaction between s-state electrons and nuclear moments. Fröhlich and Nabarro were the first to suggest that this interaction could lead to a polarization of the nuclear moments [197]. The actual form of this interaction, however, was obtained by Ruderman and Kittel [198]. In analogy with this nuclear coupling, the exchange interaction between conduction electrons and localized electrons can also lead to indirect coupling between localized electronic moments. Zener [199] proposed that this was the origin of ferromagnetism in transition metals. Kasuya investigated this interaction in more detail, particularly with respect to its effect on spin waves and electrical resistivity [200]. Yosida also employed this interaction to explain the magnetic properties of $Cu-Mn$ alloys [201]. As a result of these developments, the indirect coupling of magnetic moments by condution electrons is referred to as the Ruderman-Kittel-Kasuya-Yosida (RKKY) interaction.

The form of the RKKY interaction is easily obtained within the framework of our generalized susceptibility. Let us assume that the interaction between
a localized spin S_α located at $\boldsymbol{r}=0$ and the conduction spins $\boldsymbol{s}_i$ has the form

$$-J\sum_i \boldsymbol{S}_\alpha\cdot\boldsymbol{s}_i(\boldsymbol{r})$$

Each conduction spin therefore experiences an effective field given by

$$\boldsymbol{H}_{\text{eff}}(\boldsymbol{r})=-\frac{J}{g\mu_B}\boldsymbol{S}_\alpha\delta(\boldsymbol{r}) \tag{9.14}$$

The response of an electron gas to such a field is determined by its susceptibility $\chi(\boldsymbol{q})$. Since the Fourier transform of this field is

$$\boldsymbol{H}_{\text{eff}}(\boldsymbol{q}) = -\frac{J}{g\mu_B}\boldsymbol{S}_\alpha \tag{9.15}$$

the spin density at $\boldsymbol{r}$ is

$$\boldsymbol{s}(\boldsymbol{r}) = \frac{J}{g^2\mu_B^2 V}\sum_{\boldsymbol{q}} \chi(\boldsymbol{q})e^{i\boldsymbol{q}\cdot r}\boldsymbol{S}_\alpha \tag{9.16}$$

For a free-electron gas $\chi(\boldsymbol{q})$ is given by (3.97). The sum over $\boldsymbol{q}$ is evaluated by converting it into an integral. This integral is easily evaluated by using the integral representation [H. Levine, private communication].

$$\ln\left|\frac{2k_F+q}{2k_F-q}\right| = 2\int_0^\infty \frac{dx\sin(2k_Fx)\sin(qx)}{x} \tag{9.17}$$

The result is

$$\begin{aligned}\frac{1}{V}\sum_{\boldsymbol{q}}\chi(\boldsymbol{q})e^{i\boldsymbol{q}\cdot r} &= \frac{3g^2\mu_B^2(N/V)}{3\varepsilon_F}\frac{1}{2\pi^2 r}\int dqqF\left(\frac{q}{2k_F}\right)\sin qr \\ &= \frac{3g^2\mu_B^2(N/V)}{8\varepsilon_F}\frac{k_F^3}{16\pi}\left[\frac{\sin 2k_Fr - 2k_Fr\cos 2k_Fr}{(k_Fr)^4}\right]\end{aligned} \tag{9.18}$$

The part of this expression within the brackets is plotted in Fig. 9.6.
Thus we see that when a localized moment is introduced into a metal, the conduction spins develop an oscillating polarization in the vicinity of this moment. These spin-density oscillations have the same form as the Friedel charge-density oscillations that result when an electron gas screens out a charge impurity.

The most direct evidence for the oscillatory nature of this interaction comes from NMR studies. In **CuFe**, for example, the nuclear resonance spectrum consists of the strong Knight-shifted Cu^{63} line found in pure Cu plus various satellites corresponding to Cu nuclei which lie within the RKKY oscillations around an Fe impurity. These Cu nuclei feel a contact hyperfine field

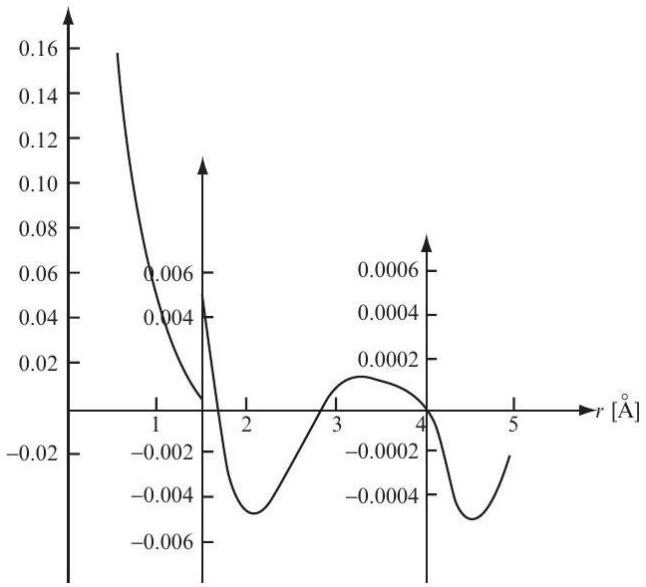


Fig. 9.6. Plot of the conduction-electron spin density produced by an impurity spin in a metallic host

$$\delta H(\boldsymbol{r}, T) = -\frac{8}{3}\pi g \mu_B \sigma(\boldsymbol{r}, T),$$

where

$$\sigma(\boldsymbol{r}, T) = \chi_{\text{imp}}\,(T) H f(\boldsymbol{r}).$$

Here $\chi_{\text{imp}}\,(T)$ is the susceptibility of a single Fe impurity and $f(\boldsymbol{r})$ characterizes the RKKY spatial dependence. An external field polarizes the impurity which in turn increases $\sigma(\boldsymbol{r}, T)$. Thus the satellites should move away from the main line with increasing field with a slope that depends upon $f(\boldsymbol{r})$. Figure 9.7 shows the field dependence of five such satellites [202]. The appearance of both positive and negative slopes results from the oscillatory nature of $f(\boldsymbol{r})$.

If there is another localized spin $\boldsymbol{S}_\beta$ at $\boldsymbol{r}$, it interacts with this induced spin density, leading to an effective interaction between the localized spins of the form

$$\mathcal{H}_{RKKY} = -\frac{J^2}{g^2 \mu_B^2 V} \sum_{\boldsymbol{q}} \chi(\boldsymbol{q}) e^{i\boldsymbol{q}\cdot\boldsymbol{r}} \boldsymbol{S}_\alpha \cdot \boldsymbol{S}_\beta. \tag{9.19}$$

This interaction manifests itself in many ways. Ruderman and Kittel [198] showed that it leads to a broadening of the nuclear magnetic resonance absorption. This interaction is also the origin of the exchange coupling between the localized moments in the rare-earth metals. As we saw in Chap. 4, its oscillatory nature leads to helimagnetism in these materials.

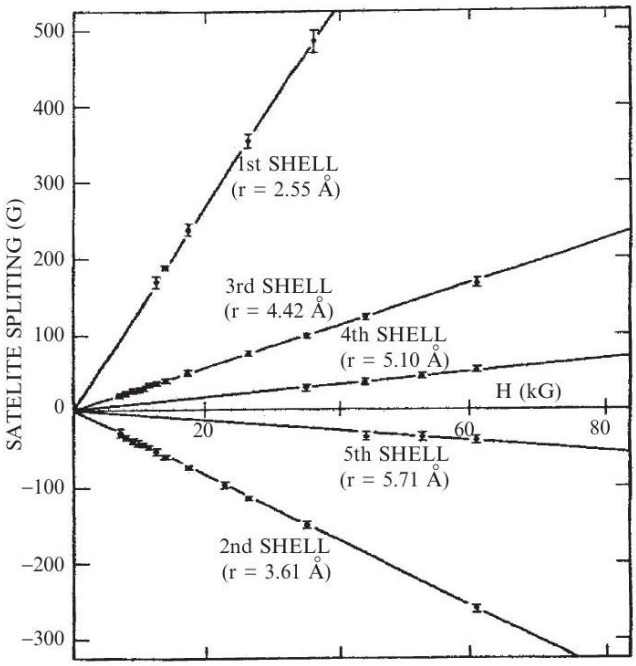


Fig. 9.7. Magnetic field dependance of satellite separations from the main $\mathrm{Cu ^{63}$ resonance at 300 K in **CuFe** [202]}

9.2.2 Quantum Well Model

When the moments being coupled are itinerant, such as in an Fe/Cu/Fe trilayer, it is appropriate to describe the whole trilayer in a band picture. An electron moving in a ferromagnet will experience an exchange potential. We represent this by a difference in the potentials for spin up and spin down. Thus, the trilayer potential for a spin up when the two ferromagnets are antiparallel is shown in Fig. 9.8a. The objective is to calculate the energy levels associated with each of these potentials and then determine the total energy when these

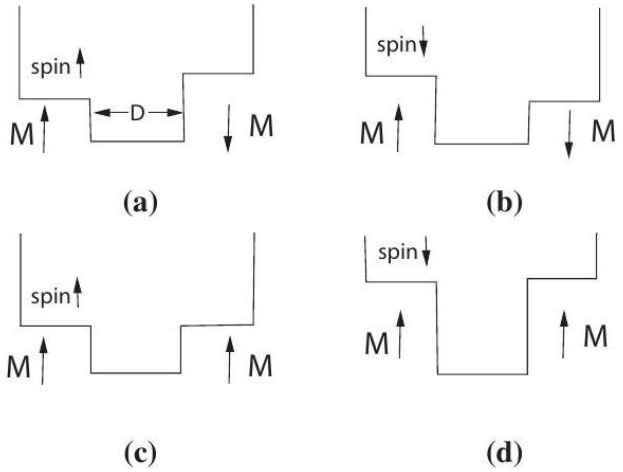


Fig. 9.8. Squarewell potentials seen by an electron moving in a trilayer consisting of two ferromagnets separated by a nonmagnetic, but conducting, spacer

levels are filled up to the Fermi level [205]. Since we are dealing with a large number of particles it is convenient to work with the grand thermodynamic potential, Ω, we introduced in our discussion of the deHaas-Van Alphen effect, (3.75):

$$\Omega = k_B T \sum_i \ln\left(1 + \exp\left[-\beta\left(E_i - E_F\right)\right]\right)$$
$$= -k_B T \int_{-\infty}^{\infty} \ln\left[1 + \exp\left(\frac{\varepsilon_F - \varepsilon}{k_B T}\right)\right] n(\varepsilon) d\varepsilon \tag{9.20}$$

where $n(\varepsilon)$ is the density of states. Integrating by parts gives

$$\Omega = -\int_{-\infty}^{\infty} N(\varepsilon) f(\varepsilon) d\varepsilon$$

where $N(\varepsilon)$ is the integrated density of states,

$$N(\varepsilon) = \int_{-\infty}^{\varepsilon} n\left(\varepsilon'\right) d\varepsilon'$$

and $f(\varepsilon)$ is the Fermi function. At $T = 0$

$$\Omega \to -\int_{-\infty}^{\varepsilon_F} N(\varepsilon) d\varepsilon$$

Thus, our problem reduces to calculating the density of states associated with the potentials in Fig. 9.8. Since we are interested in how the density of states changes with the thickness of the spacer layer we focus our attention on this layer. When an electron in the spacer layer moving to the right with momentum $k^+$is incident on the ferromagnetic layer its wave function is reflected and transmitted. Let us denote the reflection coefficient for a spin up striking the B ferromagnet when its moment is also "up" as $R_B^{\uparrow}$. The electron then moves back to the A ferromagnet where it is reflected with a coefficient $R_A^{\uparrow}$ if the A moment is also "up". Bruno [203] has shown that this confinement of up-spin electrons in the spacer bound by ferromagnetically aligned moments produces a change in the integrated density of states equal to

$$\Delta N(\varepsilon) = -\frac{2}{\pi} \operatorname{Im} \ln\left(1 - R_A^{\uparrow} R_B^{\uparrow} e^{iqD}\right), \tag{9.21}$$

where $q = k^+ - k^-$. The change in the thermodynamic potential Ω, associated with this change in the integrated density of states gives the change in energy

$$\Delta E_F^{\uparrow} = \frac{2}{\pi} \operatorname{Im} \int_{-\infty}^{\infty} \ln\left(1 - R_A^{\uparrow} R_B^{\uparrow} e^{iqD}\right) d\varepsilon \tag{9.22}$$

Considering the other three configurations and generalizing to threedimensions we obtain

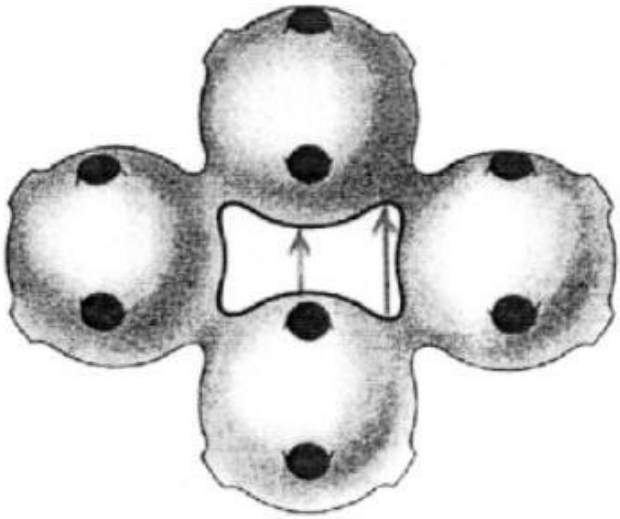


Fig. 9.9. Fermi surface of gold showing the two spanning vectors that govern the coupling between magnetic films placed on either side of the gold

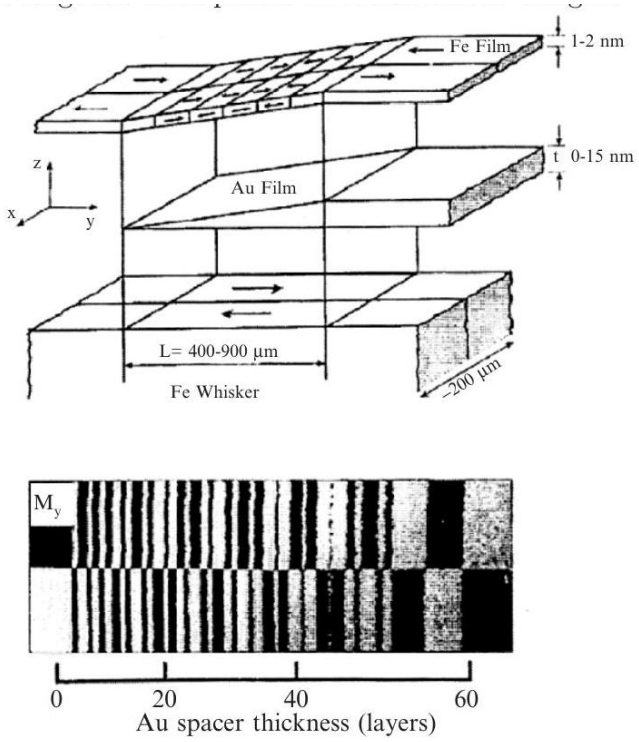


Fig. 9.10. Magnetization in the plane of an Fe film separated from an Fe whisker by a gold wedge showing that the coupling varies with the thickness of the wedge [204]

$$\Delta E_F - \Delta E_{AF} = \frac{1}{4\pi^3}\,\mathrm{Im}\int d^2k_{\parallel}\int_{-\infty}^{+\infty} f(\varepsilon) \times \ln\left[\frac{\left(1-R_A^{\uparrow}R_B^{\uparrow}e^{iq_{\perp}D}\right)\left(1-R_A^{\downarrow}R_B^{\downarrow}e^{iq_{\perp}D}\right)}{\left(1-R_A^{\uparrow}R_B^{\downarrow}e^{iq_{\perp}D}\right)\left(1-R_A^{\downarrow}R_B^{\uparrow}e^{iq_{\perp}D}\right)}\right] d\varepsilon \tag{9.23}$$

In the limit of weak confinement this may be written

$$E_F - E_{AF} = -\frac{1}{\pi^3}\,\mathrm{Im}\int d^2k_\parallel \int_{-\infty}^{\infty} f(\varepsilon)\Delta R_A \Delta R_B e^{iq_\perp D} d\varepsilon \tag{9.24}$$

Table 9.1. Exchange Oscillation Periods for $\mathrm{Fe / \mathrm{Au} / \mathrm{Fe}$ (100)}

SEMPA	2.48 ± 0.05	8.6 ± 0.3
dHvA	2.51	8.60

where $\Delta R = R^\uparrow - R^\downarrow$. The factor $\exp(iq_\perp D)$ describes the propagation through this spacer. Since $q_\perp = 2k_F$ this predicts that the coupling oscillates as a function of spacer thickness with a period given by $2k_F$. When $2k_F$ connects regions of the Fermi surfaces that are relatively flat the integral over $k_\parallel$will weight these contributions more heavily. We refer to these particular vectors as the spanning vectors of the Fermi surface. Figure 9.9 shows the Fermi surfaces of gold in the extended zone scheme. For a spacer oriented in the (100) direction there are two spanning vectors. This would suggest that the coupling between two magnetic films separated by gold with a (100) orientation should show a variation with thickness that is the superposition of a long and a short period. Such behavior has been observed in a very nice experiment on Fe/Au/Fe by Unquris, et al. [204]. As shown in Fig. 9.10, they deposit a thin iron layer on top of a gold wedge-shaped spacer which is in turn deposited on top of an iron whisker with two magnetic domains. The direction of the magnetization in the top layer is measured by Scanning Electron Microscopy with Polarization Analysis (SEMPA). The periods deduced from the SEMPA measurement are compared with those obtained from deHaas-van Alphen measurements in Table 9.1 [204].

9.2.3 Giant Magnetoresistance (GMR)

When current flows through a magnetic film, as shown in Fig. 9.11 the resistivity depends upon the angle the current makes with the magnetization according to

$$\rho = \rho_0 + \Delta\rho \cos^2\theta. \tag{9.25}$$

This is called the anisotropic magnetoresistance (AMR) and is due to the scattering of the conduction electrons from the d-moments of the magnetization.

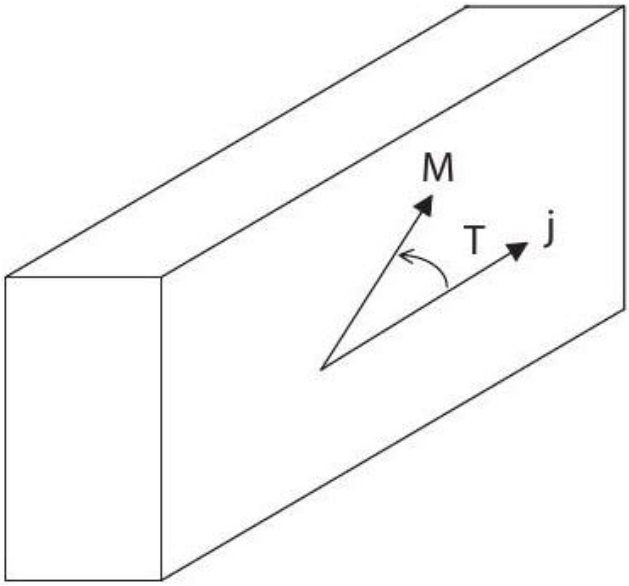


Fig. 9.11. Geometry for the AMR effect

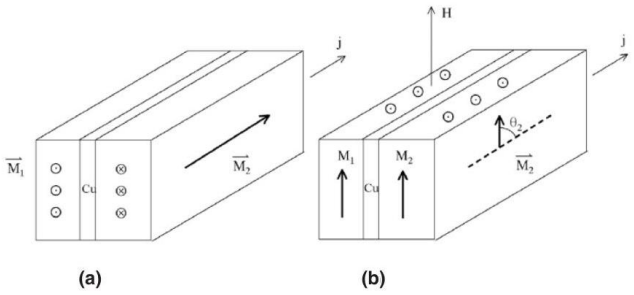


Fig. 9.12. Geometry of the Current-In-Plane (CIP) GMR effect. (a) high resistance; (b) low resistance

The largest effect has been found in permalloy where $\Delta\rho/\rho_0 = 2\%$. With the development of thin film deposition techniques, it was natural to measure the magnetoresistance of magnetic multilayers. The surprising result was that the magnetoresistance of an Fe/Cr/Fe trilayer was 200% (at 4.2 K) [206]. However, in the case of the trilayer, the resistance depends upon the relative orientation of the magnetizations in the magnetic layers.

If the metallic inner layer provides an antiferromagnetic coupling, the magnetizations in zero field will be antiparallel as shown in Fig. 9.12a. In the presence of a strong enough field, the moments become parallel as shown in Fig. 9.12b. The resistance is always found to be larger in the antiparallel configuration. This effect is also independent of the relative orientation of the magnetizations and the bias current. However, the magnitude of the effect is larger when the current flows perpendicular to the plane of the magnetization (CPP) as opposed to the current being in the plane (CIP) - as is the case in Fig. 9.12. Thus, the resistivity of this "giant magnetoresistance" (GMR) effect may be expressed as

$$\rho = \rho_0 + \Delta\rho \cos(\theta_1 - \theta_2) . \tag{9.26}$$

For a summary of experimental aspects of GMR see [207]. The microscopic explanation of the GMR effect may be understood by first considering the scattering of electrons in a ferromagnet. As we saw in Sect.5.3, the spin-up and spin-down d-bands are exchange split, the

s-bands less so as illustrated in Fig. 9.13. The current is largely carried by the s-electrons. However, when they scatter, from impurities or lattice irregularities, for example, the probability of scattering into a *d*-state is large because the density of *d*-states is larger.

Since the probability of spin flip scattering is smaller, we shall not include this process. Now, let us consider how this scattering affects the transport in a Co/Cr/Co trilayer. Consider the down-spin (1) (relative to the direction of the magnetization) in Fig. 9.14a. When this electron reaches the Cr interface it scatters back into the Co with a probability that is proportional to the Co *d*-electron down-spin density, $N_{d\downarrow}$. The up-spin (2) successfully transverses the

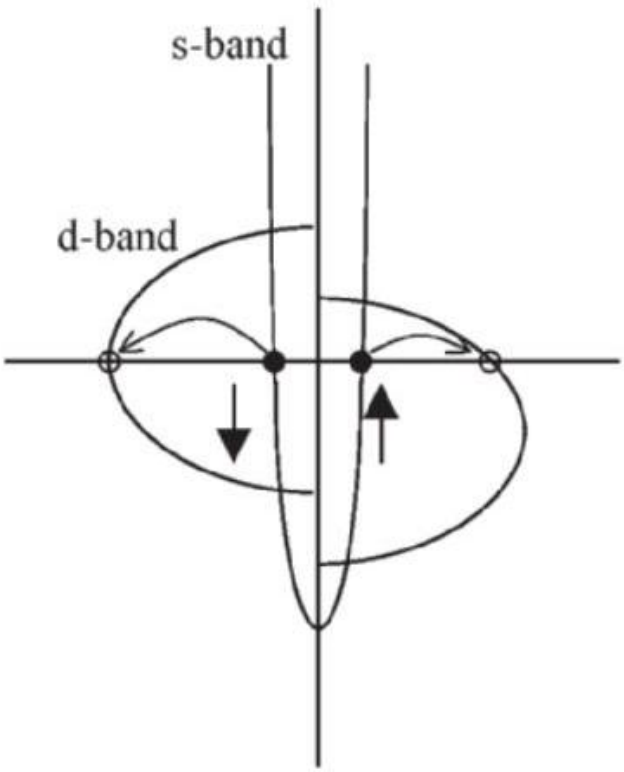


Fig. 9.13. Simplified band structure of a ferromagnet showing *s*-electrons being scattered into *d*-states

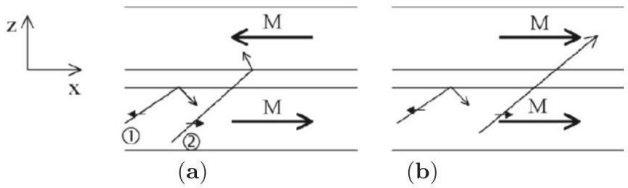


Fig. 9.14. Most likely scattering processes in a trilayer whose magnetic layers have a band structure like that shown in Fig. 9.13

Co-Cr interface because the majority spin *d*-band density of states for the bands shown in Fig. 9.13 is smaller. However, when this electron reaches the interface with the top magnetic layer there are two possibilities depending upon whether the magnetization of this layer is parallel or antiparallel to the bottom layer. If it is parallel (Fig. 9.14b), the probability is again small and the electron continues on its way. If the magnetization is antiparallel (Fig. 9.14a), then the up-spin electron finds itself in an environmment in which it is effectively a "down-spin"

and therefore, scatters more strongly. This argument suggests that the average mean free path in the case of parallel magnetizations is larger than that of antiparallel magnetizations, giving the parallel configuration a smaller resistivity.

The GMR serves as the basis for very sensitive detection of magnetic fields. In this application one of the magnetic layers is "pinned", generally by placing it in contact with an antiferromagnet and utilizing the exchange bias. [1] The

other layer is designed to have a very small coercivity, so it is free to be oriented by the field one is sensing. This configuration is called a spin valve.

In order to maximize the GMR one needs a theory that incorporates the physics necessary to predict the GMR behavior.

Kubo Formula. In keeping with the linear response approach adopted in this text one can apply the same formalism to calculate the electrical conductivity $\sigma_{\alpha\beta}$, that we used to calculate the susceptibility. We apply a perturbing electric field and calculate the current response. Since the applied field produces currents which, in turn, produce electric fields, the conductivity is defined by the current response to the total electric field, $\mathcal{E}(\boldsymbol{r},t)$,

$$J_\alpha(\boldsymbol{r},t)=\int d^3r'\int dt'\sigma_{\alpha\beta}\left(\boldsymbol{r}-\boldsymbol{r}',t-t'\right)\mathcal{E}_\beta\left(\boldsymbol{r}',t'\right) \tag{9.27}$$

The current density at $\boldsymbol{r}$ is given by averaging the velocity of the particles over volume Ω,

$$J_\alpha(\boldsymbol{r},t)=e\left\langle\sum_i v_{i\alpha}\delta\left(\boldsymbol{r}-\boldsymbol{r}_i\right)\right\rangle \tag{9.28}$$

The velocity is given by

$$\boldsymbol{v}_i=\frac{1}{m}\left[\boldsymbol{p}_i-\frac{e}{c}\boldsymbol{a}\left(\boldsymbol{r}_i\right)\right] \tag{9.29}$$

Therefore

$$J_\alpha(\boldsymbol{r},t)=\frac{e}{m\Omega}\sum_i\left\langle p_{i\alpha}\right\rangle-\frac{e^2}{mc}\sum_i\left\langle A_\alpha\left(\boldsymbol{r}_i\right)\right\rangle \tag{9.30}$$

Since the momentum is related to the current by $\boldsymbol{i}_i=\frac{e}{m}\boldsymbol{p}_i$,

$$J_\alpha(\boldsymbol{r},t)=\frac{1}{\Omega}\sum_i\left\langle i_{i\alpha}\right\rangle-\frac{e^2}{mc}\sum_i\left\langle A_\alpha\left(\boldsymbol{r}_i\right)\right\rangle \tag{9.31}$$

The vector potential is related to the electric field by $A_\alpha\left(\boldsymbol{r}_i\right)=-i\frac{c}{\omega}\mathcal{E}_\alpha\left(\boldsymbol{r}_i\right)$. The number of particles in the sum over i is $n\Omega$ where n is the particle density. This sum gives the macroscopic average,

$$J_\alpha(\boldsymbol{r},t)=\left\langle j_\alpha(\boldsymbol{r},t)\right\rangle+i\frac{ne^2}{m\omega}\mathcal{E}_\alpha(\boldsymbol{r},t) \tag{9.32}$$

To obtain $\sigma_{\alpha\beta}$ we must calculate $\left\langle j_\alpha(\boldsymbol{r},t)\right\rangle$ to lowest order in $\mathcal{E}_\alpha(\boldsymbol{r},t)$. The result is [208].

$$\sigma_{\alpha\beta}(\boldsymbol{q},\omega)=\frac{1}{\omega\Omega}\int_0^\infty dte^{i\omega t}\langle\psi|j_\alpha^\dagger(\boldsymbol{q},t),j_\beta(\boldsymbol{q},0)|\psi\rangle+i\frac{ne^2}{m\omega}\delta_{\alpha\beta} \tag{9.33}$$

[01] In real devices the design is slightly more complex. The pinned layer has magnetic charge at its edges which produces a demagnetization field which can affect this free layer. To minimize this effet a third magnetic layer is placed next to the pinned layer with a nonmagnetic spacer having a thickness that gives an antiferromagnetic coupling. This third layer is, in turn, pinned by the antiferromagnet. The two antiferromagnetically coupled ferromagnetic layers are called a "synthetic antiferromagnet" (SAF).

Thus, the conductivity is given by the average value of the current-current correlation function where the wavefunction $|\psi\rangle$ is the ground state of the many body system. Equation (9.33) is known as the Kubo formula. The problem with this approach is that one does not always know the ground state $|\psi\rangle$. Butler and his colleagues, for example, have employed the local spin density approximation to the density functional theory we discussed on page 177 to calculate the transport properties of multilayers [209].

Boltzmann Equation. An alternative approach to describing the GMR effect which relates more directly to the physical picture we gave at the beginning of this section is the semiclassical approach based on the Boltzmann equation. In this approach the electron is described by a distribution function, $f^\sigma(\boldsymbol{r}, \boldsymbol{k}, t)$, which gives the probability of finding an electron with spin σ at $\boldsymbol{r}$ with a momentum $\hbar\boldsymbol{k}$. Conceptually, this is the basis for the Fermi liquid approach we introduced in Chap. 7. This probability will change over time due to forces acting on the electron, and also simply due to the fact that the electron is moving through space. The total time rate of change is given by

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \dot{\boldsymbol{r}} \cdot \nabla_{\boldsymbol{r}} f + \dot{\boldsymbol{k}} \cdot \nabla_{\boldsymbol{k}} f + \left.\frac{\partial f}{\partial t}\right|_{\text{scat}}, \tag{9.34}$$

where the last term is the change due to scattering. In equilibrium the total time derivative is zero. Therefore

$$\frac{\partial f}{\partial t} = -\dot{\boldsymbol{r}} \cdot \nabla_{\boldsymbol{r}} f - \dot{\boldsymbol{k}} \cdot \nabla_{\boldsymbol{k}} f - \left.\frac{\partial f}{\partial t}\right|_{\text{scat}}. \tag{9.35}$$

$\dot{\boldsymbol{k}}$ is given by the forces,

$$\hbar\dot{\boldsymbol{k}} = -|e|\mathcal{E} - |e|\boldsymbol{v} \times \boldsymbol{B}. \tag{9.36}$$

The term involving the magnetic field is generally small and, therefore, neglected. Let us also restrict our consideration to the dc conductivity so that f is not a function of time.

Let $\mathcal{P}_{\boldsymbol{k},\boldsymbol{k}'}$ be the probability that an electron scatters from state $\boldsymbol{k}$ to $\boldsymbol{k}'$. We shall neglect spin-flip scattering since its probability is small. The scattering may, however, be different for the two spin directions giving rise to spin-dependent scattering. We shall suppress the spin index, keeping in mind that there are two spin channels. Then the scattering term may be written.

$$\begin{aligned}\left.\frac{\partial f}{\partial t}\right|_{\text{scat}} = &\sum_{\boldsymbol{k}'} P_{\boldsymbol{k}\boldsymbol{k}'}[1 - f(\boldsymbol{r}, \boldsymbol{k})] f(\boldsymbol{r}, \boldsymbol{k}') \\ &- P_{\boldsymbol{k}'\boldsymbol{k}}[1 - f(\boldsymbol{r}, \boldsymbol{k}')] f(\boldsymbol{r}, \boldsymbol{k})\end{aligned} \tag{9.37}$$

Microscopic reversibility requires $P_{\boldsymbol{k}'\boldsymbol{k}} = P_{\boldsymbol{k}\boldsymbol{k}'}$. Therefore

$$\left.\frac{\partial f}{\partial t}\right|_{\text{scat}} = \sum_{\boldsymbol{k}'} P_{\boldsymbol{k}\boldsymbol{k}'}[f(\boldsymbol{r}, \boldsymbol{k}') - f(\boldsymbol{r}, \boldsymbol{k})]. \tag{9.38}$$

In equilibrium, in the absence of an applied field, the distribution function is the Fermi-Dirac function

$$f_0(\boldsymbol{k}) = \frac{1}{1 + e^{[E(\boldsymbol{k}) - E_F]/k_B T}}$$

Again, in the spirit of linear response theory, we assume $f(\boldsymbol{r}, \boldsymbol{k})$ may be written as $f_0(\boldsymbol{k})$ plus a small deviation $g(\boldsymbol{r}, \boldsymbol{k})$, i.e.,

$$f(\boldsymbol{r},\boldsymbol{k}) = f_0(\boldsymbol{k}) + g(\boldsymbol{r},\boldsymbol{k}) \tag{9.39}$$

The equation for $g(\boldsymbol{r},\boldsymbol{k})$ then becomes

$$\boldsymbol{v}(\boldsymbol{k})\cdot\nabla_r g(\boldsymbol{r},\boldsymbol{k}) - |e|\mathcal{E}\cdot\boldsymbol{v}(\boldsymbol{k})\frac{\partial f_0(\boldsymbol{k})}{\partial E(\boldsymbol{k})} = \sum_{\mathbf{k}'} P_{\boldsymbol{k}\boldsymbol{k}'}\left[g\left(\boldsymbol{r},\boldsymbol{k}'\right) - g(\boldsymbol{r},\boldsymbol{k})\right] \tag{9.40}$$

To simplify the scattering term we assume that $g\left(\boldsymbol{r},\boldsymbol{k}'\right)$ is proportional to $g(\boldsymbol{r},\boldsymbol{k})$ where the coefficient of proportionality depends upon $\boldsymbol{k}$ and $\boldsymbol{k}'$, i.e.,

$$g\left(\boldsymbol{r},\boldsymbol{k}'\right) = c\left(\boldsymbol{k},\boldsymbol{k}'\right) g(\boldsymbol{r},\boldsymbol{k})$$

The scattering term then reduces to $\frac{g(\boldsymbol{r},\boldsymbol{k})}{\tau(\boldsymbol{k})}$ where

$$\frac{1}{\tau(\boldsymbol{k})} = \sum_{\mathbf{k}'} P_{\boldsymbol{k}\boldsymbol{k}'}\left[c\left(\boldsymbol{k},\boldsymbol{k}'\right) - 1\right]$$

It turns out that $c\left(\boldsymbol{k},\boldsymbol{k}'\right) - 1$ varies as $1-\cos\theta$ where θ is the angle between $\boldsymbol{k}$ and $\boldsymbol{k}'$. That is, large angle scattering is weighted more heavily in determining $\tau(\boldsymbol{k})$. Equation (9.40) becomes

$$\boldsymbol{v}(\boldsymbol{k})\cdot\nabla_r g(\boldsymbol{r},\boldsymbol{k}) - |e|\mathcal{E}\cdot\boldsymbol{v}(\boldsymbol{k})\frac{\partial f_0(\boldsymbol{k})}{\partial E(\boldsymbol{k})} = \frac{g(\boldsymbol{r},\boldsymbol{k})}{\tau(\boldsymbol{k})} \tag{9.41}$$

This simplication of the scattering term is referred to as the relaxation time approximation.

Before considering the multilayer geometry it is of interest to consider a homogeneous system where $\nabla_{\boldsymbol{r}} g(\boldsymbol{r},\boldsymbol{k}) = 0$. Then the solution to (9.41) becomes

$$g(\boldsymbol{k}) = |e|\tau(\boldsymbol{k})\frac{\partial f_0(\boldsymbol{k})}{\partial E(\boldsymbol{k})}\cdot\boldsymbol{v}(\boldsymbol{k})\cdot\mathcal{E}$$

The current density is

$$j(\boldsymbol{r}) = \frac{|e|}{\Omega}\sum_{\boldsymbol{k}}\boldsymbol{v}(\boldsymbol{k})g(\boldsymbol{k}) \tag{9.42}$$

Therefore the conductivity is

$$\sigma_{\alpha\beta} = \frac{|e|^2}{\Omega}\sum_{\boldsymbol{k}} v_\alpha(\boldsymbol{k})v_\beta(\boldsymbol{k})\tau(\boldsymbol{k})\frac{\partial f_0(\boldsymbol{k})}{\partial E(\boldsymbol{k})}. \tag{9.43}$$

While (9.43) can be derived from the Kubo formula (9.33), the derivation is not trivial (see [208], Sect. 8.1.2).

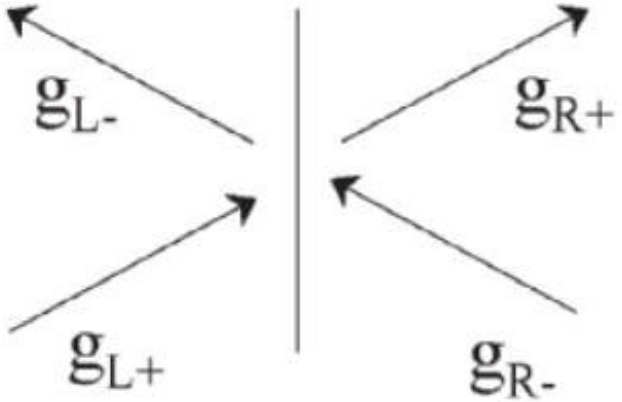
Let us now consider a trilayer structure consisting of two magnetic layers (e.g., Fe or Co.) separated by a thin conductor (e.g., Cu or Cr) [210]. Assume the electric field is applied in the plane of the trilayer (x-direction) and that the normal to the sandwich defines the z-direction. Then (9.41) reduces to

$$\frac{\partial g(z,\boldsymbol{k})}{\partial z} - \frac{g(z,\boldsymbol{k})}{v_z(\boldsymbol{k})\tau(\boldsymbol{k})} - \frac{|e|\mathcal{E}v_x(\boldsymbol{k})}{v_z(\boldsymbol{k})}\frac{\partial f_0(\boldsymbol{k})}{\partial E(\boldsymbol{k})} = 0. \tag{9.44}$$

We shall solve this equation piecewise for the different layers. Furthermore, the solution for $v_z(\boldsymbol{k}) > 0$ is different than that for $v_z(\boldsymbol{k}) < 0$. Denoting those solutions by $\pm$, we obtain for a given layer

$$g_{\pm}(z,\boldsymbol{k}) = e\tau(\boldsymbol{k})\mathcal{E}v_x(\boldsymbol{k})\frac{\partial f_0(\boldsymbol{k})}{\partial E(\boldsymbol{k})}\left[1 + A^{\pm}(\boldsymbol{k})e^{\mp z/\tau(\boldsymbol{k})|v_z(\boldsymbol{k})|}\right] \tag{9.45}$$

The coefficients $A^{\pm}(\boldsymbol{k})$ are obtained from boundary conditions. For example, consider the following boundary:



The electron moving away from the boundary on the left could come from the reflection of the electron incident on the boundary from the left or from the transmission of the electron incident on the boundary from the right, i.e.,

$$g_{L^-} = Rg_{L^+} + Tg_{R^-},$$

where R and T are the reflection and transmission coefficients, respectively.
The spin dependence arising from the spin-dependent scattering enters through the relative orientation, θ, of the magnetizations in the two layers. Camley and Barnaś dealt with this by introducing an artificial interface in the middle of the conducting spacer where the spin quantization axis is rotated by θ. Based on our discussion of the change of the spin wavefunction with quantization direction in Sect.2.2.10, the probabilities g^{σ} have the following relationships:

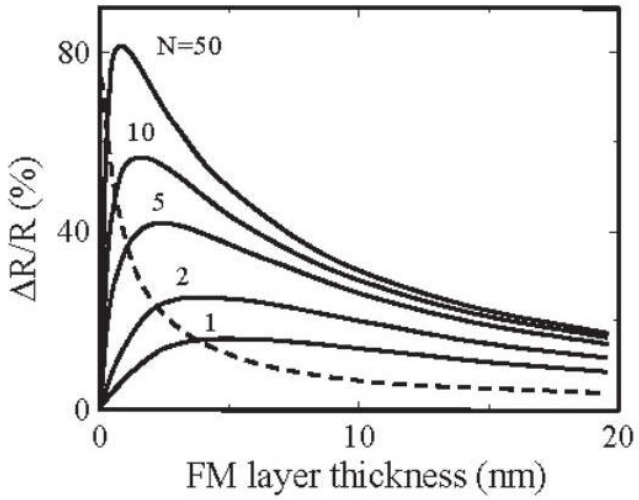


Fig. 9.15. Magnetoresistance versus ferromagnetic layer thickness in $(\mathrm{FM / \mathrm{NM})_{N} \mathrm{FM}$ multilayers for bulk (the solid lines) and interface (the dashed line) scattering as calculated using the semiclassical free-electron model. $\lambda_{NM} = 20$ nm and diffuse outer boundary scattering are assumed in the calculation. The other parameters are set as follows: $\lambda^{\uparrow}_{FM} = 12$ nm, $\lambda^{\downarrow}_{FM} = 0.6$ nm, $T^{\uparrow} = T^{\downarrow} = 0.1$ for interface spin-dependent scattering. After Dieny [214]}

$$\xrightarrow{g^{-\sigma}_{L+}} \Big| \xrightarrow[g^{\sigma}_{R+}=\cos^2\frac{\theta}{2}g^{\sigma}_{L+}+\sin^2\frac{\theta}{2}g^{-\sigma}_{L+}]{g^{\sigma}_{L+}}$$

and, similarly,

$$g^{\sigma}_{L-} = \cos^2\frac{\theta}{2}g^{\sigma}_{R-} + \sin^2\frac{\theta}{2}g^{-\sigma}_{R-}.$$

In principle, the reflection and transmission coefficients, R and T, may be calculated. However, they were regarded as parameters by Camley and Barnaś. In fact, they assumed that when the magnetic layers and the conducting spacer are transition metals, like Fe/Cr/Fe, the reflection coefficient is zero. Since this does not account for scattering at the Fe/Cr interface they introduced a diffusive scattering parameter, D, by setting $T = 1 - D$. In a subsequent paper, Johnson and Camley [211], replaced the concept of diffusive scattering at the internal interfaces by introducing "mixed" regions between the magnetic layers and the conducting layer. This gives six regions, each having their own coefficients $A^{\pm}_{\sigma}(\boldsymbol{k})$ in (9.45). There are 24 boundary conditions to determine these 24 coefficients which are solved numerically [212]. Once one knows $g^{\sigma}_{\pm}$ for all the regions the current density in each region can be calculated from (9.42). The total current is obtained by integrating $j_x(z)$ over z.

It turns out, [213], that the introduction of mixed regions is necessary to obtain an equivalence between the Kubo formalism and the Boltzmann formalism. Nevertheless, the use of a diffusive scattering parameter is convenient for obtaining qualitative understanding of the GMR effect. For example, suppose
we wish to understand how a multilayer, $(\mathrm{FM/NM})_N\mathrm{FM}$, behaves relative to a simple trilayer, FM/NM/FM. If we assume that the spin dependance arises from scattering within the ferromagnetic layers, then the scattering time, τ_σ, or, equivalently, the mean free path, λ_σ, is taken to be different for spin up and spin down, while the transmission coefficients, T_σ, are both taken to be 1 . On the other hand, if the spin dependence arises from scattering at the interface, then we set $\lambda_\uparrow = \lambda_\downarrow$. Figure 9.15 shows the magnetoresistance obtained from the Boltzmann equation as a function of the thickness of the ferromagnet for various multilayer configurations [214]. The solid lines assume the scattering within the ferromagnets is spin-dependent while the dashed line is for spin-dependent interface scattering. There are several observations. One is that the GMR for multilayers is larger than that of a trilayer (spin valve). This is due to the fact that in calculating the magnetoresistance in Fig. 9.15 it was assumed that the electrons are diffusely scattered at the outer boundaries. Therefore, for bulk spin-dependent scattering, the GMR will increase with multilayer thickness until this thickness becomes comparable to the mean free path, $\lambda_\uparrow$. Once the multilayer thickness exceeds $\lambda_\uparrow$ the role of the diffuse scattering decreases and further increase in the thickness simply provides a shunting path which reduces the GMR.

This argument suggests that the magnetoresistance is very sensitive to the outer boundary. As we mentioned above, one of the magnetic layers in a spin valve is pinned by placing it next to an antiferromagnet. The popular choice, FeMn, is generally believed to produce diffuse scattering. The GMR can be enhanced by using NiO as the pinning antiferromagnet. This enhancement is due to the fact that NiO is an insulator and reflects back the electrons, i.e., increases the specular scattering. A layer of Cu at the back of the magnetic layers also enhances the GMR. This is a "spin filtering" effect due to the fact that for Co/Cu, for example, the Fermi surface of the Co majority ($\uparrow$) spins is very similar to that of Cu whereas that of the Co minority ($\downarrow$) spins is very different. Therefore, the majority spins pass through the interface more easily.

9.2.4 Tunneling

Another trilayer structure that has enormous practical applications is that in which the intermediate layer, between the two metal electrodes, is a thin insulator. When a voltage is applied across this structure a small current is observed. This current occurs because the insulator represents a finite potential barrier of the order of the bandgap of the insulator as illustrated in Fig. 9.16. The bandgap of Al_2O_3, for example, is 6.6 eV . This means the wave function for an electron with an energy below the barrier height is an evanescent wave and an electron in the left electrode will therefore have a finite probability of tunneling into the right electrode.

There are a variety of approaches to calculating the tunneling conductance. One of these, introduced by Bardeen [215], is based on a tunneling Hamiltonian. If $c_{\boldsymbol{q}\sigma}$ destroys an electron with momentum $\mathbf{q}$ and spin σ on one

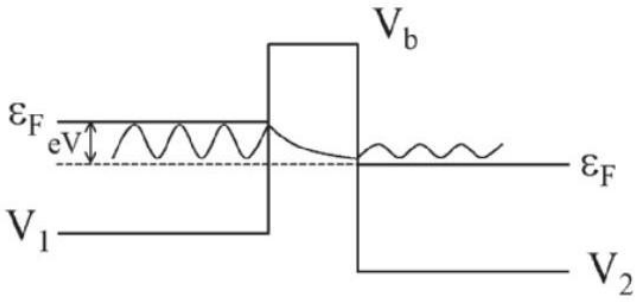


Fig. 9.16. Wavefunction associated with tunnel barrier

side of the junction and $c^{\dagger}_{\boldsymbol{k}\sigma}$ creates one with momentum $\boldsymbol{k}$ and spin σ on the other side the tunneling may be represented by a Hamiltonian of the form

$$\mathcal{H}_T = \sum_{\boldsymbol{k},\boldsymbol{q}} \left(T^{\sigma}_{\boldsymbol{k}\boldsymbol{q}} c^{\dagger}_{\boldsymbol{k}\sigma} c_{\boldsymbol{q}\sigma} + T^{\sigma}_{\boldsymbol{q}\boldsymbol{k}} c^{\dagger}_{\boldsymbol{q}\sigma} c_{\boldsymbol{k}\sigma} \right). \tag{9.46}$$

The total probability that an electron tunnels from one side, say 1 , to the other, 2 , is

$$W_{1\to 2} = \frac{2\pi}{\hbar} \sum_{l} \left| \left\langle l \right| \mathcal{H}_T \left| 0 \right\rangle \right|^2 \delta\left(E_l - E_0\right) \tag{9.47}$$

where $|0\rangle$ and $|l\rangle$ are products of the eigenstates of the Hamiltonians appropriate for each side. If one now writes

$$\delta\left(\varepsilon_{\boldsymbol{q}\sigma} + \mathrm{eV} - \varepsilon_{\boldsymbol{k}\sigma}\right) = \int_{-\infty}^{\infty} dE \delta\left(\varepsilon_{\boldsymbol{q}\sigma} - E\right) \delta\left(\varepsilon_{\boldsymbol{k}\sigma} - \mathrm{eV} - E\right) \tag{9.48}$$

and converts the sums over $\boldsymbol{q}$ and $\boldsymbol{k}$ into integrals over $\varepsilon_{\boldsymbol{q}}$ and $\varepsilon_{\boldsymbol{k}}$ one finds that the tunneling current associated with spin σ is

$$I_{\sigma} = \frac{2\pi e}{\hbar} \int_{-\infty}^{\infty} dE \langle \left| T^{\sigma}_{kq} \right|^2 \rangle N_{1\sigma}(E + eV) N_{2\sigma}(E) [f(E + eV) - f(E)] \tag{9.49}$$

where $\langle |T|^2 \rangle$ is the average of the tunneling probability. If we assume that the densities of states, $N_{1\sigma}$ and $N_{2\sigma}$, are not strong functions of energy they may be replaced by their values at the Fermi energy and taken outside the integral. The tunneling conductance $G_{\sigma} = dI_{\sigma}/dV$ then becomes

$$G_\sigma = \frac{2\pi e^2}{\hbar} \langle |T_{kq}^\sigma|^2 \rangle N_{1\sigma}(E_F) N_{2\sigma}(E_F). \tag{9.50}$$

If the two magnetic electrodes have their magnetizations parallel (see Fig. 9.17) then the tunneling takes place between the majority (M) spin bands and minority (m) spin bands. Therefore, the conductance is given by

$$G_{\uparrow\uparrow} = G_\uparrow + G_\downarrow \sim N_{M1}N_{M2} + N_{m1}N_{m2}.$$

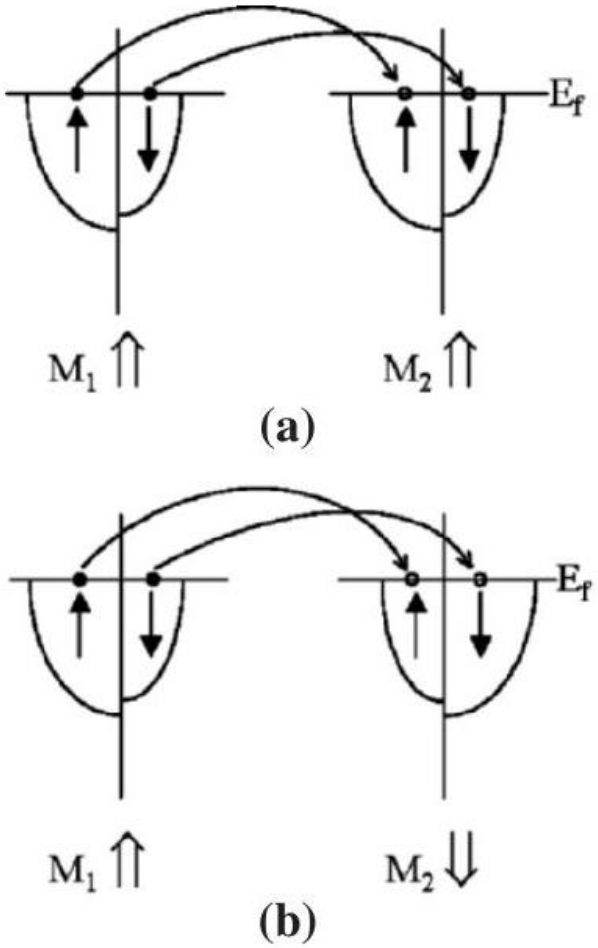


Fig. 9.17. Tunneling processes for parallel (a) and antiparallel (b) electrodes

If the magnetic electrodes are magnetized antiparallel, then

$$G_{\uparrow\downarrow} = G_\uparrow + G_\downarrow \sim N_{M1}N_{m2} + N_{m1}N_{M2}.$$

The tunneling magnetoresistance (TMR) is defined by

$$\text{TMR} = \frac{G_{\uparrow\uparrow} - G_{\uparrow\downarrow}}{G_{\uparrow\downarrow}}. \tag{9.51}$$

When we have exchange-split bands, the densities of states at the Fermi level for the different spin directions are different. These differences define the spin polarization, i.e.,

$$P_1 = \frac{N_{M1}(E_F) - N_{m1}(E_F)}{N_{M1}(E_F) + N_{m1}(E_F)}. \tag{9.52}$$

The TMR may then be expressed in terms of the spin polarizations of the two electrodes as

$$\mathrm{TMR} = \frac{2P_1P_2}{1 - P_1P_2} \tag{9.53}$$

This expression was first derived by Julliere [216] to explain his tunneling results in films of Fe/Ge/Co.

In order to test the simple relation (9.53) we need to know the polarizations of the electrode materials. There are three techniques used to measure the spin polarizations - tunneling in which one of the electrodes is a superconductor, Andreev reflection using a superconducting point contact, and spin-resolved photoemission.

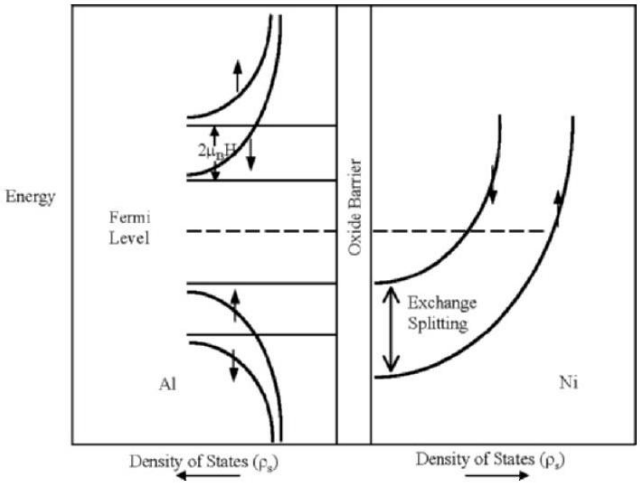


Fig. 9.18. Density of states for an $\mathrm{Al -\mathrm{Al}*{2}* *mathrm{O}*{3}-\mathrm{Ni}$ junction in which the Al is superconducting}

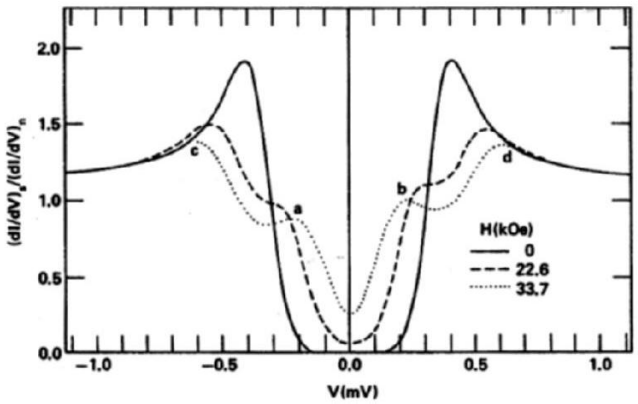


Fig. 9.19. Experimental results for the normalized conductance as a function of applied voltage for an $\mathrm{Al -\mathrm{Al}*{2}* *mathrm{O}*{3}-\mathrm{Ni}$ tunneling junction for different values of the applied field H (kOe). After Meservey and Tedrow [217]}

Spin-Dependent Tunneling. When one of the electrodes of a tunnel junction is a superconductor in its superconducting state the electrons are bound in ($\boldsymbol{k}\uparrow, -\boldsymbol{k}\downarrow$) pairs. When the pairs are excited they form quasiparticles with energies

$$E_{\boldsymbol{k}\sigma} = \left(\varepsilon_{\boldsymbol{k}}^2 + \Delta^2\right)^{\frac{1}{2}} - \sigma\mu_B H,$$

where Δ is the superconducting energy gap. The resulting density of states is shown on the left side of Fig. 9.18. The splitting in the magnetic field manifests itself in the tunneling conductance as shown in Fig. 9.19. The two peaks or shoulders in the conductance curve are the results of the superposition of tunneling from/into the two spin bands of the ferromagnet and provide a direct

Table 9.2. Measured spin polarization values for various magnetic materials

Materials	$Ni_{80}Fe_{20}$	Ni	Co	Fe	C
Spin polarization measured by using Andreev reflection method [218]	37%	43%	43%	45%	-
Spin polarization measured on FM/AlOx/Al Tunnel junction [219]	48%	33%	42%	45%	5
Spin polarization measured by using spin-resolved photoemission [220]	-	-	-	-	-

measurement of the ratio of their densities of states, that is, the polarization. In Table 9.2 we list the polarizations obtained in this manner for several magnetic materials.

Andreev Reflection. In this technique a superconducting tip is placed in contact with the magnetic sample. Current in the superconductor is carried by electron pairs. Therefore, if a spin up electron from the ferromagnet enters the superconductor it must do so by pairing with a spin down electron created at the interface. This leaves a spin up hole which is reflected back into the ferromagnet. These "Andreev reflected" holes provide a parallel conduction channel. When the sample is unpolarized the conductance of the contact is twice that when the superconductor is normal. As the polization increases the current at zero bias decreases, becoming zero for a 100% spin polarized ferromagnet. Spin polarizations obtained by this technique are also listed in Table 9.2.

We notice that the polarizations deduced from the different techniques are different. Part of this difference has to do with the fact that the spin polarization is weighted differently in the different techniques. The quality of the tunneling barrier and the point contact are also important. The point contact transport, for example, may vary from being ballistic to diffusive.

Even if we had accurate values for the spin polarizations the Julliere expression (9.53) does not capture all the features experimentally observed in magnetic tunnel junctions. For example, the TMR shows a strong decrease with bias voltage. In the tunneling Hamiltonian approach described above we treated the tunneling matrix elements as constants as functions of energy. This ignores the details of the barrier. To incorporate the details of the barrier we must solve the tunneling problem.

The contribution of an electron in state $\boldsymbol{k}$ with spin σ to the tunneling current density, J_σ, is given by its probability of getting through the barrier,
$T(\boldsymbol{k})$, times the frequency with which it hits the barrier $\nu_{\boldsymbol{k}\sigma}$. Thus the current per unit area is

$$J_\sigma = \sum_{\boldsymbol{k}} \frac{e}{A} T(\boldsymbol{k}) \nu_{\boldsymbol{k},\sigma} f_{\boldsymbol{k}\sigma}, \tag{9.54}$$

where $f_{k\sigma}$ is the Fermi function and A is the area of the junction. The frequency of striking the barrier is $v_z/2L$ where L is the thickness of the electrode. Converting the sum over $\boldsymbol{k}$ into an integral according to

$$\sum_k \frac{AL}{(2\pi)^3} \iint dk_z dk_\parallel^2 \to \frac{V}{(2\pi)^3} \int dE/\left(dE/dk_z\right) \int dk_\parallel^2 \tag{9.55}$$

the current density becomes

$$J_\sigma = \frac{AL}{(2\pi)^3} \frac{e}{AL} \frac{1}{2\hbar} \int \frac{dE}{dE/dk_z} \int dk_\parallel^2 T\left(k_\parallel, E\right) \frac{dE}{dk_z} f_{\boldsymbol{k}\sigma}. \tag{9.56}$$

If we assume that the temperature is zero and that the energy dependence of $T\left(\boldsymbol{k}_\parallel, E\right)$ over the energy range eV is not large, then the integral over E just gives eV, and

$$J_\sigma = \frac{e^2 V}{(2\pi)^3} \frac{1}{2\hbar} \int dk_\parallel^2 T\left(k_\parallel\right)$$

The transmission probability $T\left(\boldsymbol{k}_\parallel\right)$is obtained by solving the 1D Schrödinger equation. For a simple rectangular barrier as shown in Fig. 9.16 the transmission probability is given by

$$T\left(k_\parallel\right) = \frac{16 k_1 \kappa^2 k_2 e^{-2d\kappa}}{\kappa^2 \left(k_1 + k_2\right)^2 + \left(\kappa^2 - k_1 k_2\right)^2}, \tag{9.57}$$

where

$$k_1 = \sqrt{\left(2m/\hbar^2\right)\left(E_F - V_1\right) - k_\parallel^2}$$
$$k_2 = \sqrt{2m/\hbar^2\left(E_F - V_2\right) - k_\parallel^2}$$
$$\kappa = \sqrt{\left(2m/\hbar^2\right)\left(V_b - E_F\right) + k_\parallel^2}$$

In the case of a magnetic tunnel junction we have four different spin orientations analogous to Fig. 9.8. These are shown in Fig. 9.20. Slonezewski [221] applied the free-electron results above to this situation and obtained an expression for the tunneling magnetoresistance of the same form as Julliere's, i.e.,

$$\mathrm{TMR} = \frac{2P_{\mathrm{eff}}^2}{1 - P_{\mathrm{eff}}^2} \tag{9.58}$$

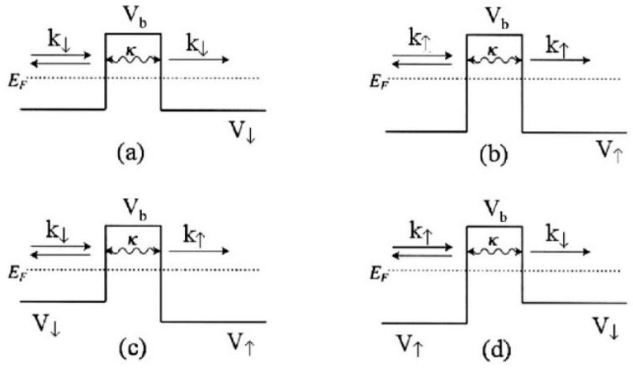


Fig. 9.20. Squarewell potentials seen by an electron moving in a trilayer consisting of two ferromagnets separated by a nonmagnetic insulating barrier

but with an effective polarization

$$P_{\text{eff}} = P\frac{\kappa_0^2 - k_\uparrow k_\downarrow}{\kappa_0^2 + k_\uparrow k_\downarrow}, \tag{9.59}$$

where $k_\uparrow$ and $k_\downarrow$ are evaluated at $k_\parallel = 0$, and κ_0 at $V_b = 0$. This simple free electron model illustrates how sensitive the tunneling magnetoresistance is to the details of the barrier. In Fig. 9.21 we show the voltage dependence of the TMR. Such a strong voltage dependence is a characteristic of magnetic tunnel junctions.

Neither Julliere's model nor the free-electron model can treat realistic electrode and barrier materials. To do this, one must employ a Kubo formalism that incorporates first-principles calculations of the electronic states. One of the great successes of such calculations was the prediction that a barrier consisting of MgO should give a high TMR [222]. Such large TMR's were subsequently observed [223].

9.2.5 Spin Transfer

In addition to the giant magnetoresistance that occurs in a trilayer structure when the current flows perpendicular to the plane (CPP) of the structure there is another interesting phenomenon. Namely, since the conduction elections in a ferromagnet are partially spin polarized by the *d*-electrons they can transfer spin angular momentum from one magnetic layer to the other. This can have the result of exciting spinwaves in the "receiving" layer or even completely reversing the magnetization of this layer. These possibilities were first predicted independently by Slonczewski [224] and Berger [225] and subsequently observed by various groups [226-228].

Let us begin by considering a single ferromagnetic layer embedded between two semi-infinite Cu leads. We shall assume this layer is relatively thick and the magnetization fixed. This situation may be analyzed by the Boltzmann

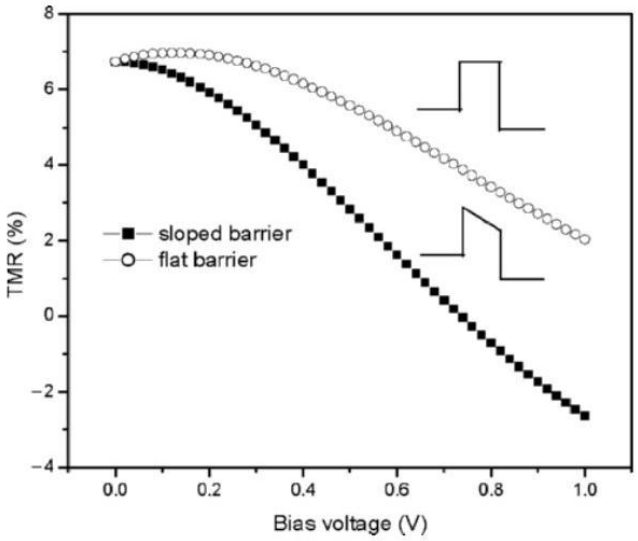


Fig. 9.21. Voltage dependence of TMR associated with shape of the barrier [Chando Park]

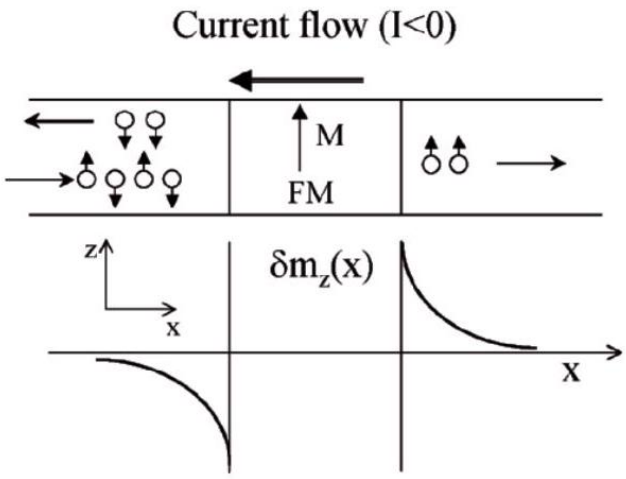


Fig. 9.22. Spin polarization in Cu leads attached to a ferromagnetic film based on a Boltzmann equation calculation [229]

equation approach we used in describing the GMR effect. Suppose we consider current flowing from right to left (which we define as a "negative" current) so that electrons are moving left to right. And suppose that the reflection coefficients are such that $R_\uparrow \ll R_\downarrow$. We also incorporate spin-flip scattering in the Cu by introducing a spin-flip mean free path, $\lambda_{\uparrow\downarrow} = v_F \tau_{\uparrow\downarrow}$ (for Cu this is estimated to be 2000 nm). If λ is the mean free path, then the spin diffusion length is given by $\sqrt{\lambda\lambda_{\uparrow\downarrow}}$. The result is that there is an imbalance in the number of up and down spins in the vicinity of the FM/Cu interface as shown in Fig. 9.22. We refer to this as spin accumulation.

Suppose we now add a second magnetic layer (free layer) with thickness d_2 separated from the first (reference layer) by a distance less than the spin diffusion length (Fig. 9.23). We now have spin-polarized electrons entering the second magnetic layer. These electrons interact with the magnetization in this layer through the $s-d$ exchange interaction, $-2J_{sd}\boldsymbol{s}\cdot\boldsymbol{S}(\boldsymbol{r})$. This interaction gives rise to a large effective field, of the order of 1000 T , acting on the the "incoming" conduction electrons. This causes the spins of these electrons to

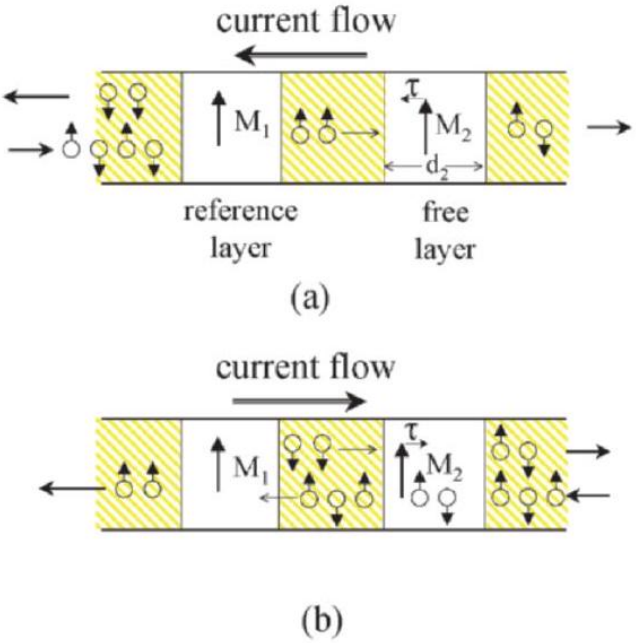


Fig. 9.23. Origin of the torque $(\boldsymbol{\tau})$ on a free layer due to spin accumulation from a reference layer}

precess rapidly around the direction of $\boldsymbol{S}(\boldsymbol{r})$. Slonczewski has shown [224] that the transverse component of the spin, after the electron has traveled a distance L_0 in the free layer of the order of $\pi/\left(k_{\uparrow}-k_{\downarrow}\right)$, where $k_{\uparrow}$ and $k_{\downarrow}$ are the spin-up and spin-down wavevectors, is zero. For Co L_0 is approximately 9.5 nm . By conservation of angular momentum this spin angular momentum is transferred to $\boldsymbol{S}(\boldsymbol{r})$. Each spin therefore contributes an angular momen $(\hbar/2)\sin\theta$ where θ is the angle between $\boldsymbol{s}$ and $\boldsymbol{S}$. The number of conduction electrons flowing through a cross sectional area in unit time is j/e, where j is the current density. Therefore, the amount of angular momentum transferred to the volume (unit area) d_2 in unit time, i.e., the torque, is $(\hbar/2ed_2)\, j\sin\theta$. We shall also introduce a spin polarization factor

$$P=\frac{n_{\uparrow}-n_{\downarrow}}{n_{\uparrow}+n_{\downarrow}}$$

to account for any loss in polarization as the electrons traverse the nonmagnetic separator. If we identify the direction of s with the magnetization of the reference layer, $\boldsymbol{M}_1$, and note that $\hat{\boldsymbol{M}}_1\times\hat{\boldsymbol{M}}_2=\sin\theta$, then the direction of the torque may be written as $\hat{\boldsymbol{M}}_2\times\left(\hat{\boldsymbol{M}}_1\times\hat{\boldsymbol{M}}_2\right)$. Since the magnetization is related to the angular momentum per unit volume through the gyromagnetic ratio γ we have

$$\frac{d\boldsymbol{M}_2}{dt}=\frac{\hbar\gamma P j}{2ed_2}\hat{\boldsymbol{M}}_2\times\left(\hat{\boldsymbol{M}}_1\times\hat{\boldsymbol{M}}_2\right). \tag{9.60}$$

The implications of this torque on the dynamics of the magnetization may be studied by adding this term to the Landau-Lifshitz-Gilbert equation we

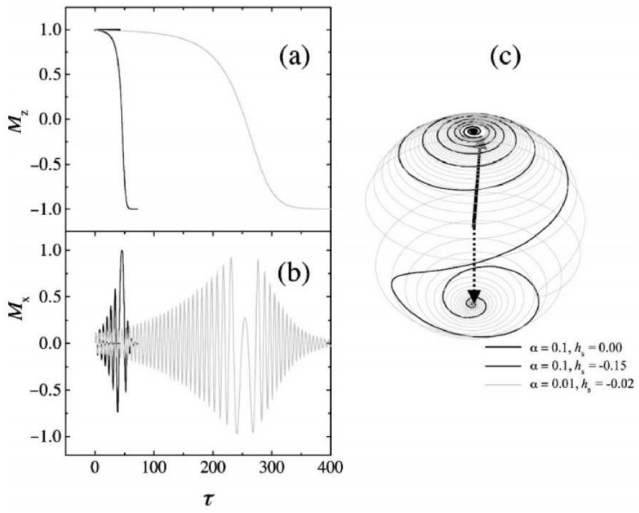


Fig. 9.24. The precession of magnetization under the influence of a spin current. Uniaxial anisotropy alone. (a) Time dependence of M_z. (b) Time dependence of M_x. (c) A 3D portrait of the spiral motion of the tip of M[230]h_s is the spin current amplitude in dimensionless units}

described in Sect. 8.4 This torque term adds a great deal of richness to the solutions. First of all, note that this term can be positive or negative. Consider, for example, Fig. 9.23a in which the two magnetic layers are initially parallel and the current is negative, i.e., electrons flow from the thick layer to the thin layer. In this case, (9.60) gives a torque that adds to the damping. In Fig. 9.23b the current is positive and the torque tends to pull the magnetization $\boldsymbol{M}_2$ away from the parallel state. Sun [230] has numerically solved the LLG equation including the spin transfer term. A solution showing switching behavior is shown in Fig. 9.24.

In the presence of a large applied magnetic field complete reversal of the magnetization is inhibited. In this case spinwaves are excited. The threshold for this excitation is readily obtained by expanding the magnetization $\boldsymbol{M}_2$ in (9.60) in terms of spinwave amplitudes according to (8.14) and adding this term to the linearized spinwave amplitude equation of motion. The result is [231]

$$\frac{d\alpha_k}{dt} = -i\omega_k\alpha_k - (\eta_k - \beta j)\,\alpha_k, \tag{9.61}$$

where $\beta = \hbar\gamma P/2ed_2M_2$ and ω_k is the frequency of the spinwave with wavevector k and η_k is the spinwave relaxation rate. The solution of this equation is

$$\alpha_k(t) = \alpha_k(0)e^{i\omega_k t}e^{-(\eta_k-\beta j)t}. \tag{9.62}$$

We see that the spinwave amplitude will grow exponentially when the current density exceeds the critical value

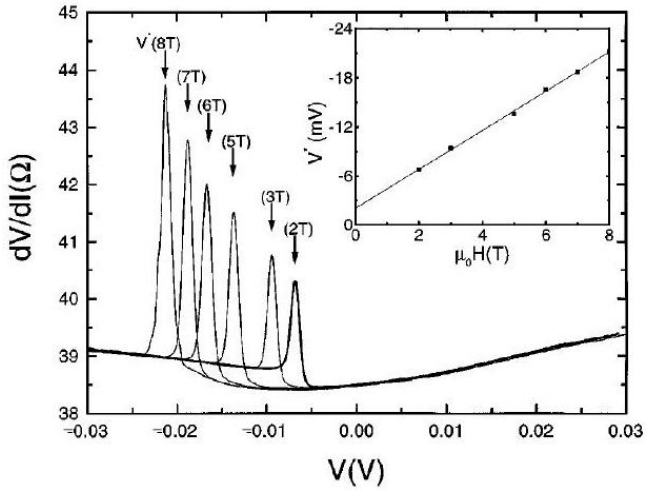


Fig. 9.25. The point contact $dV/dI(V)$ spectra for a series of magnetic fields ($2, 3, 5, 6, 7$, and 8 T) revealing an upward step and a corresponding peak in dV/dI at a certain negative bias voltage $V^{*}(H)$. The inset shows that $V^{*}(H)$ increases linearly with the applied magnetic field H [226]}

$$j_c = \frac{2ed_2M_2\eta_k}{\hbar\gamma P} \tag{9.63}$$

This "pumping" of spinwaves by a spin-polarized current is analogous to the Suhl instabilities we discussed in Sect. 8.2.6. In particular, this pumping can lead to chaotic phenomena as well [232].

The first observations of spin-polarized current-driven excitations in a magnetic multilayer were made by Tsoi, et al. [226]. Figure 9.25 shows the differential resistance of a Co/Cu multilayer stack excited by current injected through a point contact. We immediately see the asymmetry with current direction predicted above. The peaks scale linearly with field. In this experiment the magnetic field is applied perpendicular to the plane of the layers. In this configuration the frequency of the $\boldsymbol{k} = 0$ spinwave, or ferromagnetic resonance (FMR) mode is $\omega_{\boldsymbol{k}=0} = \gamma\,(H - 4\pi M_s)$. Therefore, Fig. 9.25 is evidence for the current-driven excitation of $\boldsymbol{k} = 0$ spinwaves.

9.2.6 Spin Hall Effect

Spin accumulation can also occur in the direction transverse to the current flow in a single conducting film. There are, in fact, two mechanisms by which this can occur. The most direct (Fig. 9.26) is the spin polarization that results from the response of the metal to the magnetic field produced by the current itself through $\boldsymbol{\nabla} \times \boldsymbol{H} = 4\pi\boldsymbol{j}/c$. This response is proportional to the Pauli susceptibility and is very small. A larger effect occurs through spin-orbit scattering, (Fig. 9.27), which is also the origin of the "anomalous" Hall effect in

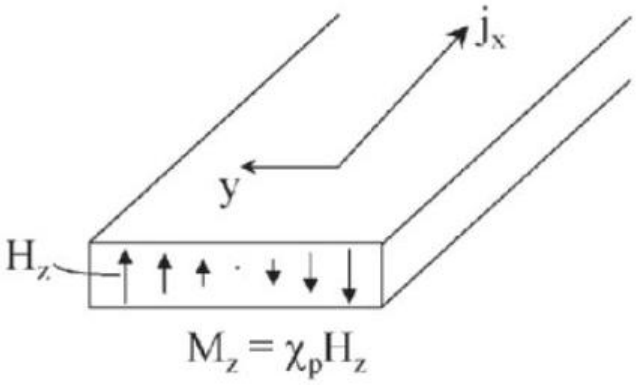


Fig. 9.26. Spin Polarization across film due to field of current

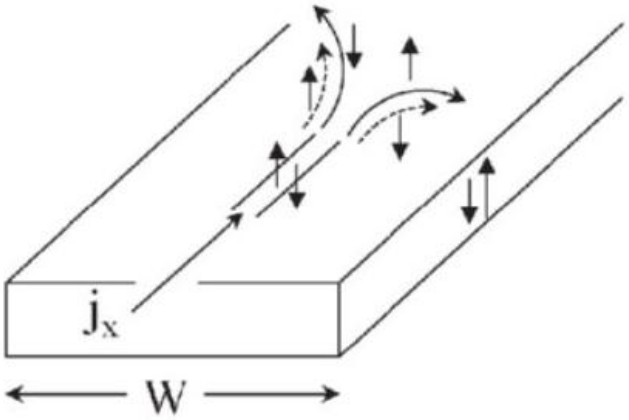


Fig. 9.27. Spin polarization at edges of film due to spin-dependent scattering

ferromagnets. These two mechanisms are quite different. In the case of "the Pauli polarization" the Fermi levels of the up- and down-spins are equal but vary across the film. In the case of spin-orbit scattering the Fermi levels of the up- and down-spins are not equal and are maintained in this nonequilibrium state by the current flow.

Let us consider the scattering process. When an electron is incident on a spinless scattering center the potential it sees is given by

$$V(r) = V_c(r) + V_s(r)\boldsymbol{\sigma} \cdot \boldsymbol{L}, \tag{9.64}$$

where $V_c(r)$ is the Coulomb potential and $V_s(r)$ is the spin-orbit scattering potential. As a result of the $V_s(r)$ term, up- and down-spin electrons will be scattered differently toward the left and right.

There are two effects of this spin-orbit scattering. In one, the amplitude of the scattered electron wave as it propagates away from the scattering center is different for spin-up and spin-down. This is called "skew scattering". In the other, the apparent center from which the scattered wave propagates is different for spin-up and spin-down. This is called "side jump". The side jump may also be understood as arising from an "anomalous" velocity. In particular, the velocity operator is obtained from the Hamiltonian by $\boldsymbol{v} = (1/\hbar)\partial\mathcal{H}/\partial\boldsymbol{k}$. If $\mathcal{H} = p^2/2m$ we obtain the usual velocity operator $\boldsymbol{v} = \hbar\boldsymbol{k}/m$. But when the spin-orbit potential (9.64) is included we obtain an additional "anomalous" contribution $(\hbar\alpha/\tau^\sigma)\,\boldsymbol{k} \times \boldsymbol{\sigma}$, where α is the spin-orbit coupling constant and τ^σ is the non-slip-flip relaxation time [233].

The sign and magnitude of the spin accumulation depends upon $V_s(r)$. If the current is polarized, as in an itinerant ferromagnet, then there will also be an imbalance of charge created. This is the anomalous Hall effect. The

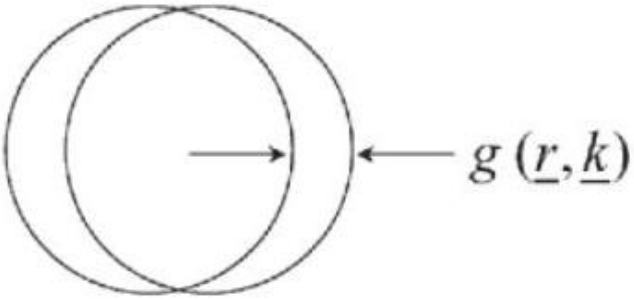


Fig. 9.28. Change in electron distribution function due to shift of Fermi sphere

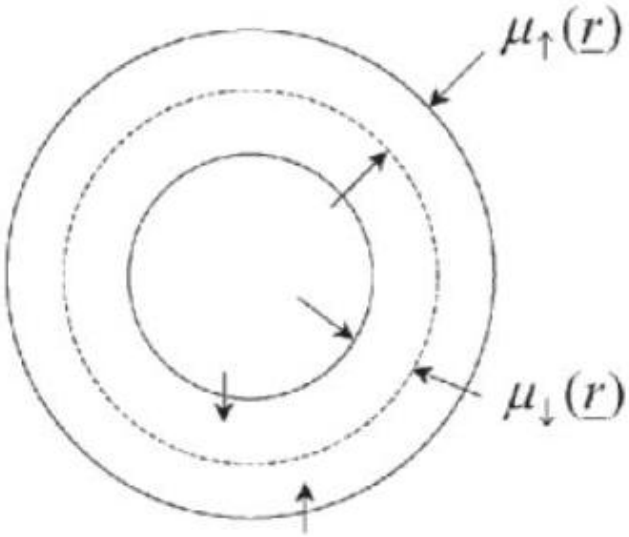


Fig. 9.29. Change in electron distribution due to change in Fermi radius

resistivity ($\mathcal{E}_y/j_x$) is written

$$\rho_H = R_0 B + 4\pi R_s M. \tag{9.65}$$

Therefore the experimental anomalous Hall coefficient R_s provides a measure of the spin Hall voltage. In particular, for a current of up spins $M = n_\uparrow \mu_B$ and the Hall voltage associated with (9.65) is

$$V_H = 4\pi R_s w j_x n_\uparrow \mu_B. \tag{9.66}$$

For down spins there will be an equal (with $n_\downarrow$) but opposite voltage. Thus, since $n_\uparrow = n_\downarrow = n/2$ in a paramagnet there will not be a charge buildup, but there will be a spin accumulation. We may therefore describe this as a spin Hall voltage [234],

$$V_{SH} = 2\pi R_s w j_x n \mu_B. \tag{9.67}$$

The spin accumulation will vary across the film. One approach to determining this spatial variation is to solve the Boltzmann equation introduced in (9.34) for the electron distribution

function $f^\sigma(\boldsymbol{r},\boldsymbol{k})$. In the case of GMR the change in the distribution function arises from a displacement of the Fermi surface (Fig. 9.28). In the case of spin accumulation we must also allow for a difference in the radii of the spin-up and spin-down Fermi surfaces (Fig. 9.29). Therefore, we describe the distribution function by two nonequilibrium terms,

$$f^\sigma(\boldsymbol{r},\boldsymbol{k}) = f^0(\boldsymbol{k}) + [g^\sigma(\boldsymbol{r},\boldsymbol{k}) + e\mu^\sigma(\boldsymbol{r})]\left(-\frac{\partial f^0}{\partial \varepsilon_F}\right). \tag{9.68}$$

We have explicitly recognized that this is a Fermi surface phenomenon by introducing the factor $(\ -\partial f^0/\partial\varepsilon_F\ )$, and we have written the k-independent part as a voltage. When this assumed form for $f^\sigma(\boldsymbol{r},\boldsymbol{k})$ is substituted into the Boltzmann equation, and a spin-flip relaxation term is included, we obtain

$$\boldsymbol{k}\cdot\nabla g^\sigma(\boldsymbol{r},\boldsymbol{k}) + e\boldsymbol{\varepsilon}^\sigma(\boldsymbol{r})\cdot\boldsymbol{k} = -g^\sigma(\boldsymbol{r},\boldsymbol{k})/\tau_\sigma + \frac{e\mu^\sigma(\boldsymbol{r}) - e\mu^{-\sigma}(\boldsymbol{r})}{\tau_{\uparrow\downarrow}} \tag{9.69}$$

where $\mathcal{E}^\sigma(\boldsymbol{r}) = \mathcal{E} - \nabla\mu^\sigma(\boldsymbol{r})$, showing that $\mu^\sigma(\boldsymbol{r})$ plays the role of a spindependent chemical potential. If we assume $\tau_{\uparrow\downarrow}$ is very long relative to τ_σ and that $g^\sigma(\boldsymbol{r},\boldsymbol{k})$ is slowly varying in space we find

$$g^\sigma(\boldsymbol{r},\boldsymbol{k}) = -e\tau_\sigma\mathcal{E}^\sigma(\boldsymbol{r})\cdot\boldsymbol{k} \tag{9.70}$$

and that $\mu^\sigma(\boldsymbol{r})$ satisfies the diffusion equation

$$\nabla^2\left[\mu^\sigma(\boldsymbol{r}) - \mu^{-\sigma}(\boldsymbol{r})\right] = \frac{\mu^\sigma(\boldsymbol{r}) - \mu^{-\sigma}(\boldsymbol{r})}{D^2} \tag{9.71}$$

where the spin diffusion coefficient is given by $\frac{1}{D^2} = \frac{1}{D_\uparrow^2} + \frac{1}{D_\downarrow^2}$ and $D_\sigma = \frac{\hbar k_F^\sigma}{m}\sqrt{\tau_\sigma\tau_{\uparrow\downarrow}/3}$. This equation has the solution

$$\mu^\uparrow(y) - \mu^\downarrow(y) = Ae^{y/D} + Be^{-y/D} \tag{9.72}$$

where the coefficients A and B are determined by the boundary conditions on the spin current density, $j_y^\sigma(y)$. In particular, for the geometry shown above, $j_y^\uparrow(y) + j_y^\downarrow(y) = 0$.

The current-density is given by the average of the current operator. When the spin-orbit interaction is included in the Hamiltonian the current operator is $\hbar\boldsymbol{k}/m$ plus the anomalous contribution introduced above, $(\hbar\alpha/\tau^\sigma)\,\boldsymbol{k}\times\boldsymbol{\sigma}$. Using (9.68) to calculate the current density analogous to (9.42) gives [235]

$$\boldsymbol{j}^\sigma(\boldsymbol{r}) = c^\sigma\mathcal{E}^\sigma(\boldsymbol{r}) + c_H^\sigma\mathcal{E}^\sigma(\boldsymbol{r})\times\boldsymbol{\sigma} \tag{9.73}$$

where

$$c^\sigma = \frac{e^2\tau^\sigma\left(k_F^\sigma\right)^3}{6\pi^2 m}$$
$$c_H^\sigma = \frac{e^2\alpha\left(k_F^\sigma\right)^3}{6\pi^2}$$

Notice that $\mu^\sigma(\boldsymbol{r})$ enters the current density through $\mathcal{E}^\sigma(\boldsymbol{r})$. Taking the gradient of (9.72) the boundary condition enables us to determine the coefficients A and B. For a paramagnetic metal where $c^\uparrow = c^\downarrow$ and $c_H^\uparrow = c_H^\downarrow$,

$$\mu^\uparrow(y) - \mu^\downarrow(y) = \frac{c_H}{2c^2}wj\times\frac{2D}{w}\frac{\sinh(y/D)}{\cosh(w/2D)}$$

Figure 9.30 shows $\left(\mu^{\uparrow} - \mu^{\downarrow}\right) / \left(c_H w j_x / 2c^2\right)$ as a function of the distance across the film. From (9.65) $\varepsilon_y / j_x = 4\pi R_s M$. Also, since $\varepsilon_y / j_x = \sigma_{xy} / \sigma_{xx}^2 = c_H / c^2$, the value of $\mu^{\uparrow}(\pm w/2) - \mu^{\downarrow}(\pm w/2)$ is the same as Hirch's V_{SH} (9.67) when $w \ll D$.

Kato et al. [234] have used Kerr rotation microscopy to image the spin polarization across a semiconductor channel and observed a dependence very much like curve (c) in Fig. 9.30.

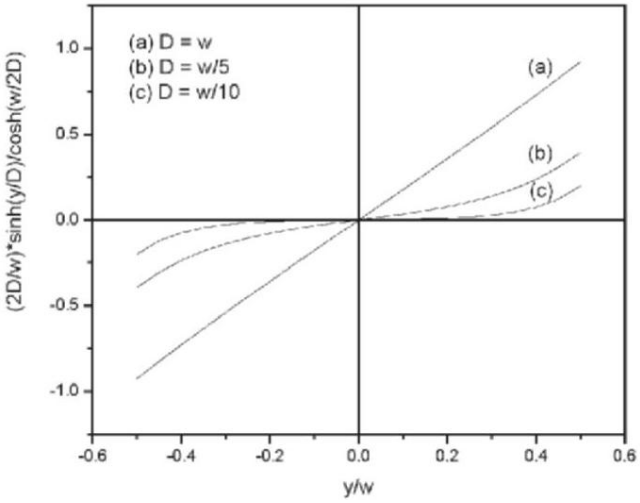


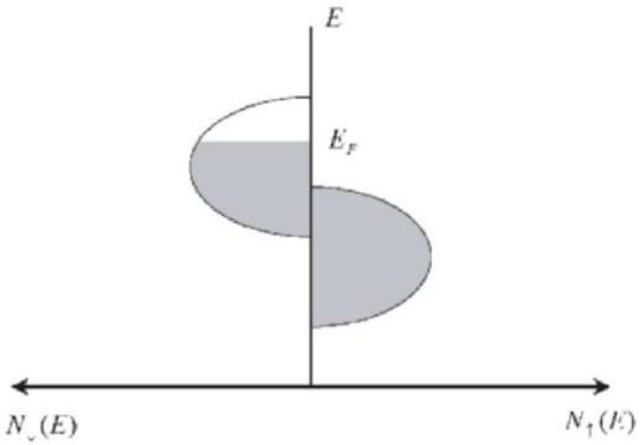
Fig. 9.30. Predicted dependence of spin accumulation as a function of distance across a semiconductor channel

Problems

9.1. Can the presence of a localized spin in a metal induce a uniform magnetization? (Hint: determine how the average spin density, $s(\boldsymbol{r})$, averaged over the range $0 < q < \frac{1}{r}$ varies with r.)

9.2. Discuss how you would expect the RKKY interraction to be modified by (a) exchange enhancement and (b) a finite electron mean free path (see [237]).

9.3. What value of GMR would be expected for magnetic layers having the following density of states:



9.4. Solve the Boltzmann equation for a single layer of Cu bounded by insulating magnetic layers which completely spin polarize the Cu electrons when they are reflected from the interface. What is the GMR?

10.11 Neutron Scattering

The scattering of neutrons provides us with an exceedingly powerful probe for studying condensed matter. This usefulness arises from the fact that thermal (300 K) neutrons have a wavelength $\lambda \simeq 1.6\text{Å}$. This means that their energy and wavelength are comparable to those of the excitations in solids. In fact, neutron scattering has become almost the last word in determining spin orderings, spin-density distributions, and spin-wave dispersion relations. A detailed discussion of the experimental techniques employed in neutron scattering, with reference to the original papers, may be found in [238].

There are two sources of neutrons for scattering. One is a nuclear reactor in which a fissile nucleus, generally ^{239}U or ^{239}Pu, captures a neutron and splits into two nearly equal-mass fragments together with approximately 2.5 neutrons, each having an average energy of 2 MeV . The second source of neutrons is spallation, in which high energy particles from an accelerator, such as 1 GeV protons, collide with a heavy nucleus, such as uranium or lead. The excited nucleus decays by evaporating neutrons. The spallation process produces approximately 20 neutrons, with average energies of 2 MeV , for each colliding proton. Most spallation sources operate in a pulsed mode with pulse widths less than one microsecond and pulse rates between 10 Hz and 60 Hz . The MeV neutrons must be slowed down to thermal velocities ($\sim$ 30meV), or even colder, by passing them through moderators consisting of beryllium or heavy water (D_2O).

10.11.1 10.1 Neutron Scattering Cross Section

Let us consider sending a collimated beam of monochromatic neutrons into a sample and measuring the energy spectrum of the neutrons scattered in some direction. This is the differential scattering cross section. For the rest of this section we shall be concerned with how the scattering cross section depends on the state of the crystal [239, 240].

If the neutrons interact with the scatterer through some interaction $V(\boldsymbol{r})$, then the probability that the neutron will be scattered from an initial state ($\boldsymbol{k}, m_s$) into a final state ($\boldsymbol{k}', m_s'$) while the system goes from state α to α' is, in the Born approximation,

$$\frac{2\pi}{\hbar}\left|\langle\alpha', m_s'|\,\frac{1}{L^3}\int d\boldsymbol{r}e^{i\boldsymbol{\kappa}\cdot\boldsymbol{r}}V(\boldsymbol{r})\right|\alpha, m_s\Big\rangle\,|^2\,\delta\left(\frac{\hbar^2k'^2}{2m_0}+E_{\alpha'}-\frac{\hbar^2k^2}{2m_0}-E_\alpha\right), \tag{10.1}$$

where $\boldsymbol{\kappa}\equiv\boldsymbol{k}-\boldsymbol{k}'$, L^3 is the volume of the sample, and m_0 is the neutron mass. The number of neutrons scattered into the solid angle $d\Omega'$ in the direction $\boldsymbol{k}'$ with momentum between $\hbar k'$ and $\hbar\,(k'+dk')$ is obtained by multiplying (10.1) by the number of states in this incremental volume of phase space, which is

$$\frac{k'^2dk'd\Omega'}{(2\pi/L)^3}. \tag{10.2}$$

Since the neutron energy and momentum are related by $E'=\hbar^2k^2/2m_0$, this becomes

$$\frac{m_0L^3k'}{8\pi^3\hbar^2}d\Omega'dE'. \tag{10.3}$$

The differential scattering cross section is defined as the scattering intensity per unit solid angle per unit incident flux. The incident flux associated with a beam of neutrons with velocity v is

$$\frac{v}{L^3}=\frac{\hbar k}{m_0L^3}. \tag{10.4}$$

Therefore the differential scattering cross section becomes

$$\left.\frac{d^2\sigma}{d\Omega' dE'}\right|_{\alpha,m_s\to\alpha',m_s'} = \frac{k'}{k}\left(\frac{m_0}{2\pi\hbar^2}\right)^2 \left|\left\langle \alpha' m_s'\right| V(-\kappa) \left|\, \alpha, m_s\right\rangle\right|^2 \times \delta\left(\frac{\hbar^2 k'^2}{2m_0} + E_{\alpha'} - \frac{\hbar^2 k^2}{2m_0} - E_\alpha\right). \tag{10.5}$$

We must now sum over the final states α' and m_s' as well as average over the initial states α and m_s, using a probability factor P_α. This factor might have the Boltzmann form

$$P_\alpha = \frac{\exp\left(-E_\alpha/k_B T\right)}{\sum_\alpha \exp\left(-E_\alpha/k_B T\right)} \tag{10.6}$$

or some other form such as the Bose-Einstein form in the case of phonons. Thus

$$\left.\begin{aligned} \tfrac{d^2\sigma}{d\Omega' dE'} = {} & \textstyle\sum_{\alpha,m_s} P_\alpha P_{m_s} \sum_{\alpha' m_s'} \tfrac{k'}{k}\left(\tfrac{m_0}{2\pi\hbar^2}\right)^2 \left|\left\langle \alpha', m_s'\right| V(-\kappa) \left|\, \alpha, m_s\right\rangle\right|^2 \\ & \times\delta\left(\tfrac{\hbar^2 k'^2}{2m_0} + E_{\alpha'} - \tfrac{\hbar^2 k^2}{2m_0} - E_\alpha\right). \end{aligned}\right\} \tag{10.7}$$

This is the general result which forms the basis of our discussion of neutron scattering.

The neutron interacts with the crystal in two ways: through the nuclear interaction and through the magnetic dipole-dipole interaction.

10.11.2 10.2 Nuclear Scattering

Although our main interest will be in magnetic scattering, we shall briefly consider the nuclear contribution to see how it complicates the experimental results. Let us begin by considering the nuclear scattering from a single fixed zero-spin nucleus located at $\boldsymbol{R}$. Since the interaction has such a short range, the scattering will be isotropic at these neutron energies. This is because at short distances from the nucleus the neutron's angular momentum is very small. Hence we have only s-wave scattering. The only form of $\boldsymbol{V}(\boldsymbol{r})$ which gives isotropic scattering in the Born approximation is a delta function.

Thus we set

$$V(\boldsymbol{r}) = \frac{2\pi\hbar^2}{m_0} b\delta(\boldsymbol{r} - \boldsymbol{R}) \tag{10.8}$$

where b is called the scattering length. It represents a phenomenological description of the nuclear scattering process which is very difficult to compute directly. Equation (10.8) is called the Fermi pseudopotential and gives a total cross section

$$\sigma = 4\pi b^2 \tag{10.9}$$

The scattering length is taken to be a constant, independent of the neutron energy. This is valid at the low neutron energies used.

Now consider scattering from many fixed nuclei situated at lattice sites $\boldsymbol{n}$. If we have a compound or if there are isotopes present, the scattering amplitudes of the different nuclei may be different. In this case the total pseudopotential is

$$V(\boldsymbol{r}) = \frac{2\pi\hbar^2}{m_0} \sum_{\boldsymbol{n}} b_{\boldsymbol{n}} \sigma(\boldsymbol{r} - \boldsymbol{n}) \tag{10.10}$$

This leads to the differential cross section

$$\frac{d^2\sigma}{d\Omega' dE'} = \sum_{\alpha, m_s} P_\alpha P_{m_s} \sum_{\alpha', m_s'} \frac{k'}{k} \left| \langle \alpha', m_s' | \sum_{\boldsymbol{n}} b_{\boldsymbol{n}} e^{i\boldsymbol{\kappa}\cdot\boldsymbol{n}} \right| \alpha, m_s \Big\rangle |^2 \, \delta \quad \text{(energy)}. \tag{10.11}$$

Here, α refers to all the quantum numbers necessary to describe the crystal. In particular, α and α' might refer to the isotope distribution and the nuclearspin orientations. To a good approximation the energy is independent of these quantities. Thus,

$$\begin{aligned} \frac{d\sigma}{d\Omega'} &= \sum_{\alpha, m_s} P_\alpha P_{m_s} \sum_{,} m \exp[i\boldsymbol{\kappa}\cdot(\boldsymbol{n}-\boldsymbol{m})] \langle \alpha, m_s | \, b_m^* b_n \, | \alpha, m_s \rangle \\ &= \sum_{n,m} \exp[i\boldsymbol{\kappa}\cdot(\boldsymbol{n}-\boldsymbol{m})] \langle b_m^* b_n \rangle \end{aligned} \tag{10.12}$$

Since there is no correlation between $b_{\boldsymbol{n}}$ and $b_{\boldsymbol{m}}^*$ for $\boldsymbol{n} \neq \boldsymbol{m}$,

$$\langle b_{\boldsymbol{m}}^* b_{\boldsymbol{n}} \rangle = \langle b_{\boldsymbol{m}}^* \rangle \langle b_{\boldsymbol{n}} \rangle = |\langle b \rangle|^2$$

Therefore we write

$$\langle b_{\boldsymbol{m}}^* b_{\boldsymbol{n}} \rangle = |\langle b \rangle|^2 + \left(\langle |b|^2 \rangle - |\langle b \rangle|^2 \right) \Delta(\boldsymbol{n}-\boldsymbol{m}). \tag{10.13}$$

This enables us to separate the cross section into two parts,

$$\frac{d\sigma}{d\Omega'} = \left(\frac{d\sigma}{d\Omega'} \right)_{\text{coh}} + \left(\frac{d\sigma}{d\Omega'} \right)_{\text{incoh}} \tag{10.14}$$

where the coherent cross section is

$$\left(\frac{d\sigma}{d\Omega'} \right)_{\text{coh}} = |\langle b \rangle|^2 \left| \sum_n e^{i\boldsymbol{\kappa}\cdot\boldsymbol{n}} \right|^2 \tag{10.15}$$

and the incoherent cross section is

$$\left(\frac{d\sigma}{d\Omega'} \right)_{\text{incoh}} = N \left(\langle |b|^2 \rangle - |\langle b \rangle|^2 \right). \tag{10.16}$$

We see that only the mean scattering potential $\langle b \rangle$ gives rise to interference effects and coherent scattering, while the incoherent scattering is proportional to the mean-square deviation $\langle |b - \langle b \rangle|^2 \rangle$. If the nucleus possesses a nonzero spin, then the nuclear scattering amplitude will depend upon the relative orientation of this spin with the neutron spin. Since these nuclear spins are disordered, they lead to a large incoherent scattering. This is beautifully illustrated in Fig. 10.1, where the cross section for thorium hydride is compared with that of thorium deuteride. The large incoherent scattering observed in the ThH_2 is associated with incoherent nuclear spin scattering.

10.2.1 Bragg Scattering

The coherent elastic scattering from a rigid lattice is also referred to as Bragg scattering. This involves the quantity.

$$\sum_n e^{i\kappa\cdot n}$$

which contains N terms. It can be shown that

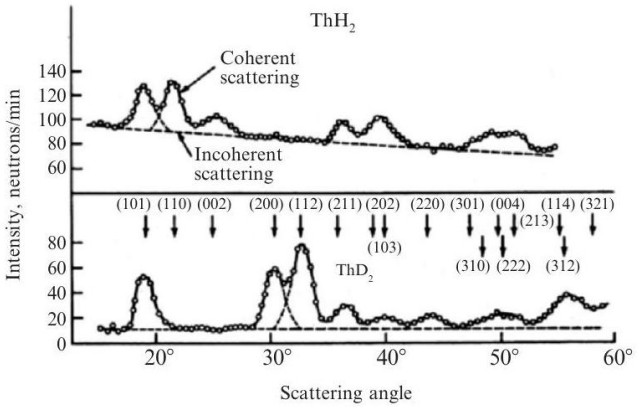


Fig. 10.1. Neutron diffraction patterns for polycrystalline $\mathrm{ThH_2}$ and $\mathrm{ThD_2}$[241]

$$\left|\sum_{\boldsymbol{n}} e^{i\boldsymbol{\kappa}\cdot\boldsymbol{n}}\right|^2 = \frac{(2\pi)^2 N}{V_0}\sum_{G}\delta(\boldsymbol{\kappa}-\boldsymbol{G}), \tag{10.17}$$

where V_0 is the volume of the unit cell and $\boldsymbol{G}$ is a reciprocal lattice vector. This is an important result which we shall encounter many times. The coherent, or Bragg, cross section then becomes

$$\left(\frac{d\sigma}{d\Omega'}\right)_{\mathrm{coh}} = |\langle b\rangle|^2\frac{(2\pi)^3 N}{V_0}\sum_{\boldsymbol{G}}\delta(\boldsymbol{\kappa}-\boldsymbol{G}), \tag{10.18}$$

If the lattice has more than one atom per unit cell, say r atoms at $\rho_1, \rho_2, \ldots$, relative to some reference point in the unit cell, then

$$\left(\frac{d\sigma}{d\Omega'}\right)_{\mathrm{coh}} = \frac{(2\pi)^3 N}{V_0}\sum_{G}|F(\boldsymbol{G})|^2\delta(\boldsymbol{\kappa}-\boldsymbol{G}), \tag{10.19}$$

where

$$F(\boldsymbol{G}) = \sum_{i=1}^{r}\langle b_i\rangle\exp\left(i\boldsymbol{G}\cdot\boldsymbol{\rho}_i\right)$$

is the unit cell structure factor. The important point is that there is no coherent scattering unless

$$\kappa = k - k' = G. \tag{10.20}$$

By squaring this relation and making use of the fact that the magnitude of the reciprocal lattice vector in some direction hkl is equal to some multiple of 2π times the inverse spacing of the planes in that direction, d (hkl), we obtain the familiar Bragg law

$$\sin\theta = \frac{n\lambda}{2d(\mathrm{hkl})} \tag{10.21}$$

for the angle 2θ between the indicent beam and the diffracted beam.

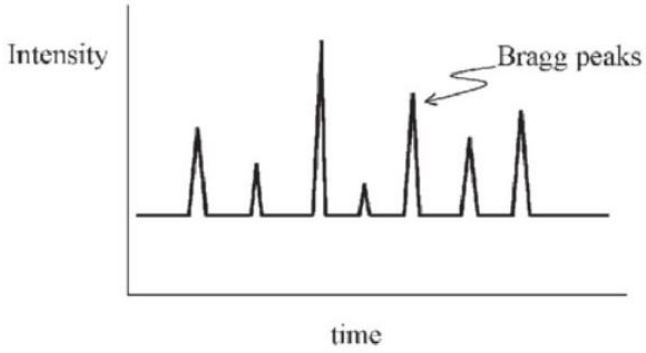


Fig. 10.2. Time-of-flight diffraction pattern

Time-of-flight differaction provides a convenient technique for measuring the lattice parameters d(hkl). In a pulsed neutron source the neutrons have a broad energy spectrum similar to the spectrum of fission neutrons. But since they are all created at the same time, time-of-flight techniques may be used to separate neutrons with different energies. In particular, combining the de Broglie relation $\lambda = h/mv$ with Bragg's law, $\lambda = 2d\sin\theta$, the d-spacing in Å at a given scattering angle is

$$d = t/252.8L2\sin\theta,$$

where t is the time of flight in μ sec and L is the flight path in m. Thus a time-of-flight diffraction pattern will look like the one in Fig. 10.2. One advantage of a time-of-flight measurement is that the resolution, $\Delta d/d$, is constant over the whole pattern.

10.2.2 Scattering of Phonons

Let us now consider the cross section when we allow the ions to deviate from their rigid lattice position $\boldsymbol{n}$ by an amount $\boldsymbol{u}_n$. The general position is then

$$\boldsymbol{R}_n = \boldsymbol{n} + \boldsymbol{u}_n, \tag{10.22}$$

and the cross section becomes

$$\frac{d^2\sigma}{d\Omega' dE'} = \sum_{a,m_s} P_a P_{m_{s'}} \sum \frac{k'}{k} \times \left| \langle \alpha', m_s' | \sum_n b_n \exp(i\boldsymbol{\kappa}\cdot\boldsymbol{n}) \exp\left(i\boldsymbol{\kappa}\cdot\boldsymbol{u}_n\right) \right| \alpha, m_s \Big\rangle |^2 \, \delta(\text{ energy }). \tag{10.23}$$

The system coordinates a and a' now refer to phonon states as well as to the nuclear spin projections m_I and m_I'. Writing out the latter explicitly, we have for the matrix element above

$$\sum_n \langle m_s', m_I' | \, b_n \, | m_s, m_I \rangle \exp(i\boldsymbol{\kappa}\cdot\boldsymbol{n}) \, \langle \alpha' | \exp\left(i\boldsymbol{\kappa}\cdot\boldsymbol{u}_n\right) | \alpha \rangle. \tag{10.24}$$

The ionic displacement $\boldsymbol{u}_n$ may be expanded in boson operators corresponding to the creation and annihilation of phonons as [242].

$$\boldsymbol{u_n} = \sum_{\boldsymbol{q},s} \sqrt{\frac{\hbar}{2NM\omega_{\boldsymbol{qs}}}} \left[a_{\boldsymbol{qs}} e^{i\boldsymbol{q}\cdot\boldsymbol{n}} \hat{\epsilon}_{\boldsymbol{qs}} + a^\dagger_{\boldsymbol{qs}} e^{-i\boldsymbol{q}\cdot\boldsymbol{n}} \hat{\epsilon}_{\boldsymbol{qs}} \right]. \tag{10.25}$$

$\hat{\epsilon}_{\boldsymbol{qs}}$ is the unit polarization vector associated with the phonon wave vector $\boldsymbol{q}$ and polarization s. The phonon coordinates α in the matrix element above may now be explicitly written as the number of phonons in each mode. Thus the factor involving $\boldsymbol{u_n}$ becomes

$$\begin{aligned}&\left\langle n'_{\boldsymbol{q}_1 s_1}, n'_{\boldsymbol{q}_2, s_2, \ldots}\right| \\ &\times \exp\left[i\sum_{\boldsymbol{q},s}\sqrt{\frac{\hbar}{2NM\omega_{\boldsymbol{qs}}}}\left(a_{\boldsymbol{qs}}e^{\,i\boldsymbol{q}\cdot\boldsymbol{n}_{\hat{\epsilon}_{\boldsymbol{qs}}}\cdot\boldsymbol{\kappa}}+a^{\dagger}_{\boldsymbol{qs}}e^{-i\boldsymbol{q}\cdot\boldsymbol{n}_{\hat{\epsilon}_{\boldsymbol{qs}}}\cdot\boldsymbol{\kappa}}\right)\right] \\ &\times \left|n_{\boldsymbol{q}_1 s_1}, n_{\boldsymbol{q}_2, s_2}, \ldots\right\rangle\end{aligned} \tag{10.26}$$

Using the relation

$$e^{A+B} = e^A e^B e^{-(1/2)[A,B]}, \tag{10.27}$$

where A and B are noncommuting operators, and expanding the exponentials involving the phonon amplitudes, the cross section may be written as the sum of multiphonon processes,

$$\frac{d^2\sigma}{d\Omega' dE'} = \sum_{n=-\infty}^{\infty}\frac{d^2\sigma_n}{d\Omega' dE'} \tag{10.28}$$

where n is the net number of phonons created.

Elastic Scattering. Let us consider the case where the net number of phonons created is 0 . This corresponds to elastic scattering, but at a finite temperature. With $n'_{\boldsymbol{qs}} = n_{\boldsymbol{qs}}$, the matrix element $n = 0$ in (10.28) becomes, to order $1/N$,

$$\prod_{\boldsymbol{qs}}\left[1 - \frac{\hbar}{2NM\omega_{\boldsymbol{qs}}}\left(n_{\boldsymbol{qs}} + \frac{1}{2}\right)|\boldsymbol{\kappa}\cdot\hat{\varepsilon}_{\boldsymbol{q},s}|^2\right]. \tag{10.29}$$

If this is substituted into the expression for $d^2\sigma_0/d\Omega' dE'$, the sum over α, which is the number of phons in each mode, converts $n_{\boldsymbol{q}_s}$ into its thermal equilibrium value $\bar{n}_{\boldsymbol{qs}}$. In the limit $N \to \infty$ this may be written as

$$\exp\left[-\sum_{\boldsymbol{q},s}\frac{\hbar}{2NM\omega_{\boldsymbol{qs}}}\left(\bar{n}_{\boldsymbol{qs}} + \frac{1}{2}\right)|\boldsymbol{\kappa}\cdot\hat{\varepsilon}_{\boldsymbol{qs}}|^2\right]. \tag{10.30}$$

Using this along with (10.24,10.24) results in a total elastic cross section,

$$\begin{aligned}\frac{d^2\sigma_0}{d\Omega' dE'} = &\sum_{m_I, m_s} P_{m_s} P_{m_I} \sum_{m'_I, m'_s}\frac{k'}{k}\sum_{n,m}\langle m_s, m_I|\, b^*_m\,|m'_s, m'_I\rangle \\ &\times \langle m'_s, m'_I|\, b_n\,|m_s, m_I\rangle \exp[i\boldsymbol{\kappa}\cdot(\boldsymbol{n}-\boldsymbol{m})] \\ &\times \exp\left[-\sum_{\boldsymbol{q},s}\frac{\hbar}{NM\omega_{\boldsymbol{qs}}}\left(n_{\boldsymbol{qs}} + \frac{1}{2}\right)|\boldsymbol{\kappa}\cdot\hat{\varepsilon}_{\boldsymbol{qs}}|^2\right]\delta(\text{ energy })\end{aligned} \tag{10.31}$$

Removing the sum over (m'_s, m'_I) by closure and integrating over energy, we obtain

$$\frac{d\sigma_0}{d\Omega'} = \sum_{\boldsymbol{n},\boldsymbol{m}}\exp[i\boldsymbol{\kappa}\cdot(\boldsymbol{n}-\boldsymbol{m})]\,\langle b^*_m b_n\rangle\exp(-2W), \tag{10.32}$$

where

$$W = \sum_{\boldsymbol{q},s} \frac{\hbar}{2NM\omega_{\boldsymbol{q}s}} \left(\bar{n}_{\boldsymbol{q}s} + \frac{1}{2}\right) |\boldsymbol{\kappa} \cdot \hat{\varepsilon}_{\boldsymbol{q}s}|^2 \tag{10.33}$$

is called the Debye-Waller factor. Comparison of (10.32) and (10.12) shows that the coherent and incoherent cross sections are the same as those we obtained for the rigid lattice case except for the factor $\exp(-2W)$. Notice that this is nonzero even at $T = 0$ owing to zero-point vibrations.

One-Phonon Scattering. We now turn to inelastic processes involving one phonon. In particular, let us consider those processes in which the number of phonons annihilated exceeds by 1 the number of phonons created. The contribution to the matrix element given by (10.26) for these processes is

$$\begin{aligned} &\sum_n \langle m_s', m_I'| b_n |m_s, m_I\rangle \, e^{i\boldsymbol{\kappa}\cdot\boldsymbol{n}} i \sqrt{\frac{\hbar}{2NM\omega_{\boldsymbol{q}s}}} \sqrt{n_{\boldsymbol{q}s}} e^{i\boldsymbol{q}\cdot\boldsymbol{n}} \boldsymbol{\kappa} \cdot \hat{\varepsilon}_{\boldsymbol{q}s} \\ &\sum_{\boldsymbol{q}',s'} \left[1 - \frac{\hbar}{2NM\omega_{\boldsymbol{q}'s'}} \left(n_{\boldsymbol{q}'s'} + \frac{1}{2}\right) |\boldsymbol{\kappa} \cdot \hat{\varepsilon}_{\boldsymbol{q}s}|^2\right] \end{aligned} \tag{10.34}$$

In the computation of the corresponding cross section the sum over the distribution of occupation numbers again introduces the thermal averages of the $n_{\boldsymbol{q}s}$. Thus we obtain

$$\begin{aligned} \frac{d^2\sigma_{-1}}{d\Omega' dE'} = &\sum_{\boldsymbol{q},s} \langle b_m^* b_n \rangle \exp[i(\boldsymbol{\kappa} + \boldsymbol{q}) \cdot (\boldsymbol{n} - \boldsymbol{m})] \\ &\times \frac{\hbar}{2NM\omega_{\boldsymbol{q}s}} \bar{n}_{\boldsymbol{q}s} |\boldsymbol{\kappa} \cdot \hat{\varepsilon}_{\boldsymbol{q}s}|^2 e^{-2W} \delta\left[\frac{\hbar^2}{2m_0}\left(k^2 - k'^2\right) - \hbar\omega_{\boldsymbol{q}}\right] \end{aligned} \tag{10.35}$$

From (10.10), (10.13), (10.15), (10.16) we may write

$$\langle b_m^* b_n \rangle = \frac{\sigma_{\text{coh}}}{4\pi} + \frac{\sigma_{\text{incoh}}}{4\pi} \Delta(\boldsymbol{m} - \boldsymbol{n}), \tag{10.36}$$

where $\sigma_{\text{coh}} = 4\pi |\langle b \rangle|^2$ and $\sigma_{\text{incoh}} = 4\pi \left\{ \langle |b|^2 \rangle - |\langle b \rangle|^2 \right\}$, so that the coherent part of the cross section becomes

$$\begin{aligned} \left(\frac{d^2\sigma_{-1}}{d\Omega' dE'}\right)_{\text{coh}} = &\frac{(2\pi)^3 N}{V_0} \frac{\sigma_{\text{coh}}}{4\pi} \sum_{\boldsymbol{q},s} \frac{k'}{k} \sum_{\boldsymbol{G}} \delta(\boldsymbol{\kappa} + \boldsymbol{q} - \boldsymbol{G}) \\ &\times \frac{\hbar}{2NM\omega_{\boldsymbol{q}s}} \bar{n}_{\boldsymbol{q}s} |\boldsymbol{\kappa} \cdot \hat{\varepsilon}_{\boldsymbol{q}s}|^2 e^{-2W} \delta \text{ (energy)} \end{aligned} \tag{10.37}$$

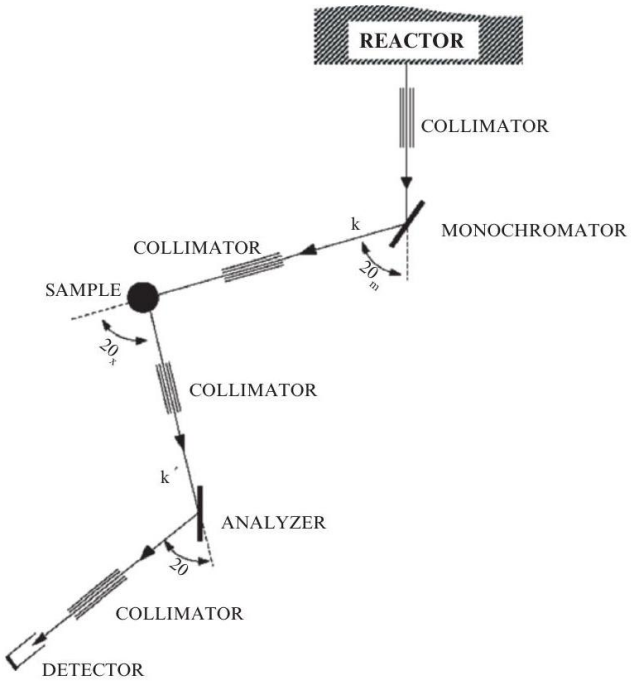


Fig. 10.3. Schematic diagram of a triple-axis spectrometer

Similarly, the incoherent part becomes

$$\left(\frac{d^2\sigma_{-1}}{d\Omega' dE'}\right)_{\text{incoh}} = \frac{N\sigma_{\text{incoh}}}{4\pi}\sum_{\boldsymbol{q},s}\frac{k'}{k}\frac{\hbar}{2NM\omega_{\boldsymbol{q}s}}\bar{n}_{\boldsymbol{q}s}\left|\boldsymbol{\kappa}\cdot\hat{\varepsilon}_{\boldsymbol{q}s}\right|^2 e^{-2W}\delta(\text{ energy }). \tag{10.38}$$

We can obtain similar expressions for one-phonon creation processes. For more details we refer the reader to [239].

In summary, a typical neutron diffraction pattern will consist of Bragg peaks plus a diffuse background. The Bragg peaks enable us to determine the crystal structure. The diffuse background arises from three sources: incoherent elastic scattering, coherent inelastic scattering, and incoherent inelastic scattering. By using certain experimental techniques we may separate the coherent inelastic scattering contribution and thereby determine the phonon dispersion relation. If we can also separate the incoherent inelastic scattering, this leads directly to the phonon density of states. In the next section we shall see how magnetic scattering contributes to this spectrum.

A triple-axis spectrometer (Fig. 10.3) is often used to measure the energy and momentum transfer associated with the inelastic scattering of neutrons from a reactor source [243]. The high-energy neutrons are moderated down to thermal energies and emerge from the reactor with a Maxwellian distribution of energies. A single energy is selected by Bragg reflection from the first monochromator crystal in Fig. 10.3. This monoenergetic beam is incident on the sample with the wave vector $\boldsymbol{k}$. The scattered beam $\boldsymbol{k}'$ is Bragg reflected by the analyzer crystal into the detector.

10.11.3 10.3 Magnetic Scattering

We now turn our attention to the interaction between the magnetic moment of the neutron and those of the crystal. The vector potential at $\boldsymbol{r_e}$ produced by a neutron's magnetic moment $\boldsymbol{\mu}$ at $\boldsymbol{r}$ is

$$\boldsymbol{A} = \frac{\boldsymbol{\mu} \times (\boldsymbol{r_e} - \boldsymbol{r})}{|\boldsymbol{r_e} - \boldsymbol{r}|^3}. \tag{10.39}$$

As we saw in Sect. 2.2, the magnetic interaction between a vector potential $\boldsymbol{A}$ and an electron with momentum $\boldsymbol{p}$ is given by

$$\boldsymbol{V}(\boldsymbol{x}) = -\frac{e}{2mc}[\boldsymbol{p} \cdot \boldsymbol{A}(\boldsymbol{x}) + \boldsymbol{A}(\boldsymbol{x}) \cdot \boldsymbol{p}] - \frac{e\hbar}{2mc}\boldsymbol{\sigma} \cdot \boldsymbol{\nabla}_x \times \boldsymbol{A}(\boldsymbol{x}), \tag{10.40}$$

where $\boldsymbol{x} \equiv \boldsymbol{r_e} - \boldsymbol{r}$ and $\boldsymbol{p}$ operates on $\boldsymbol{r_e}$. In the Born approximation we consider the matrix element of this interaction between the initial and final states. For the neutron these are just $\exp(i\boldsymbol{k} \cdot \boldsymbol{r})$ and $\exp(i\boldsymbol{k'} \cdot \boldsymbol{r})$, respectively. Therefore, if we integrate over the neutron coordinates, we have

$$\begin{aligned} \langle \boldsymbol{k'} | \, V(\boldsymbol{x}) | \boldsymbol{k} \rangle = &- \frac{e}{2mc} \int d\boldsymbol{r} e^{-i\boldsymbol{k'} \cdot \boldsymbol{r}} [\boldsymbol{p} \cdot \boldsymbol{A}(\boldsymbol{x}) + \boldsymbol{A}(\boldsymbol{x}) \cdot \boldsymbol{p}] e^{i\boldsymbol{k} \cdot \boldsymbol{r}} \\ &- \frac{e\hbar}{2mc}\boldsymbol{\sigma} \cdot \int d\boldsymbol{r} e^{i\boldsymbol{k'} \cdot \boldsymbol{r}} \nabla_{\boldsymbol{x}} \times \boldsymbol{A}(\boldsymbol{x}) e^{-i\boldsymbol{k} \cdot \boldsymbol{r}} \end{aligned} \tag{10.41}$$

The Born approximation is valid in this case because the average interaction energy $\mu_e \mu_N / x^3 \simeq 10^{-3}$ K, which is very small compared with the kinetic energy of the neutron. If we write $\boldsymbol{r} = \boldsymbol{r}_e - \boldsymbol{x}$ and $\boldsymbol{\kappa} = \boldsymbol{k} - \boldsymbol{k'}$, the partial matrix element (10.41) becomes

$$\begin{aligned} -\tfrac{e}{2mc} \textstyle\int \boldsymbol{dx} \exp(-i\boldsymbol{\kappa} \cdot \boldsymbol{x}) \exp(-i\boldsymbol{\kappa} \cdot \boldsymbol{r}_e) \, [\boldsymbol{p} \cdot \boldsymbol{A}(\boldsymbol{x}) + \boldsymbol{A}(\boldsymbol{x}) \cdot \boldsymbol{p}] \\ -\tfrac{e\hbar}{2mc}\boldsymbol{\sigma} \cdot \textstyle\int d\boldsymbol{x} \exp(-i\boldsymbol{\kappa} \cdot \boldsymbol{x}) \exp(i\boldsymbol{\kappa} \cdot \boldsymbol{r}_e) \, \nabla_x \times \boldsymbol{A}(\boldsymbol{x}) \end{aligned} \tag{10.42}$$

From the explicit form of $\boldsymbol{A}(\boldsymbol{x})$ we have that $\boldsymbol{\nabla}_{\boldsymbol{x}} \cdot \boldsymbol{A}(\boldsymbol{x}) = 0$, as well as

$$\int d\boldsymbol{x} e^{i\boldsymbol{\kappa} \cdot \boldsymbol{x}} \boldsymbol{A}(\boldsymbol{x}) = \frac{4\pi i}{\kappa^2}(\boldsymbol{\mu} \times \boldsymbol{\kappa}) \tag{10.43}$$

and

$$\int d\boldsymbol{x} e^{i\boldsymbol{\kappa} \cdot \boldsymbol{x}} \, [\nabla_x \times \boldsymbol{A}(\boldsymbol{x})] = \frac{4\pi}{\kappa^2}\boldsymbol{\kappa} \times (\boldsymbol{\mu} \times \boldsymbol{\kappa}) = 4\pi[\boldsymbol{\mu} - (\hat{\boldsymbol{\kappa}} \cdot \boldsymbol{\mu})\hat{\boldsymbol{\kappa}}]. \tag{10.44}$$

Therefore (10.41) reduces to

$$i\frac{4\pi e}{mc}\frac{1}{\kappa^2}(\boldsymbol{\mu} \times \boldsymbol{\kappa}) \cdot \exp(i\boldsymbol{\kappa} \cdot \boldsymbol{r}_e) \, \boldsymbol{p} - \frac{4\pi e\hbar}{2mc}[\boldsymbol{\sigma} \cdot \boldsymbol{\mu} - (\boldsymbol{\sigma} \cdot \hat{\boldsymbol{\kappa}})(\boldsymbol{\mu} \cdot \hat{\boldsymbol{\kappa}})] \exp(i\boldsymbol{\kappa} \cdot \boldsymbol{r}_e) \tag{10.45}$$

The first term represents the interaction between the neutron's moment and the orbital current, and the second term is the spin-spin interaction.

It is interesting to consider what happens when $\boldsymbol{\kappa} \to 0$, that is, the case of forward scattering. We rewrite the first term as

$$i\frac{4\pi e}{mc}\boldsymbol{\mu} \cdot \left[\frac{\exp(i\boldsymbol{\kappa} \cdot \boldsymbol{r}_e)}{\kappa^2}\boldsymbol{\kappa} \times \boldsymbol{p}\right] \tag{10.46}$$

Expanding the exponential, we have

$$i\frac{4\pi e}{mc}\boldsymbol{\mu} \cdot \frac{\boldsymbol{\kappa} \times \boldsymbol{p}}{\kappa^2} - \frac{4\pi e}{mc}\boldsymbol{\mu} \cdot \left(\frac{\boldsymbol{\kappa} \cdot \boldsymbol{r}_e}{\kappa^2}\boldsymbol{\kappa} \times \boldsymbol{p}\right) + \ldots \tag{10.47}$$

We take the mean value of the limit as $\boldsymbol{\kappa} \to 0$. In this limit the first term of (10.47) vanishes and the second term becomes

$$-\frac{4\pi e}{mc}\boldsymbol{\mu}\cdot\frac{1}{3}\left(\boldsymbol{r}_e\times\boldsymbol{p}\right)=\frac{8\pi}{3}\mu_B\boldsymbol{\mu}\cdot\boldsymbol{l} \tag{10.48}$$

The spin term of (10.45) becomes

$$4\pi\mu_B\left[\boldsymbol{\sigma}\cdot\boldsymbol{\mu}-\frac{1}{3}\boldsymbol{\sigma}\cdot\boldsymbol{\mu}\right]=\frac{8\pi}{3}\mu_B\boldsymbol{\mu}\cdot\boldsymbol{\sigma}. \tag{10.49}$$

Thus the partial matrix element is

$$\frac{8\pi}{3}\mu_B\boldsymbol{\mu}\cdot(\boldsymbol{l}+2s). \tag{10.50}$$

This shows that the forward-scattering cross section is proportional to the total magnetic moment of the electron.

Let us return now to the case $\boldsymbol{\kappa}\neq 0$. In many materials the orbital momentum is quenched. Therefore let us consider only the spin term. The matrix element entering (10.1), which is essentially the magnetic scattering amplitude, is

$$4\pi\mu_B\left\langle\alpha',m_s'\right|\sum_i\left[\boldsymbol{\sigma}_i\cdot\boldsymbol{\mu}-\left(\boldsymbol{\sigma}_i\cdot\hat{\boldsymbol{\kappa}}\right)\left(\boldsymbol{\mu}\cdot\hat{\boldsymbol{\kappa}}\right)\right]\exp\left(i\boldsymbol{\kappa}\cdot\boldsymbol{r}_i\right)\left|\alpha,m_s\right\rangle, \tag{10.51}$$

where the sum over i includes all the electrons in the sample. The factor $\boldsymbol{\sigma}_i\exp\left(i\boldsymbol{\kappa}\cdot\boldsymbol{r}_i\right)$ may be written as $\int d\boldsymbol{r}\boldsymbol{\sigma}_i\delta\left(\boldsymbol{r}-\boldsymbol{r}_i\right)\exp(i\boldsymbol{\kappa}\cdot\boldsymbol{r})$. By analogy with (1.48),

$$\sum_i\boldsymbol{\sigma}_i\delta\left(\boldsymbol{r}-\boldsymbol{r}_i\right)$$

is the spin density at $\boldsymbol{r}$, which we shall write as $2\boldsymbol{S}(\boldsymbol{r})$. Therefore, the total scattering amplitude is proporitional to the Fourier transform of the spin density. The neutron's magnetic moment is related to its spin by

$$\boldsymbol{\mu}=-1.91\frac{e\hbar}{m_0c}s$$

and the scattering amplitude becomes

$$-1.91\frac{4\pi e^2\hbar^2}{mm_0c^2}\left\langle a',m_s'\right|\boldsymbol{S}(-\boldsymbol{\kappa})\cdot\boldsymbol{s}-[\boldsymbol{s}\cdot\hat{\boldsymbol{\kappa}}][\boldsymbol{S}(-\boldsymbol{\kappa})\cdot\hat{\boldsymbol{\kappa}}]\left|a,m_s\right\rangle. \tag{10.52}$$

The cross section is given by

$$\begin{aligned}\frac{d^2\sigma}{d\Omega' dE'}=&4\left(\frac{1.91e^2}{mc^2}\right)^2\sum_{\alpha,m_s}\sum_{\alpha',m_s'}P_\alpha P_{m_s}\frac{k'}{k}\\&\times\left|\left\langle\alpha',m_s'\right|\boldsymbol{s}\cdot\boldsymbol{S}(\boldsymbol{\kappa})-[\boldsymbol{s}\cdot\hat{\boldsymbol{\kappa}}][\boldsymbol{S}(\boldsymbol{\kappa})\cdot\hat{\boldsymbol{\kappa}}]\right|\alpha,m_s\rangle\left|^2\delta,(\text{ energy }\right)\end{aligned} \tag{10.53}$$

Since the energy of the system does not depend on the neutron polarization, we may sum over m_s' to obtain

$$\begin{aligned}\frac{d^2\sigma}{d\Omega' dE'}=&4\left(\frac{1.91e^2}{mc^2}\right)^2\sum_{\alpha,\alpha'}P_\alpha\frac{k'}{k}\\&\times\sum_{\mu,\eta}\langle\alpha|S_\nu(\boldsymbol{\kappa})-\hat{\boldsymbol{\kappa}}_\nu[\boldsymbol{S}(\boldsymbol{\kappa})\cdot\hat{\boldsymbol{\kappa}}]\left|\alpha'\right\rangle\left\langle\alpha'\right|S_\mu(-\boldsymbol{\kappa})-\hat{\boldsymbol{\kappa}}_\mu[\boldsymbol{S}(-\boldsymbol{\kappa})\cdot\hat{\boldsymbol{\kappa}}]|\alpha\rangle\\&\times\sum_{m_s}P_{m_s}\left\langle m_s\right|s_\nu s_\mu\left|m_s\right\rangle\delta\text{ (energy)}\end{aligned} \tag{10.54}$$

Since

$$\sum_{m_s} P_{m_s} \langle m_s | \, s_\nu s_\mu \, | m_s \rangle = \frac{1}{4} \delta_{\nu\mu}$$

this becomes

$$\begin{aligned} \frac{d^2\sigma}{d\Omega' dE'} = & \left(\frac{1.91e^2}{mc^2} \right) \sum_{\mu,\nu} \left(\delta_{\nu\mu} - \hat{\kappa}_\nu \hat{\kappa}_\mu \right) \\ & \times \sum_{\alpha,\alpha'} P_\alpha \frac{k'}{k} \langle \alpha | S_\nu(\boldsymbol{\kappa}) \, | a' \rangle \, \langle \alpha' | \, S_\mu(-\boldsymbol{\kappa}) | \alpha \rangle \delta(\text{ energy }) \end{aligned} \tag{10.55}$$

Finally, we may incorporate the energy conservation into the matrix elements themselves by writing the delta function as

$$\delta(\text{ energy }) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dt \exp\left[i \left(\omega - E_{\alpha'} + E_\alpha \right) t \right] \tag{10.56}$$

where $\omega \equiv \hbar^2 \left(k^2 - k'^2 \right) / 2m_0$. Then

$$\begin{aligned} \frac{d^2\sigma}{d\Omega' dE'} = & \left(\frac{1.91e^2}{mc^2} \right)^2 \sum_{\mu,\nu} \left(\delta_{\nu\mu} - \hat{\kappa}_\nu \kappa_\mu \right) \frac{k'}{k} \frac{1}{2\pi} \int dt e^{i\omega t} \\ & \times \sum_{\alpha,\alpha'} P_\alpha \langle \alpha | \exp(i\mathcal{H}t) S_\nu(\boldsymbol{\kappa}) \exp(-i\mathcal{H}t) \, | \alpha' \rangle \, \langle \alpha' | \, S_\mu(-\boldsymbol{\kappa}) | \alpha \rangle \end{aligned} \tag{10.57}$$

The sum over a' may now be removed by closure. Furthermore, the sum over α is just what we mean by the thermal average. Therefore we have

$$\frac{d^2\sigma}{d\Omega' dE'} = \left(\frac{1.91e^2}{mc^2} \right)^2 \frac{k'}{k} \sum_{\mu,\nu} \left(\delta_{\nu\mu} - \hat{\kappa}_\nu \hat{\kappa}_\mu \right) \frac{1}{2\pi} \int dt e^{i\omega t} \left\langle S_\nu(\boldsymbol{\kappa}, t) S_\mu(-\boldsymbol{\kappa}) \right\rangle \tag{10.58}$$

It is convenient to separate the cross section into a Bragg part and a diffuse part:

$$\left(\frac{d^2\sigma}{d\Omega' dE'} \right)_{\text{Bragg}} = \left(\frac{1.91e^2}{mc^2} \right)^2 \sum_{\mu,\nu} \left(\delta_{\nu\mu} - \hat{\kappa}_\nu \hat{\kappa}_\mu \right) \delta(\omega) \left\langle S_\nu(\boldsymbol{\kappa}) \right\rangle \left\langle S_\mu(-\boldsymbol{\kappa}) \right\rangle \tag{10.59}$$

$$\begin{aligned} \left(\frac{d^2\sigma}{d\Omega' dE'} \right)_{\text{diffuse}} = & \left(\frac{1.91e^2}{mc^2} \right)^2 \frac{k'}{k} \sum_{\mu,\nu} \left(\delta_{\nu\mu} - \hat{\kappa}_\nu \hat{\kappa}_\mu \right) \frac{1}{2\pi} \int dt e^{i\omega t} \\ & \times \left[\left\langle S_\nu(\boldsymbol{\kappa}, t) S_\mu(-\boldsymbol{\kappa}) \right\rangle - \left\langle S_\nu(\boldsymbol{\kappa}) \right\rangle \left\langle S_\mu(-\boldsymbol{\kappa}) \right\rangle \right] \end{aligned} \tag{10.60}$$

The operators $S_\nu(\boldsymbol{\kappa})$ entering the cross section are the Fourier transform of the spin density of the entire sample:

$$\boldsymbol{S}(\boldsymbol{\kappa}) = \sum_i \exp\left(i\boldsymbol{\kappa} \cdot \boldsymbol{R}_i \right) \boldsymbol{s}_i \tag{10.61}$$

The unpaired electrons on a given site will form a total spin $\boldsymbol{S}_n$ according to Hund's rule. The Wigner-Eckart theorem tells us that a matrix element of the form $\langle \alpha' | \, \boldsymbol{s}_i | \alpha \rangle$ is proportional to $\langle \alpha' | \, \boldsymbol{S}_n | \alpha \rangle$. Thus

$$\langle a'|\, \boldsymbol{S}(\boldsymbol{\kappa})|a\rangle = \sum_{\text{site } n} e^{-i\boldsymbol{\kappa}\cdot\boldsymbol{n}} f(\boldsymbol{\kappa})\, \langle \alpha'|\, \boldsymbol{S}_n|\alpha\rangle. \tag{10.62}$$

The coefficient of proportionality $f(\boldsymbol{\kappa})$ is the Fourier transform of the normalized spin density associated with the nth site and is referred to as the magnetic form factor.

In the paramagnetic state the magnetic Bragg contribution vanishes, leaving only the diffuse scattering:

$$\left(\frac{d^2\sigma}{d\Omega' dE'}\right)_{\text{diffuse}} = \left(\frac{1.91e^2}{mc^2}\right)^2 |f(\boldsymbol{\kappa})|^2 \frac{k}{k'} \sum_{\mu,\nu} (\delta_{\mu\nu} - \hat{\kappa}_\mu \hat{\kappa}_\nu) \tag{10.63}$$
$$\sum_{n,m} \exp[i\boldsymbol{\kappa}\cdot(\boldsymbol{n}-\boldsymbol{m})] \frac{1}{2\pi} \int dt \exp(i\omega t)\, \langle S_{n,\nu}(t) S_{m,\mu}\rangle$$

In a paramagnet without exchange, we can integrate over the energy with the result that the diffuse peak becomes an elastic peak with a cross section

$$\left(\frac{d\sigma}{d\Omega'}\right)_{\text{diffuse}} = N \left(\frac{1.91e^2}{mc^2}\right)^2 |f(\boldsymbol{\kappa})|^2 \frac{2}{3} S(S+1). \tag{10.64}$$

The angular dependence of this cross section enables us to determine the magnetic form factor.

Let us now consider the cross sections in ordered magnetic systems.

10.11.4 10.3.1 Bragg Scattering

Let us assume that in the ordered state, the moments are aligned along the z-axis,

$$\langle S_{m,\nu}\rangle = \langle S_z\rangle\, \delta_{\nu z}$$

The cross section then becomes

$$\frac{d\sigma}{d\Omega'} = \left(\frac{1.91e^2}{mc^2}\right)^2 |f(\boldsymbol{\kappa})|^2 \left(1 - \hat{\kappa}_z^2\right) \langle S^z\rangle^2 \sum_{n,m} \exp[i\boldsymbol{\kappa}\cdot(\boldsymbol{n}-\boldsymbol{m})] \tag{10.65}$$

Or, from (10.17),

$$\left(\frac{d\sigma}{d\Omega'}\right)_{\text{Bragg}} = \frac{(2\pi)^3 N}{V_0} \left(\frac{1.91e^2}{mc^2}\right)^2 \langle S^z\rangle^2 \sum_G |f(\boldsymbol{G})|^2 \left(1 - G_z^2\right) \delta(\boldsymbol{\kappa} - \boldsymbol{G}), \tag{10.66}$$

where $\boldsymbol{G}$ is a vector in the magnetic reciprocal lattice. Thus we see that this cross section consists of peaks at positions determined by the magnetic structure. Furthermore, the amplitudes of these peaks are proportional to the magnetic form factor.

As an example of Bragg scattering let us consider the spin arrangements in MnF_2. Chemically, this crystal has the body-centered tetragonal lattice shown in Fig. 2.2. The unit cell has six ions at the following locations:

Mn:

$$(0,0,0) \quad \left(\frac{a}{2}, \frac{a}{2}, \frac{c}{2}\right)$$

F :

$$\begin{matrix} (au, au, 0) & (a-au, a-au, 0) \\ \left(\frac{a}{2}+au, \frac{a}{2}-au, \frac{c}{2}\right) & \left(\frac{a}{2}-au, \frac{a}{2}+au, \frac{c}{2}\right) \end{matrix}.$$

The reciprocal lattice vectors are given by

$$\boldsymbol{G} = 2\pi\left(\frac{l}{a}, \frac{m}{a}, \frac{n}{c}\right), \tag{10.67}$$

where l, m, and n are integers, and $u = 0.31$. The unit cell structure factor for nuclear scattering is

$$\begin{aligned} \mathrm{F}(\boldsymbol{G}) &= \sum_{i=1}^{r} \langle b_i \rangle \exp\left(i\boldsymbol{G}\cdot\boldsymbol{\rho}_i\right) \\ &= \begin{cases} 2\langle b\rangle_{\mathrm{Mn}} + 4\langle b\rangle_{\mathrm{F}} \cos(2\pi ul)\cos(2\pi um) & l+m+n = \text{even} \\ -4\langle b\rangle_{\mathrm{F}} \sin(2\pi ul)\sin(2\pi um) & l+m+n = \text{odd} \end{cases} \end{aligned} \tag{10.68}$$

The experimental scattering intensity at a temperature above the magnetic ordering temperature is shown in Fig. 10.4b. Notice that there are no peaks corresponding to [100] and [001], in accordance with (10.69). Figure 10.4a shows that the scattering intensity in the magnetically ordered state. We

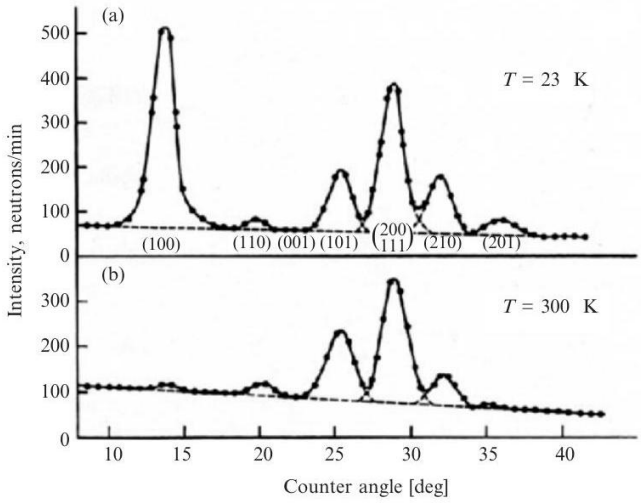


Fig. 10.4. Neutron diffraction pattern for $\mathbf{M n F _{2}$ in (a) the antiferromagnetic state (23 K) and (b) the paramagnetic state (300 K) [244]}

immediately notice the presence of additional peaks. These are magnetic peaks. Since they can be indexed in the same way as the nuclear peaks, the magnetic unit cell is the same as the chemical unit cell. By their mere existence these peaks tell us that the spin density is different from the electronic charge density; this is, the spin arrangement is not ferromagnetic. Since there are two magnetic ions in the unit cell, we try an antiferromagnetic arrangement. This introduces a sign change into the magnetic form factor. If the up spins are on the corners of the unit cell and the down spins are at the centers, the magnetic structure factor is

$$F_M(\boldsymbol{G}) = \sum_i (\pm)_i \exp\left(i\boldsymbol{G}\cdot\boldsymbol{\rho}_i\right) = 1 - e^{i\pi(l+m+n)} = \begin{cases} 0 & l+m+n = \text{ even} \\ 2 & l+m+n = \text{ odd} \end{cases} \tag{10.69}$$

We see that a peak is expected at [100], as observed in Fig. 10.4a. However, we also expect one at [001]. The fact that this is not observed indicates that the spins are oriented in this direction since the factor $1 - \hat{\kappa}_z^2$ then vanishes.

Izuyama et al. [245] have shown that in a metal the Bragg scattering cross section is essentially the same as that give by (10.66), except that $\langle S^z \rangle^2$ is replaced by $(N_\uparrow - N_\downarrow)^2$, where N_σ is the total number of electrons with spin σ.

Neutron scattering is also particularly useful in studying structures which are not commensurate with the crystal lattice. The rare-earth metals listed in Fig. 4.3, for example, exhibit a variety of incommensurate structures. These were all determined by neutron scattering.

An interesting study of incommensurate structures is found in MnP. This orthorhombic material is ferromagnetic below 291 K , with an additional magnetic phase transition at 50 K . The lower temperature phase is characterized by a spiral spin structure with a wave vector along the α-axis. Application of a magnetic field along the b-axis produces a fan structure with a periodicity governed by $\boldsymbol{q} = 2\pi(\delta/a)\boldsymbol{A}$. Figure 10.5 shows the phase diagram of MnP [246].

The point at which the three phases meet was recognized as a possible Lifshitz point. This particular multicritical point is characterized by three features: (1) it is the meeting point of three phases: a paramagnetic phase, an ordered phase whose order parameter has a fixed wave vector (zero in the ferromagnetic case), and an ordered phase (the fan phase) whose wave vector varies continuously with the thermodynamic variables; (2) the transitions from the paramagnetic phase to both ordered phases must be second order; and (3) the wave vector in the fan phase must approach zero continuously as the Lifshitz point is approached.

Neutron scattering was essential in verifying the last condition. Figure 10.6 shows a triple-axis scan through the (200) position of MnP, indicating the magnetic satellites due to the fan structure. The smooth decrease of δ to zero as shown in the insert in Fig. 10.5 confirms that this is indeed a Lifshitz point [247].

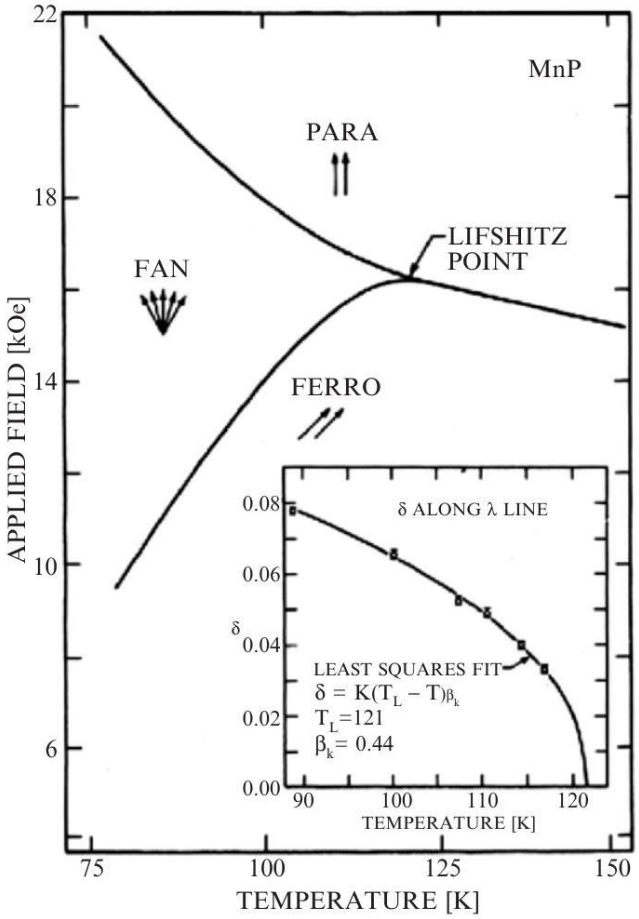


Fig. 10.5. Phase diagram of MnP near the Lifshitz point. The magnetic field is parallel to the *b*-axis. The insert shows the temperature dependence of the fan wave vector along the fan-para phase boundary

From the amplitudes of the Bragg peaks we obtain the magnetic form factor $f(\boldsymbol{\kappa})$. The experimental values for $f(\boldsymbol{\kappa})$ for metallic iron are shown in Fig. 10.7. The arrows indicate pairs of reflections for which the scattering angle is the same for the pair. Since the experimental errors are indicated by the size of the circles, these pairs tell us that the form factor, and hence the spin density, is not spherically symmetric. The curves are the theoretical results of a Hartree-Fock calculation which includes exchange effects by Watson and Freeman. The deviations from these spherical calculations are also apparent. By Fourier transforming such data it is possible to construct a spin-density "contour map" of the unit cell. This is illustrated for iron in Fig. 10.8.

Thus we find that from the Bragg scattering we can obtain information about the static magnetic state of the crystal-the spin configuration and the spin-density distribution as well as temperature and field dependence of the magnetic moment. Let us now turn to the diffuse-scattering contribution. We shall find that this contains information about the dynamic properties of the system-spin waves and critical fluctuations.

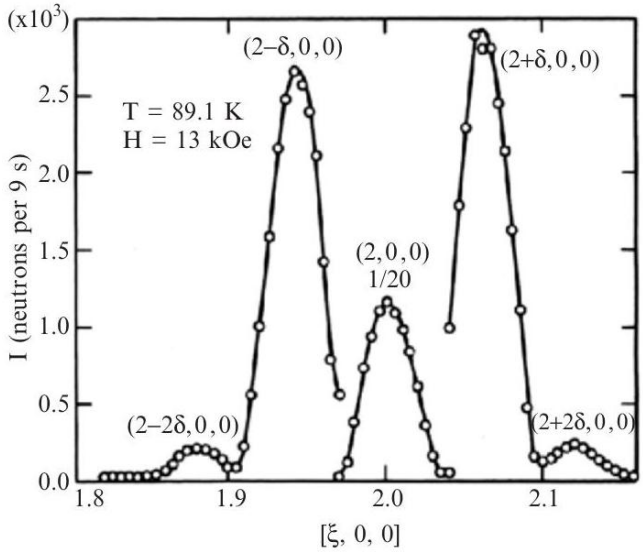


Fig. 10.6. An elastic triple-axis scan through the (200) position of MnP showing first- and second-order magnetic satellites. Note that the central (200) peak is shown on a reduced scale

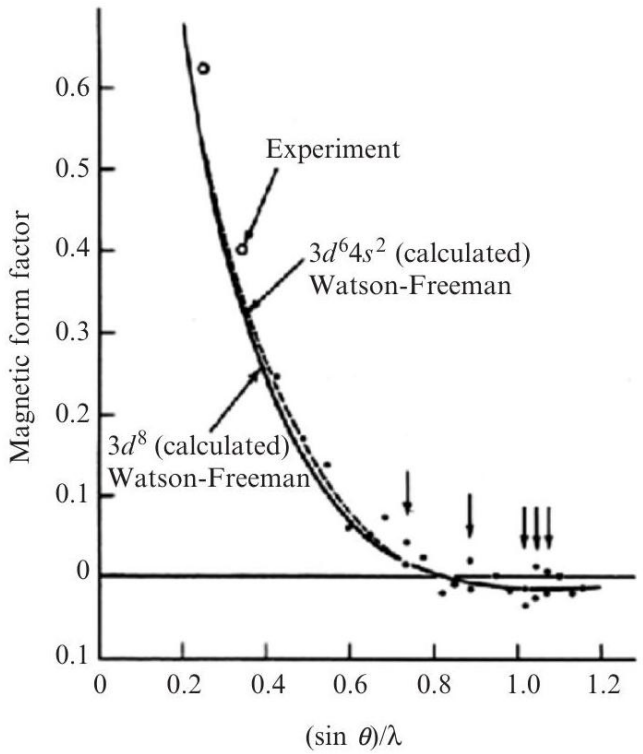


Fig. 10.7. Comparison between experimental and calculated values of the magnetic form factor for iron. The solid curve is that calculated for the spherically symmetric free atom [248]

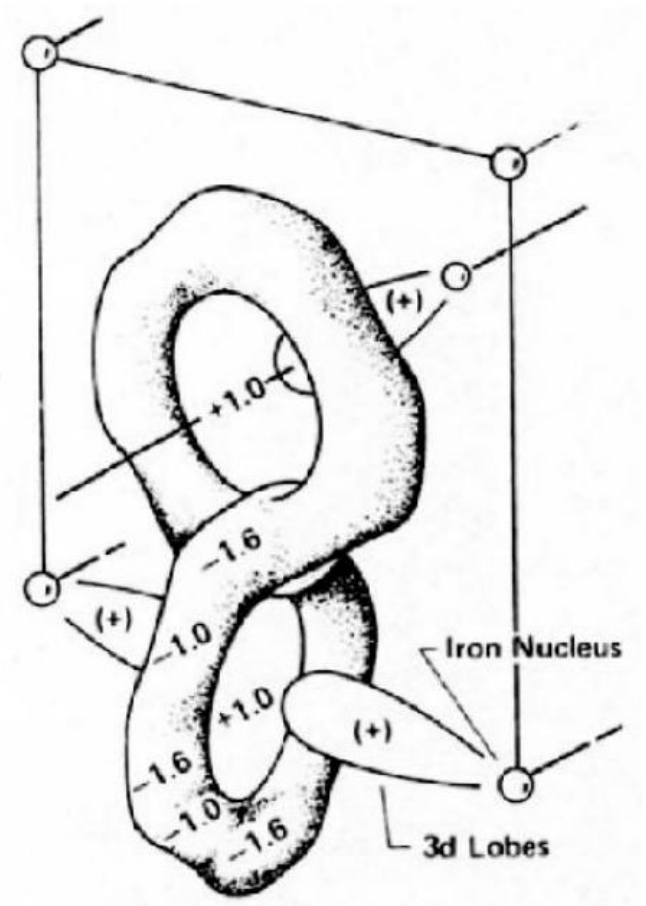


Fig. 10.8. Model of the magnetization distribution in the body-centered-cubic unit cell of iron. The concentration of the very large positive $3d$-shell magnetization along the cube edges is overemphasized to show its relationship to the negative magnetization ring [249]

10.3.2 Spin Dynamics

Let us consider the diffuse-scattering cross section (10.60). The sum over ν and μ enables us to replace $\langle S_\nu(\boldsymbol{\kappa},t)S_\mu(-\boldsymbol{\kappa})\rangle$ by the correlation function $\langle\{S_\nu(\boldsymbol{\kappa},t)S_\mu(-\boldsymbol{\kappa})\}\rangle$. In deriving the fluctuation-dissipation theorem, (1.87), we neglected long-range order. When this is included, the correlation function in (1.87) is identical to the quantity in brackets in (10.60). Therefore,

$$\left(\frac{d^2\sigma}{d\Omega' dE'}\right)_{\text{diffuse}} = \left(\frac{1.91e^2}{mc^2}\right)^2 \frac{k'}{k} \sum_{\mu,\nu} (\delta_{\nu\mu} - \hat{\kappa}_\nu \hat{\kappa}_\mu) \times \left(1/\pi g^2 \mu_B^2\right) \hbar V \coth(\beta\hbar\omega/2) \chi''_{\nu\mu}(\boldsymbol{\kappa},\omega)_s \tag{10.70}$$

The importance of the generalized susceptibility has been emphasized throughout this text. Now we see that the diffuse scattering of neutrons provides us with a direct measure of this quantity. This is why neutron scattering plays such an important role in magnetism.

The susceptibility appearing in (10.70) is the susceptibility associated with the total spin density of the system. In deriving explicit expressions from this quantity we have assumed throughout that the spin density was either localized or completely itinerant. This gave us what we might call a reduced susceptibility. In order to apply these results, we must relate this reduced
susceptibility to the total susceptibility. From (10.62) it readily follows that for localized spins

$$\chi(\boldsymbol{q},\omega)_{\text{total}} = |f(\boldsymbol{q})|^2 \chi(\boldsymbol{q},\omega)_{\text{reduced}}\,, \tag{10.71}$$

where $f(\boldsymbol{q})$ is the atomic magnetic form factor.
The corresponding relation for an itinerant system is essentially the same. To see this, let us second quantize the magnetic-moment operator in terms of Bloch states. The field operator (1.115) then becomes

$$\psi(\boldsymbol{r}) = \frac{1}{\sqrt{V}} \sum_{\boldsymbol{k},\sigma} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} u_{\boldsymbol{k}}(\boldsymbol{r}) \eta_\sigma c_{\boldsymbol{k}\sigma} \tag{10.72}$$

The Bloch function may be expanded in terms of the Wannier functins introduced in (2.78) as

$$e^{i\boldsymbol{k}\cdot\boldsymbol{r}} u_{\boldsymbol{k}}(\boldsymbol{r}) = \sum_{\alpha} e^{i\boldsymbol{k}\boldsymbol{R}_\alpha} \phi\left(\boldsymbol{r} - \boldsymbol{R}_\alpha\right). \tag{10.73}$$

In terms of these functions the field operator may be written

$$\psi(\boldsymbol{r}) = \frac{1}{\sqrt{V}} \sum_{\boldsymbol{k},\sigma} \sum_{\alpha} e^{i\boldsymbol{k}\cdot\boldsymbol{R}_\alpha} \phi\left(\boldsymbol{r} - \boldsymbol{R}_\alpha\right) \eta_\sigma c_{\boldsymbol{k}\sigma}. \tag{10.74}$$

The magnetization operator is

$$\mathcal{M}(\boldsymbol{r}) = -\mu_B \psi^\dagger(\boldsymbol{r}) \sigma \psi(\boldsymbol{r}). \tag{10.75}$$

Using (10.74) and taking the Fourier transform we obtain, for example,

$$\mathcal{M}_+(\boldsymbol{q}) = -2\mu_B \sum_{\boldsymbol{k}} F(\boldsymbol{q},\boldsymbol{k}) c^\dagger_{\boldsymbol{k}-\boldsymbol{q},\uparrow} c_{\boldsymbol{k}\downarrow} \tag{10.76}$$

where

$$F(\boldsymbol{q},\boldsymbol{k}) \equiv \frac{N}{V} \sum_{\boldsymbol{l}} e^{-\boldsymbol{k}\cdot\boldsymbol{l}} \int d\boldsymbol{r} e^{-i\boldsymbol{q}\cdot\boldsymbol{r}} \phi^*(\boldsymbol{r}) \phi(\boldsymbol{r}+\boldsymbol{l}) \tag{10.77}$$

Here $\boldsymbol{l}$ is the vector connecting lattice sites, that is, $\boldsymbol{l} = \boldsymbol{R}_\alpha - \boldsymbol{R}_{\alpha'}$. Since the Wannier function $\phi\left(\boldsymbol{r} - \boldsymbol{R}_\alpha\right)$ is strongly peaked around the site $\boldsymbol{R}_\alpha$, the dominant contribution to (10.72) arises from the $\boldsymbol{l} = 0$ term. The function $F(\boldsymbol{q},\boldsymbol{k})$ is then independent of $\boldsymbol{k}$. By comparing $\mathcal{M}_+(\boldsymbol{q})$ in this case to the free-electron expression, we again obtain the relation (10.71).

Spin-Wave Scattering. There are several applications of diffuse scattering which are particularly interesting. One is the very low temperature region where one has inelastic spin-wave scattering. As we saw in Chap. 8, the susceptibilities of both localized spin systems and metals contain poles which correspond to spin-wave excitations. Therefore at low temperatures the differential diffuse-scattering cross section is proportional to $|f(\boldsymbol{q})|^2 \delta(\omega - \omega\boldsymbol{q})$.

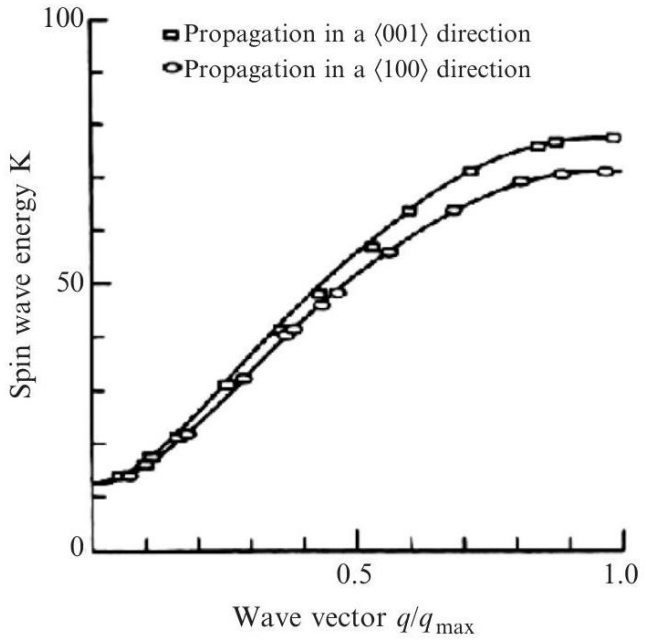


Fig. 10.9. Antiferromagnetic spin-wave dispersion in $\mathrm{MnF_2}$ at 4.2 K observed by neutron inelastic scattering. The solid lines are obtained from spin-wave theory, with the parameters $J_1 = 0.32$ K (ferromagnetic), $J_2 = -1.76$ K (antiferromagnetic), $J_3 = 0$ K, and an anisotropy field $H_A = 1.06$ K (at zero wave vector) [250]}

By measuring the energy loss of those neutrons scattered at direction $\boldsymbol{q}$ with respect to the incident beam we can thus obtain the spin-wave spectrum. The spin-wave spectrum of the antiferromagnet MnF_2 obtained in this way is shown in Fig. 10.9. An example of a metallic spin-wave spectrum which also reveals the presence of phonons is that of the rate earth terbium, shown in Fig. 10.10. We saw in Sect. 7.1 that the exchange interaction between localized moments in rare-earth metals arises through the conduction electrons and therefore depends upon the band structure. Thus from a spin-wave spectrum such as that shown in Fig. 10.10 we can obtain information about the band structure.

Critical Scattering. Diffuse scattering is also interesting in the vicinity of the magnetic phase transition. At the Curie point the Bragg peak transforms to a sharply peaked diffuse intensity, and then when the temperature in increased this diffuse intensity spreads out to approach the paramagnetic limit. The temperature dependence of diffuse scattering at a fixed angle is dramatically illustrated in Fig. 10.11. This is what is referred to as critical scattering.

The origin of this critical scattering in a localized spin system is contained in our result (4.12). The quantity measured is $d\sigma/d\Omega'$, obtained by integrating (10.70) over ω. If we assume that the major frequency response occurs for frequencies below $k_B T/\hbar$, then $d\sigma/d\Omega'$ is proportional, by the Kramers-Kronig relation (1.64), to the static susceptibility $\chi(\boldsymbol{q})$. In the vicinity of a Bragg peak momentum transfer is very small. In the limit $q \to 0$ the susceptibility

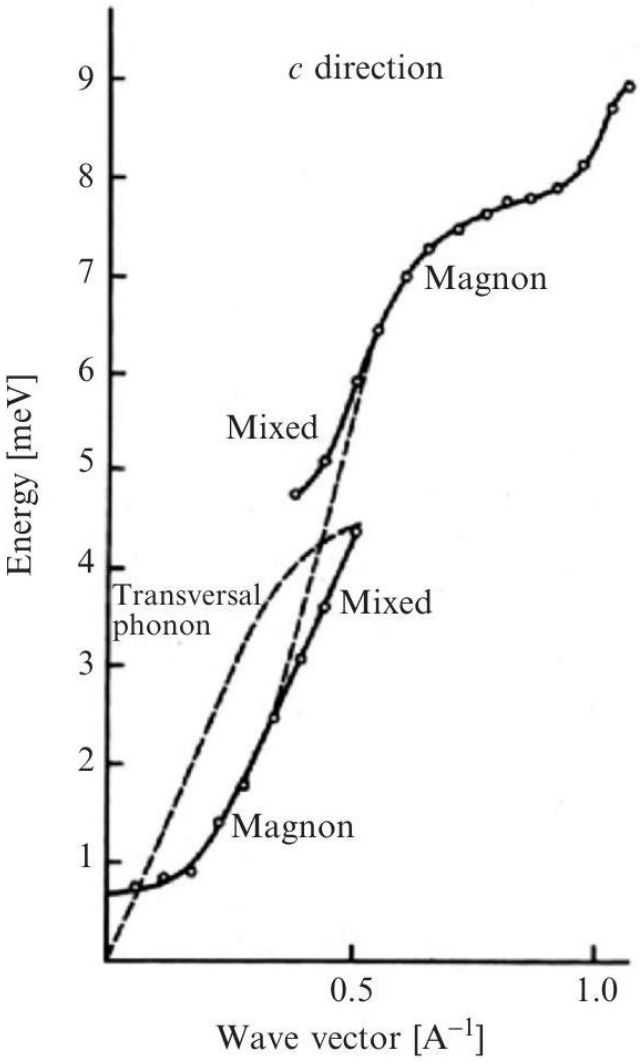


Fig. 10.10. Neutron excitation energies in the c direction of $\mathrm{Tb
-10 % \mathrm{Ho}$ at 110 K . The scan used for these measurements was such that
transverse phonons were not observed [251]}

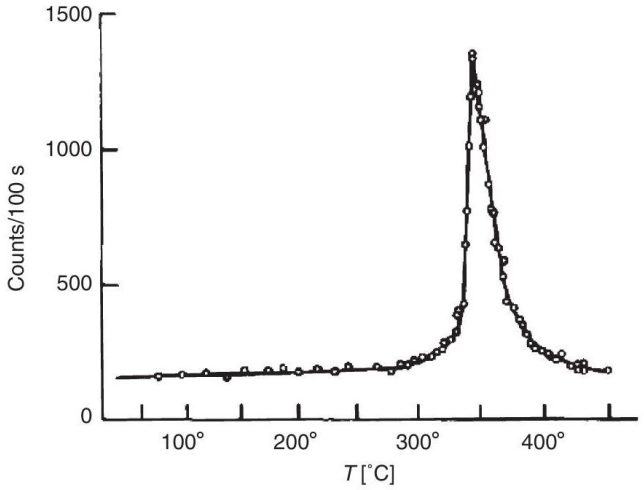


Fig. 10.11. Critical scattering by Ni. Curve gives total intensity of $4.75 - \mathring{A}$ neutrons scattered at an angle of 2° [252]

(4.12), and hence the cross section, varies as $(\kappa^2 + \boldsymbol{q}^2)^{-1}$, where κ is given by (4.49). Therefore the critical scattering should increase, in the mean field approximation, as $(T - T_c)^{-1}$. The correlation-range parameter κ^2 is found to diverge more like $(T - T_c)^{-4/3}$, which is in agreement with a more sophisticated theoretical treatment of the correlation function.

In principle, we can measure $\chi(\boldsymbol{q}, \omega)$ for every point in Fourier space and for all temperatures. An example of this is the work of Lowde and coworkers on nickel [253]. Such results enable us to study the complete evolution of the magnetic state.

Diffuse Elastic Scattering. Finally, let us mention the use of diffuse scattering to study binary alloys. Shull and Wilkinson [254] were among the first to employ neutrons in the fashion. Consider an alloy containing $N_1 = cN$ atoms of type 1 and $N_2 = (1 - c)N$ atoms of type 2 . If we assume that all atoms of type 1 have a moment μ_1, independent of their environment, while all atoms of type 2 have a moment μ_2, then the elastic diffuse cross section becomes

$$\left(\frac{d\sigma}{d\Omega'}\right)_{\text{diffuse}} = \frac{N}{4}\left(\frac{1.91e^2}{mc^2}\right)^2 \left(1 - \kappa_z^2\right) c(1-c)\left[f_1(\boldsymbol{\kappa})\mu_1 - f_2(\boldsymbol{\kappa})\mu_2\right] \tag{10.78}$$

Marshall [255] extended this formula to include environmental effects. From an analysis of the elastic diffuse scattering one can obtain the moments and their form factors. In Ni-Cu alloys, for example, Medina and Cable [256] found that the reduction in the net moment of the alloy is associated with a negative polarization of the Ni sites, and that there is no polarization of the Cu atoms.

10.4 Example: Manganese Oxides

Neutron scattering has proven particularly useful in illuminating the magnetic properties of the manganese oxides, or "manganites", $AMnO_3$ where $A = La, Ca, Ba, Sr, Pb, Nd$, or Pr. In $LaMnO_3$ the manganese is trivalent, Mn^{3+}, with a d^4 configuration. In a cubic crystal field the energy levels have the form shown on the left in Fig. 10.12. An orthorhombic distortion splits these states as shown on the right. If the resulting gain in electronic energy exceeds the elastic energy system will spontaneously distort. This is known as a coopertative Jahn-Teller distortion. $LaMnO_3$ is indeed an orthorhombic insulating antiferromagnet with a Néel temperature $T_N = 140$ K. In $CaMnO_3$ the manganese is tetravalent, $Mn4+$. $CaMnO_3$ has the cubic perovskite structure and is also an antiferromagnetic insulator.

What makes the manganites so interesting is when A consists of a mixture of trivalent and divalent ions, such as La^{3+} and Ca^{2+}. This results in the presence of both Mn^{3+} and Mn^{4+} with very interesting electrical properties. Figure 10.13, for example shows the resistivity as a function of temperture for $La_{0.67}Ca_{0.33}MnO_3$. We see that there is a metal-insulator transition as a function of temperature and a large magnetoresistance. These materials are

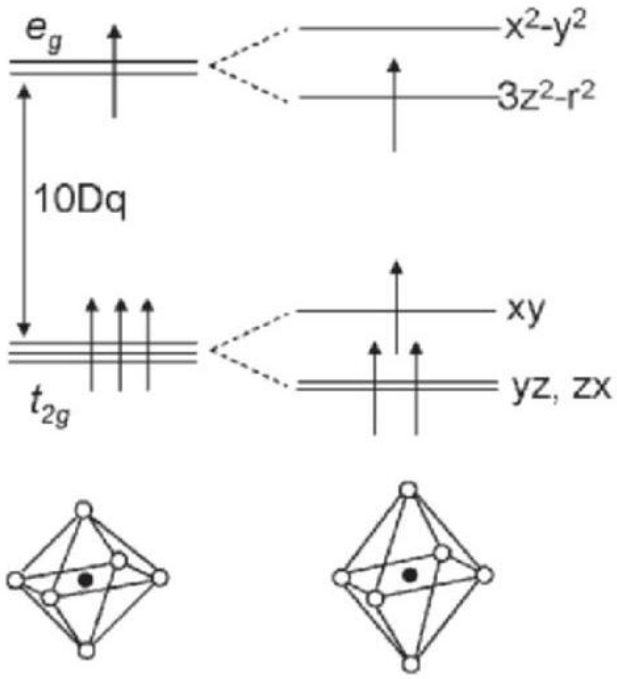


Fig. 10.12. Splitting of the $e_{g} $ and t_{zg} orbitals as the result of a Jahn-Teller distortion. (After [257])}

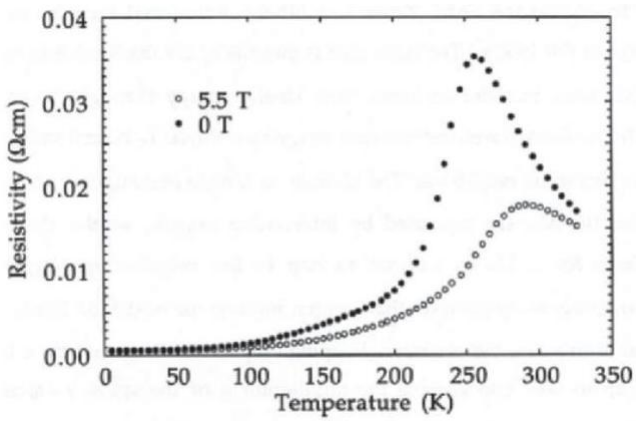


Fig. 10.13. Resistivity of $\mathrm{La_{0.67}Ca_{0.33}MnO_3}$ thin film as a function of temperature [258]}

referred to as "colossal" magnetoresistance oxides. Magnetization measurements indicate that $La_{0.67}Ca_{0.33}MnO_3$ is ferromagnetic below $T_c = 257$ K.

What makes this system so interesting and, at the same time, so challenging, is the interplay among the charge, the magnetic, and the lattice (orbital) degrees of freedom. J.W. Lynn and

his colleagues have made extensive neutron measurements using a triple-axis spectrometer on $La_{0.7}Ca_{0.3}MnO_3$ and related materials in an effort to understand their properties [259].

Figure 10.14 shows the diffuse magnetic scattering as a function of energy at a momentum transfer of $\boldsymbol{q} = (0.09, 0, 0)$ (the nuclear incoherent scattering has been subtracted out). This magnetic fluctuation spectrum shows spinwave peaks at ±2meV as well as a quasielastic peak centered at $E = 0$. What makes this peak unusual is that if one fits the intensity of this peak to a Lorentzian shape, $I \sim (q^2 + \kappa^2)^{-1}$ the correlation length $\xi \left(= \frac{1}{\kappa}\right)$ does not

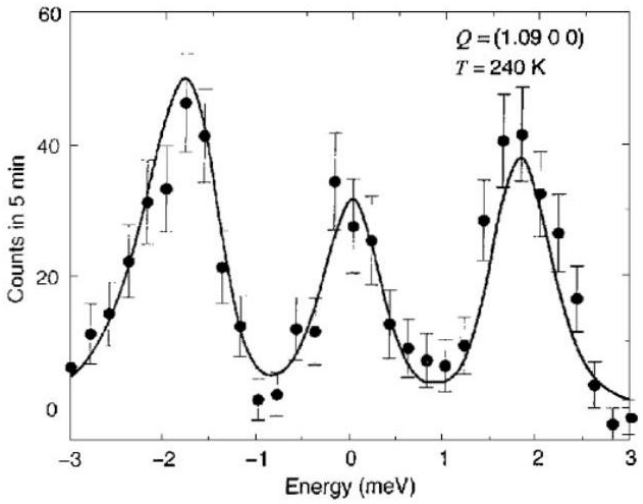


Fig. 10.14. Constant-Q scan taken at $T=240 \mathrm{˜K $, well below Curie temperature of $T_c = 257$ K, at a reduced wave vector of ($0.09, 0, 0$)}

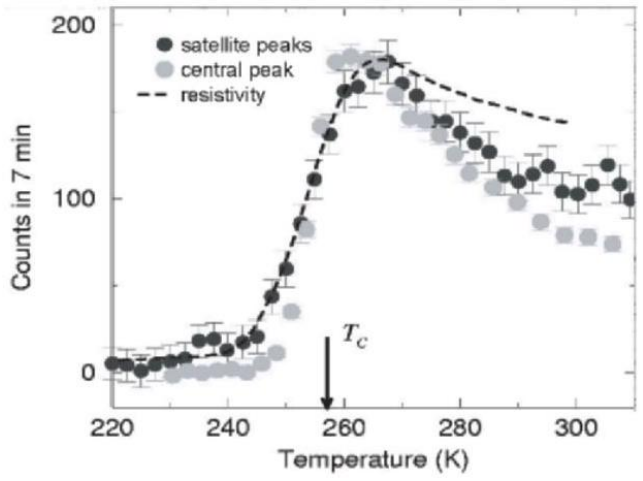


Fig. 10.15. Temperature dependence of the central peak in Fig. 10.14 and the satellite peaks compared with the resistivity. The data has been scaled so the peak heights match

diverge as one approaches T_c from above as it would for a conventional ferromagnet. Another interesting revelation from the neutron measurements is the development of nuclear peaks above T_c at the non-Bragg locations $\left(\frac{1}{4}, \frac{1}{4}, 0\right)$. The intensities of these satellite lines as well as the central peak in the magnetic spectrum have the same temperature dependence as the resistivity as shown in Fig. 10.15.

These neutron results, combined with the electrical and magnetic measurements [260] lead to the following picture. At low temperatures the material is a ferromagnetic metal. The "extra"

electron on a Mn^{3+} hops freely to a neighboring Mn^{4+}. This gives rise to the ferromagnetism through the double exchange mechanism discussed in Sects. 2.2.10. At high temperatures there is a Jahn-Teller distortion that traps the e_g electron forming a "small polaron". This gives a hopping conductivity with an activation energy of the order of

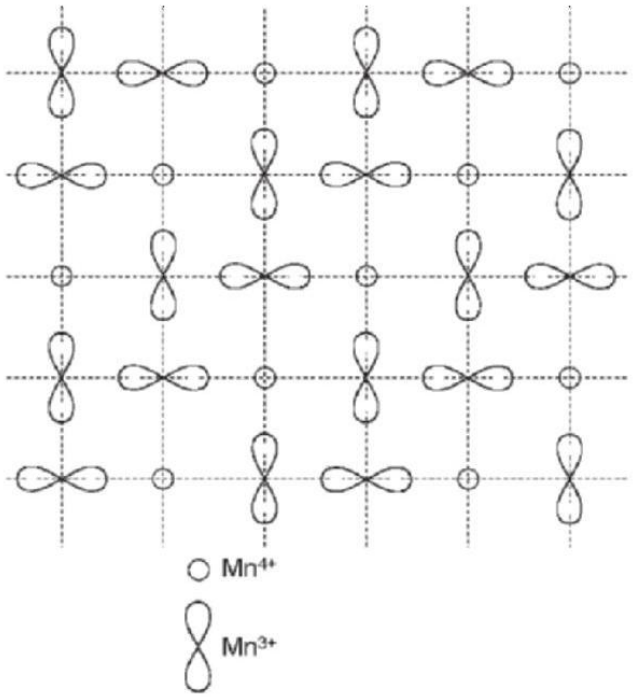


Fig. 10.16. Orbital ordering

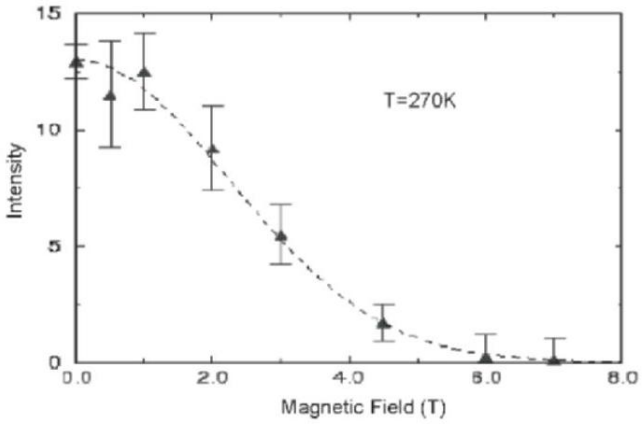


Fig. 10.17. Field dependence of the intensity of the satellite peaks associated with the polaron-polaron correlations

1000 K_B [258]. At a Ca concentration of the order of $\frac{1}{3}$ the polarons interact with each other through the Coulomb interaction which leads to ordering of both the Mn^{3+}, i.e., charge ordering, as well as ordering of the direction of the e_g orbitals, i.e., orbital ordering. It is this ordering that gives rise to the $\left(\frac{1}{4}, \frac{1}{4}, 0\right)$ peaks. An illustration of the sort of ordering that might occur is shown in Fig. 10.16. Notice that this ordering produces charge stripes running along the $(1, 1, 0)$ diagonal. The intensity of the $\left(\frac{1}{4}, \frac{1}{4}, 0\right)$ satellite peaks decreases with magnetic field as shown in Fig. 10.17. The alignment of the spins makes it easier for the e_g electrons to hop

onto neighboring Mn^{4+} sites destroying the order. This also reduces the resistivity giving rise to the colossal magnetoresistance.

10.11.5 10.5 Example: Quantum Phase Transitions

Another example where neutron scattering plays a critical role is that of identifying the mechanism responsible for magnetic field-induced transitions in certain magnetic insulators.

The magnetic phase transitions we described in Sect. 4.5 were characterized by a coherence length ξ which diverged as one approaches the transition temperature from above. As these "islands" of long range order grow the spins within them will remain ordered longer and longer. That is, their fluctuation frequency, ω_f will become smaller and smaller. This is referred to as "critical slowing down". If the transition occurs at a finite temperature, Tc , then there will always come a point at which these characteristic fluctuations $\hbar\omega_f$ become less than the thermal fluctuations, k_B T_c. In such case the critical behavior is always governed by the thermal fluctuations and we refer to these transitions as "classical". At $T = 0$, on the other hand, it is possible to envision a situation where one can induce a transition by varying some parameter of the system, such as the strength of the exchange interaction. In such a transition the zero-point, or quantum, fluctuations dominate and we refer to these as quantum phase transitions. It turns out that a vestige of such transitions extends into the finite temperature regime where they can be studied.

A good example of a quantum phase transition is found in $TlCuCl_3$. The phase diagram for this material is shown in Fig. 10.18. At low magnetic fields the material is magnetically not ordered. Some refer to this as a quantum spin liquid. Above a critical field the material has a uniform magnetization parallel to the field as well as antiferromagnetic ordering perpendicular to the field. What is happening as one increases the field at $T \simeq 0$?

The crystal structure of $TlCuCl_3$ consists of coupled two-leg ladders with rungs formed by pairs of Cu^{2+} ions. The $S = \frac{1}{2}$ spins of these ions are coupled by an antiferromagnetic exchange along the rungs leading to dimers with singlet ground states and triplet excited states. The singlet wavefunction has

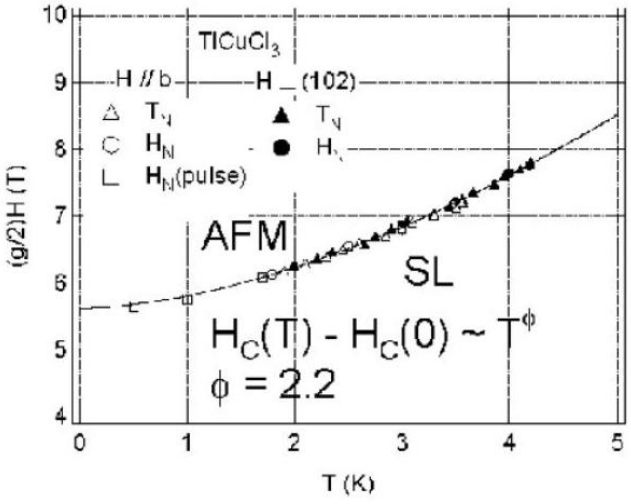


Fig. 10.18. The phase diagram of $\mathrm{TlCuCl
_{3}[261]$}

the form $(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)/\sqrt{2}$, i.e. $S = 0$. There are weaker exchange couplings along the legs of the ladders and between ladders, but the quantum fluctuations inherent in the singlet state of the dimer prevent long range magnetic order even at T = 0. The additional exchange

interactions lead to propagation of the $S = 1$ triplet states among the dimers with dispersion as shown in Fig. 10.19. When a magnetic field is applied the triplet states split as shown by the neutron data in Fig. 10.20. The lowest component at $\boldsymbol{Q} = (0, 0, 2\pi)$ "goes soft" at $H_c = 5.6$ T. When one of these triplet "magnons" is excited it creates magnetization in the transverse direction. This is different from the magnons described in Chap. 8 which decrease the magnetization. Because these magnons are bosons, and because the transverse magnetization above H_c is macroscopic this has led some to suggest that the transition at H_c is a Bose-Einstein condensation (BEC) [265, 266].

One piece of evidence that the field-induced transition is associated with Bose condensation comes from the shape of the phase boundary in Fig. 10.18. Since, in our case, we cannot have more than one triplet on a site, we have a Bose system with hard core interactions. It is this feature that limits the number of triplet bosons and makes it possible to introduce a chemical potential,

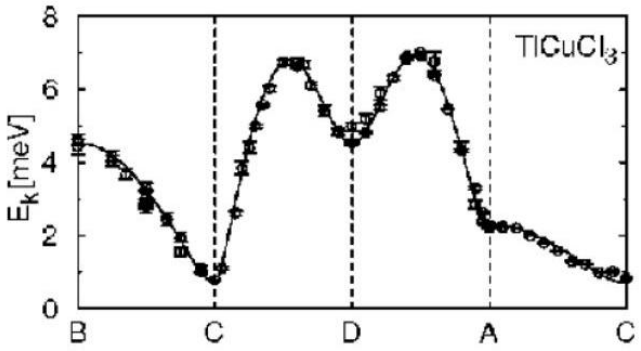


Fig. 10.19. Zero-field triplet dispersion relations for $\mathrm{TlCuCl
_{3}$ at $T = 1.5$ K[262]}

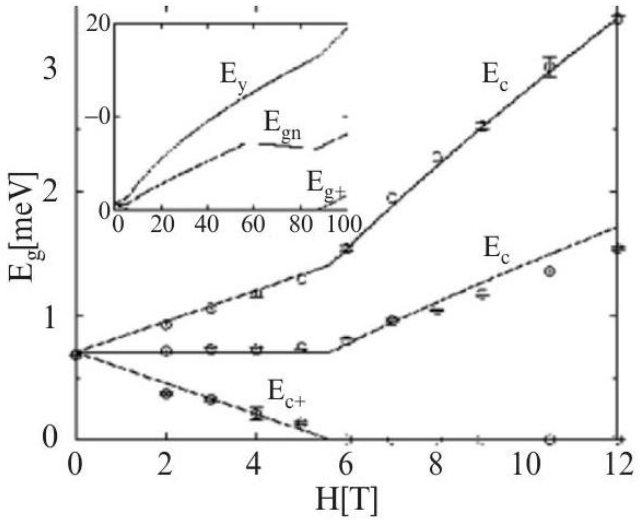


Fig. 10.20. Field dependence of the energy gap of the three magnon modes in $\mathrm{TlCuCl
_{3}$. Points are data from the neutron experiment of [267] and [269], and solid lines the results of thoery [264]. Inset: field dependence at high fields}

$\mu = g\mu_B\,(H - H_c)$. Ordinary magnons, like photons and phonons, have zero chemical potential [268]. It can be shown that the value of the chemical potential at which Bose condensation

occurs for such a system varies with temperature as $T^{\frac{3}{2}}$. This implies that $H_c(T) - H(0) \sim T^{\frac{3}{2}}$. In Fig. 10.18 we see that the boundary varies as $T^{2.2}$. A later analysis of the experimental data down to lower temperatures shows that the critical exponent is consistent with the predicted value of 3/2 [263]. In our discussion of classical phase transitions we mentioned that different systems may be characterized by the same critical exponents, defining a "universality class". Thus, proponents of the BEC model would say that $TlCuCl_3$ belongs to the Bose-Einstein condensation universality class.

Neutron scattering provides additional evidence for the existence of Bose condensation. As mentioned above, the theory of the field-induced condensate [264] predicts a transverse staggered magnetization. If the low field phase has rotational symmetry, i.e., $0(z)$, then the transition involves breaking this symmetry. As we described in Sect. 4.5, one of the consequences of a broken symmetry is the appearance of modes associated with variations in the order parameter, Goldstone modes. In the case of $TlCuCl_3$ this Goldstone mode is predicted to be a phase mode with zero energy gap and a linear wavevector dependence, i.e., a rotation of the ordered moment in the plane perpendicular to the field. Figure 10.21 shows the inelastic neutron scattering results around the Bragg point $(0, 4, 0)$. Rüegg et al. [269] suggest that $\omega \to 0$ as $k \to 0$, consistent with a Goldstone mode. However, the data does not extend below 0.2 meV . Glazkov et al. [270] have carried out magnetic resonance measurements in the range $9 - 80$GHz and found a gap of 0.08 meV for this field geometry. This indicates that the system has an anisotropy of the order of 1% of the exchange and is not completely rotationally symmetric. This would argue against a BEC. There are other dimeric copper systems

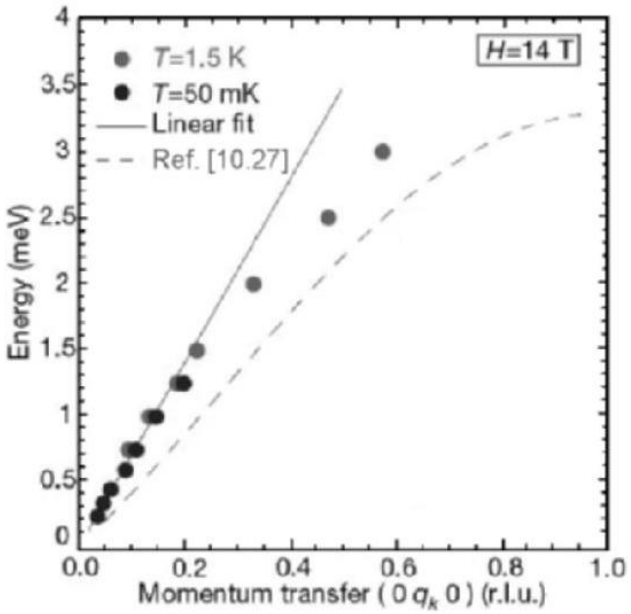


Fig. 10.21. Energy dispersion of low-lying magnetic excitations in $\mathrm{TlCuCl
_{3}$ [269]}

that exhibit signature of a magnon BEC . In particular, $BaCuSiO_6$ [273] and Cs_2CuCl_4[274] are other good candidates.

This quantum phase transition can also be induced by pressure. In this case the full 0(3) symmetry is broken with the result that one expects two degenerate Goldstone modes and a characteristic amplitude mode related to the condensate [271]. Evidence for such modes have been reported by Rüegg et al. [272] in inelastic neutron scattering performed on a triple-axis spectrometer at the Swiss spallation neutron source.

10.12 References

Chapter 1

1. J.H. Van Vleck, "The Theory of Electric and Magnetic Susceptibilities". Chaps. 1 and 4 (Oxford University Press, London 1932)
2. J.D. Jackson, "Classical Electrodynamics". 3rd edn. (John Wiley & Sons, New York 1999) Sect. 6.6
3. J.D. Jackson, "Classical Electrodynamics". 3rd edn. (John Wiley & Sons, New York 1999) Sect. 5.16
4. C. Kittel, "Thermal Physics". (John Wiley & Sons, New York 1969) Sect. 23
5. See, for example, K. Huang, "Statistical Mechanics". 2nd edn. (Wiley, New York 1987)

Chapter 2

6. L.I. Schiff: Quantum Mechanics. 3rd edn. (McGraw-Hill, New York 1968) Chap. 8
7. J.D. Bjorken, S.D. Drell: Relativistic Quantum Mechanics (McGraw-Hill, New York 1965) Chap. 4
8. A. Abragam: The Principles of Nuclear Magnetism (Oxford University Press, London 1961) p. 165
9. C.P. Slichter: Principles of Magnetic Resonance. 2nd edn. Springer Series in Solid-State Sciences, Vol. 1 (Springer, Berlin Heidelberg New York 1980) Chap. 9
10. J.S. Griffiths: The Theory of Transition-Metal Ions (Cambridge University Press, Cambridge 1961)
11. C.J. Ballhausen: Introduction to Ligand Field Theory (McGraw-Hill, New York 1962)
12. M.T. Hutchings: "Point-Charge Calculations of Energy Levels of Magnetic Ions in Crystalline Electric Fields". In Solid State Physics, Vol. 16, ed. by F. Seitz, D. Turnbull (Academic Press, New York 1964)
13. M. Tinkham: Group Theory and Quantum Mechanics (McGraw-Hill, New York 1964)
14. C. Herring: In Magnetism, Vols. IIB, IV, ed. by G. . Rado, H. Suhl (Academic Press, New York 1966)
15. W. Heitler, F. London, Z. Phys. 44, 455 (1927)
16. J.H. Van Vleck, Rev. Univ. Nac Tucuman Ser. A 14, 189 (1962)
17. J.B. Goodenough: Magnetism and the Chemical Bond (Interscience, New York 1963) Chap. 3
18. J. Kanamori, Phys. Chem. Solids 10, 87 (1959)
19. K. Ueda, H. Tabata, T. Kawai, Science 280, 1064 (1998)
20. P.W. Anderson: In Magnetism, Vol. I, ed. by G.T. Rado, H. Suhl (Academic Press, New York 1963)

21. P.J. Hay, J.C. Thibeault, R. Hoffmann, J. Am. Chem. Soc. 974884 (1975)

22. R.D. Willett, C.P. Landee, J. Appl. Phys. 52, 2004 (1981)

23. H.G. Bohn, W. Zinn, B. Dorner, A. Kollmar, J. Appl. Phys., 2228 (1981)

24. H.M. Mc Connell: Proc. Robert Welch Foundation Conference on Chemical Research, Houston, Texas (1968) p. 144

25. J.S. Miller, A.J. Epstein, Angew. Chem. Int. Ed. Engl. 33, 385 (1994)

26. C. Zener, Phys. Rev. 82, 403 (1951)

27. P.W. Anderson, H. Hasegawa, Phys. Rev. 100, 675 (1955)

28. P.-G. deGennes, Phys. Rev. 18, 141 (1960)

29. C.F. Hirjibehedin, C.P. Lutz, and A.J. Heinrich, Science 312, 1021 (2006)

30. J.W. Corbett, G.D. Watkins, Phys. Rev. Lett. 1, 314 (1961)

31. D.K. Biegelsen: Nuclear and Electron Resonoance Spectroscopies Applied to Materials Science. Materials Research Society Symposia Proceedings, Vol. 3 (Elsevier, North-Holland, Amsterdam 1981) p. 85

32. M. Tachiki, J. Phys. Soc. Jpn. 25, 686 (1968)

33. I. Dzialoshinski, Phys. and Chem. Solids 4, 241 (1958); T. Moriya, Phys. Rev. 120, 91 (1960)
34a. E.O. Kane: Semiconductos and Semimetals, Physics of III-V Compounds, Vol. 1 (Academic Press, New York 1966) pp. 75-100
34b. L.M. Roth, B. Lax, S. Zwerdling, Phys. Rev. 114, 90 (1959)

34. Y. Yafet: In Solid State Physics, Vol. 14, ed. by F. Seitz, D. Turnbull (Academic Press, New York, 1943) p. 93

Chapter 3

36. See, for example, K. Huang, "Statistical Mechanics", 2nd Edn. (Wiley, New York 1987)

37. J.H. Van Vleck: The Theory of Electric and Magnetic Susceptibilities (Oxford University Press, London 1932) p. 276

38. J.S. Griffith: The Theory of Transition-Metal Ions (Cambridge University Press, Cambridge 1961) pp. 437-439

39. M. Tinkham: Group Theory and Quantum Mechanics (McGraw-Hill, New York 1964) p. 132, Eq. (5.69)

40. L.D. Landau, Z. Phys. 64, 629 (1930) [English transl.: In Collected Papers of L.D. Landau, ed. by D. ter Haar (Gordon and Breach, New York (1965)]

41. R. Peierls, Z. Phys. 80, 763 (1933); 81, 186 (1933)

42. A.H. Wilson, Proc. Cambridge Philos. Soc. 49, 292 (1953)

43. E.H. Sondheimer, A.H. Wilson, Proc. R. Soc. London A210, 173 (1951)

44. J. Bardeen: In Handbuch der Physik, Vol. 15, Low Temperature Physics II, ed. by S. Flügge (Springer, Berlin Göttingen Heidelberg 1956) p. 305

45. H. Fukuyana, Phys. Lett. A 32, 111 (1970)

46. R. Kubo, Y. Obata, J. Phys. Soc. Jpn. 11, 547 (1956)

47. W.J. de Haas, P.M. van Alphen, Proc. Amsterdam Acad. 33, 1106 (1936)

48. J.M. Ziman:"Principles of the Theory of Solids". (Cambridge Univ. Press, Cambridge 1964) p. 275

49. K.V. Klitzing, G. Dorda, M. Pepper, Phys. Rev. Lett. 45, 494 (1981)

50. D.C. Tsui, A.C. Gossard, Appl. Phys. Lett. 38, 550 (1981)

51. R.E. Prange, Phys. Rev. B 23, 4802 (1981)

52. D.C. Tsui, S.J. Allen, Phys. Rev. B 24, 4082 (1981)

53. R.B. Laughlin, Phys. Rev. B 23, 5632 (1981)

54. D.C. Tsui, H.L. Stormer, A.C. Gossard, Phys. Rev. Lett. 48, 1559 (1982)

55. R.B. Laughlin, Phys. Rev. Lett. 50, 1395 (1983)

56. "Landolt-Börnstein Zahlenwerte und Funktionen aus Physik, Chemie, Astronomie, Geophysik und Technik". 6th edn., Vol.2, Part 9 Magnetic Properties I (Springer, Berlin Heidelberg New York 1962)

57. C.J. Krissman, H.B. Callen, Phys. Rev. 94, 837 (1954)

58. R.M. Bozorth, H.J.W. Williams, D.E. Walsh, Phys. Rev. 103, 572 (1956)

59. S. Foner, Rev. Sci. Instrum. 30, 548 (1959); S. Foner, E.J. McNiff, Jr., Rev. Sci. Instrum. 39, 171 (1968)

60. F.R. Kim, W.P. Wolf, J. Sci. Intrum. 34, 64 (1957)

61. R.T. Schumacher, C.P. Slichter, Phys. Rev. 101, 58 (1956)

62. A.M. Clogston, V. Jaccarino, Y. Yafet, Phys. Rev. A134, 1650 (1964)

63. J. Friedel, Nuovo Cimento Suppl. 2, 287 (1958)

64. A. Clogston, B. Matthias, M. Peter, H. Williams, E. Corenzwit, R. Sherwood, Phys. Rev 125, 541 (1962)

65. P.W. Anderson, Phys. Rev. 124, 41 (1961)

66. In alkali-metal hosts: Phys. Rev. Lett. 57, 388 (1986); in sp-band hosts: Phys. Rev. Lett. 61, 2129 (1988); in noble- and transition-metal hosts: Phys. Rev. Lett. 62, 316 (1989)

67. P.A. Wolff, Phys. Rev. 124, 1030 (1961)

68a. G.T. Rado, H. Suhl (Eds.): "Magnetism", Vol. V (Academic Press, New York 1973)
68b. G. Grüner, A. Zawadowski, Rep. Prog. Phys. 37, 1497 (1974); Prog. Low Temp. Phys. B 7, 591 (1978)
69. J.R. Schrieffer, P.A. Wolff, Phys. Rev. 149, 491 (1966)
70. J. Kondo, Prog. Theor. Phys. 32, 37 (1964)
71. N. Andre, Phys. Rev. Lett. 45, 379 (1980)
72. P.B. Wiegmann, J. Phys. C 14, 1463 (1981)
73. K. Yosida, A. Yoshimori: In "Magnetism", Vol. V, ed. by G.T. Rado, H. Suhl (Academic Press, New York 1973)
74. P.W. Anderson, Phys. Rev. 164, 352 (1967)
75. C. Lacroix, J. Phys. F11, 2389 (1981)

Chapter 4

76. J.S. Smart: Effective Field Theories of Magnetism (Saunders, Philadelphia 1966)

77a. L.D. Landau, E.M. Lifshitz: Statistical Physics (Pergamon, Oxford 1958) Sect. 149
77b. E. Lieb and D. Mattis, Phys. Rev. 125, 164 (1962)
77c. See, for example, R. Crook et al. Science 312, 1359 (2006)
78. R.J. Elliott (Ed.): Magnetic Properties of Rare-Earth Metals (Plenum, New York 1972)
79. A. Yoshimori, J. Phys. Soc. Jpn. 14, 807 (1959)
80. G. Low, Proc. Phys. Soc. 82, 992 (1963)
81. S. Geller, J.P. Remeika, R.C. Sherwood, H.J. Williams, G.P. Espinosa, Phys. Rev. 137, 1034 (1965)
82. S.F. Andrews, P.W. Anderson, J. Phys. F 5, 965 (1975)
83. C.A.M. Mulder, A.J. van Duyneveldt, J.A. Mydosh, Phys. Rev. B 23, 1384 (1981)
84. R.V. Chamberlin, M. Hardiman, R. Orbach, J. Appl. Phys. 52, 1771 (1981)
85. R.B. Griffiths, Phys. Rev. Lett. 24, 715 (1970)
86. K. Huang: Statistical Mechanics 2nd Edn. (Wiley, New York 1987)
87. P. Pfeuty, G. Toulouse: Introduction to the Renormalization Group and to Critical Phenomena (Wiley, New York 1977)
88. M. Fisher, Rev. Mod. Phys. 46, 597 (1974)
89. N. Mermin, H. Wagner, Phys. Rev. Lett. 17, 1133 (1966)
90. J. Kötzler, W. Scheithe, K. Knorr, W.B. Yelon, J. Phys. C 9, 1291 (1974)
91. J.C. Norvell, W.P. Wolf, L.M. Corliss, J.M. Hastings, R. Nathans, Phys. Rev. 186, 557 (1969)
92. R.J. Birgeneau, J. Skalyo, G. Shirane, J. Appl. Phys. 41, 1303 (1970)
93. P. Heller, Phys. Rev. 146, 403 (1966)
94. A. Kastner, R.J. Birgeneau, G. Shirane, Y. Endoh, Rev. Mod. Phys. 70, 897 (1998)
95. S. Chakravarty, B.I. Halperin, D.R. Nelson, Phys. Rev. Lett. 60, 1057 (1988)
96. B. Keimer, N. Belk, R.J. Birgeneau, A. Cassanho, C.Y. Chen, M. Greven, M.A. Kastner, A. Aharony, Y. Endoh, R.W. Erwin, G. Shirane, Phys. Rev. B 46, 14034 (1992)
97. V.L. Berezinskii, Sov. Phys. JETP 32, 493 (1970)
98. J.M. Kosterlitz, D.J. Thouless, J. Phys. C6, 1181 (1973)
99. E.C. Stoner, E.P. Wohlfarth: Phil Trans. R. Soc. Lond., Ser. A, 240, 599 (1948)
100. M.P. Sharrock, J. Appl. Phys. 76(10), 6413 (1994)
101. S.M. Stinnett, W.D. Doyle, P.J. Flanders, C. Dawson, IEEE Trans. Magn. 34, 1828 (1998)
102. See, for example, R. Street, S.D. Brown, J. Appl. Phys. 76, 6386 (1994)
103. S.M. Stinnett, W.D. Doyle, O. Koshkina, L. Zhang, J. Appl. Phys. 85, 5009 (1999)

Chapter 5

104. H. Kanazawa, N. Matsudaira, Prog. Theor. Phys. 23, 433 (1960)

105. P. Weiss, R. Forrer, Ann. Phys. 5, 153 (1926)

106. L.D. Landau, Sov. Phys. JETP 2, 920 (1956); 5, 101 (1957)

107. V.P. Silin, Sov. Phys. JETP 6, 945 (1958)

108. C. Herring: In Magnetism, Vol. 4, ed. by G.T. Rado, H. Suhl; (Academic Press, New York 1966)

109. P. Fulde, V. Zevin, G. Zwicknagl, Z. Phys. B 92, 133 (1993)

110. J.C. Slater, Phys. Rev. 49, 537 (1936)

111. E.C. Stoner, Proc. R. Soc. London A 165, 372 (1938)

112. P.A. Wolff, Phys. Rev. 120, 814 (1960)

113. A.W. Overhauser, Phys. Rev. 128, 1437 (1962)

114. D. Wohlleben, Phys. Rev. Lett. 21, 1343 (1968)

115. J.C. Slater, Phys. Rev. 81, 385 (1951)

116. W. Kohn, L.J. Sham, Phys. Rev. A 140, 1133 (1965)

117. T.M. Hattox, J.B. Conklin, Jr., J.C. Slater, S.B. Trickey, J. Phys. Chem. Solids 34, 1627 (1973)

118. P. Hohenberg, W. Kohn, Phys. Rev. B 136, 864 (1964)

119. P. Thiry, D. Chandesris, J. Lecante, C. Guillot, R. Pinchaux, Y. Pétroff, Phys. Rev. Lett. 43, 82 (1979)

120. F.J. Himpsel, P. Heimann, D.E. Eastman, J. Appl. Phys. 52, 1658 (1981)

121. J.C. Slater, J. Appl. Phys. 8, 385 (1937)

122. L. Pauling, Phys. Rev. 54, 899 (1938)

123. H.A. Mook, Phys. Rev. 148, 495 (1966)

124. R.M. Moon, Phys. Rev. A 136, 195 (1964)

125. J. Hubbard, J. Appl. Phys. 52, 1654 (1981)

126. M.V. You, V. Heine, A.J. Holden, P.J. Lin-Chung, Phys. Rev. Lett. 44, 1282 (1980)

127. J. Hubbard, Proc. R. Soc. London A 276, 238 (1963); Ibid., 277, 237; Ibid., 281, 401

Chapter 6

128. F. Bloch, Phys. Rev. 70, 460 (1946)

129a. F. Bloch, W.W. Hansen, M. Packard, Phys. Rev. 70, 474 (1946)
129b. N. Bloembergen, Phys. Rev. 78, 572 (1950).
129c. H.W. Schumacher, C. Chappert, R.C. Sousa, P.P. Freitas, J.M. Itat, Phys. Rev. Lett. 90, 017204 (2003)
130. J.H. Van Vleck, Phys. Rev. 74, 1168 (1948)
131. R. Kubo, K. Tomita, J. Phys. Soc. Jpn. 9, 888 (1954)
132. See G.E. Pake, "Paramagnetic Resonance". (W.A. Benjamin, New York 1962) p. 142
133. A. Abragam: The Principles of Nuclear Magnetism (Oxford University Press, New York 1961) Chap. 10
134. R.E. Dietz, F.R. Merritt, R. Dingle, D. Hone, G.G. Silbernagel, P.M. Richards, Phys. Rev. 26, 1186 (1971)
135. T. Moriya, Prog. Theor. Phys. 16, 23, 641 (1956)
136. C.P. Slichter: Principles of Magnetic Resonance. 2nd edn. Springer Series in Solid-State Sciences, Vol. 1 (Springer, Berlin Heidelberg New York 1980)
137. E. Hahn, Phys. Rev. 80, 580 (1950)
138. N. Bloembergen, E.M. Purcell, R.V. Pound, Phys. Rev. 73, 679 (1948)
139. V.L. Safonov, H.N. Bertram, Phys. Rev. B 71, 224402 (2005)

Chapter 7

140. S. Doniach, Proc. Phys. Soc. London 91, 86 (1967)

141. N.F. Berk, J.R. Schrieffer, Phys. Rev. Lett. 17, 433 (1966); S. Doniach, S. Engelsberg, Phys. Rev. Lett. 17, 750 (1966)

142. N.D. Mermin, Phys. Rev. B1, 2362 (1970)

143. V.P. Silin, Sov. Phys. JETP 8, 870 (1959)

144. P.M. Platzman, P.A. Wolff, Phys. Rev. Lett. 18, 280 (1967)

145. F.J. Dyson, Phys. Rev. 98, 349 (1955); G. Feher, A.F. Kip, Phys. Rev. 98, 337 (1955)

146. S. Schultz, G. Dunifer, Phys. Rev. Lett. 18, 283 (1967)

147. H. Hasegawa, Prog. Theor. Phys. 21, 483 (1959)

148. S. Schultz, M.R. Shanabarger, P.M. Platzman, Phys. Rev. Lett. 19, 749 (1967)

149. P.S. Pershan, J. Appl. Phys. 38, 1482 (1967)

Chapter 8

150. F. Keffer: In Handbuch der Physik, Vol. 18, Part 2, Ferromagnetism, ed. by S. Flügge (Springer, Berlin Heidelberg New York 1966)

151. R.M. White, M. Sparks, I. Ortenberger, Phys. Rev. A 139, 450 (1965)

152. C. Kittel, Phys. Rev. 73, 155 (1948)

153. N. Zagury, S. Rezende, Phys. Rev. B 4, 201 (1971)

154. L.R. Walker, J. Applied Phys. 29, 318 (1958)

155. R.L. White, I.H. Solt, Jr., Phys. Rev. 104, 56 (1956)

156. J.F. Dillon, Jr., J. Appl. Phys. 31, 1605 (1960)

157. R.W. Damon, J.R. Eshbach, J. Phys. Chem. Solids 19, 308 (1981)

158. M. Chen, M.A. Tsankov, J. Nash, C. Patton, Phys. Rev. B 49, 12773 (1994)

159. F. Bloch, Z. Phys. 61, 206 (1930)

160. G. Low, Proc. Phys. Soc. London 82, 992 (1963)

161. F.J. Dyson, Phys. Rev. 102, 1217 (1956)

162. M. Sparks: "Ferromagnetic Relaxation Theory" (McGraw-Hill, New York 1964)

163. M. Sparks, R. Loudon, C. Kittel, Phys. Rev. 122, 791 (1961)

164. N. Bloembergen, R.W. Damon, Phys. Rev. 85, 699 (1952)

165. H. Suhl, Proc. IRE 44, 1270 (1957)

166. G. Gibson, C. Jeffries, Phys. Rev. A 19, 811 (1984)

167. S.M. Rezende, F.M. de Aguiar, Proc. IEEE 78, 893 (1990)

168. S.J. Allen, R. Loudon, P.L. Richards, Phys. Rev. Lett. 16, 463 (1966)

169. R.L. Greene, D.D. Sell, W.M. Yen, A.L. Schawlow, R.M. White, Phys. Rev. Lett. 15, 656 (1965); D.D. Sell, R.L. Greene, R.M. White, Phys. Rev. 158, 489 (1967)

170. P.A. Fleury, R. Loudon, Phys. Rev. 166, 514 (1968)

171. P.M. Richards, unpublished

172. R. Kubo, J. Phys. Soc. Jpn. 12, 570 (1957)

173. P.M. Richards, P.K. Leichner, Phys. Rev. B7, 453 (1973)

174. J.C. Burgiel, M.W.P. Strandberg, J. Appl. Physics 35, 852 (1964)

175. L.D. Landau, E.M. Lifshitz, Phys. Z. Sowjet 8, 153 (1935). [English translation: L.D. Landau, Collected Papers, ed. by D. ter Haar (Gordon and Breach, New York 19676) p. 101

176. W.F. Brown, Phys. Rev. 58, 736 (1940)

177. J.-G. Zhu, Y. Sheng, "Spin Dynamics in Confined Magnetic Structure I", ed. by B. Hillebrands, K. Ounadjela (Springer, Berlin Heidelberg New York 2002) pp. 289-325

178. G. Binning, C.F. Quate, C. Gerber, Phys. Rev. Lett. 56, 930 (1986)

179. C. Mao, Ph.D. Thesis, Carnegie Mellon University (1999)

180. C.Y. Mao, J.-G. Zhu, R.M. White, J. Appl. Phys. 87, 5416 (2000)

181. See T.L. Gilbert, IEEE Trans. Mag. 40, 3443 (2004)

182. T. Izuyama, E.J. Kim, R. Kubo: J. Phys. Soc. Jpn. 18, 1025 (1963)

183. C. Herring: In "Magnetism", ed. by G.T. Rado, H. Suhl (Academic Press, New York 1966) Vol. IV, Sect. XIV,3

184. G. Barnea, G. Horwitz, J. Phys. C6, 738 (1973)

185. R. Lowde, Proc. R. Soc. London A235, 305 (1956)

186. J. Lynn, Phys. Rev. B 11, 2624 (1975)

187. Mook, J. Lynn, Nicklow, Phys. Rev. Lett. 30, 556 (1973)

188. X. Tao, D.P. Landau, T.C. Schulthess, G.M. Stocks, Phys. Rev. Lett. 95, 08707 (2005)

189. M.H. Seavey, P.E. Tannenwald, Phys. Rev. Lett. 1, 168 (1958)

190. J.R. Sandercock, W. Wettling, IEEE Trans. M-14, 442 (1978)

Chapter 9

191. W.H. Meiklejohn, C.P. Bean, Phys. Rev. 102, 1413 (1956)

192. A.P. Malozemoff, J. Appl. Phys. 63, 3874 (1988)

193. Imry, Ma, Phys. Rev. Lett. 35, 1399 (1975)

194. S.M. Rezende, A. Azevedo, M.A. Lucena, F.M. de Aguir, Phys. Rev. B 63, 214418 (2001)

195. C.F. Majkrzak, J.W. Cable, J. Kwo, M. Hong, D.B. McWhan, Y. Yafet, J.V. Waszcczak, C. Vettier, Phys. Rev. Lett. 56, 2700 (1986)

196. P. Grünberg, R. Schreiber, Y. Pang, M.B. Brodsky, H. Sowers, Phys. Rev. Lett. 57, 2442 (1986)

197. H. Fröhlich, F.R.N. Nabarro, Proc. R. Soc. London A175, 382 (1940)

198. M.A. Ruderman, C. Kittel, Phys. Rev. 96, 99 (1954)

199. C. Zener, Phys. Rev. 81, 440 (1951)

200. T. Kasuya, Prog. Theor. Phys. 16, 45, 58 (1956)

201. K. Yosida, Phys. Rev. 106, 893 (1957)

202. J.B. Boyce, C.P. Slichter, Phys. Rev. B B3, 379 (1976)

203. P. Bruno, Phys. Rev. 52, 411 (1995)

204. J. Unquris, R. Celotta, D. Pierce, J. Appl. Phys. 75, 6437 (1994); Phys. Rev. Lett 67, 140 (1991)

205. M. Stiles, Phys. Rev. B 48, 7238 (1993)

206. M.N. Baibich, J.M. Broto, A. Fert, F. Nguyen Van Dau, F. Petroff, P. Etienne, P. Creuzet, A. Friederich, J. Chazelas, Phys. Rev. Lett. 61, 2472 (1988)

207. E.Y. Tsymbal, D.G. Pettifor, Solid State Physics 56, 113 (2001)

208. G. Mahan: "Many-Particle Physics", 3rd Edn. (Kluwer Academic/Plenum, New York 2000) pp. 163-165

209. W.H. Butler, X.-G. Zhang, D.M.C. Nicholson, J.M. MacLaren, Phys. Rev. B 52, 13399 (1995)

210. R.E. Camley, J. Barnaś, Phys. Rev. Lett. 63, 664 (1989)

211. B.L. Johnson, R.E. Cmley, Phys. Rev. 44, 9997 (1991)

212. R.E. Camley, R.L. Stamps, J. Phys.: Condensed Matter 5, 3727 (1993)

213. H.E. Camblong, P.M. Levy, Phys. Rev. Lett. 69, 2835 (1992)

214. B. Dieny, Europhys. Lett. 17, 261 (1992)

215. Bardeen, Phys. Rev. Lett. 6, 57 (1961)

216. M. Julliere, Phys. Lett. A 54, 225 (1975)

217. R. Meservey, P.M. Tedrow, Phys. Rep. 238, 174 (1994)

218. R.J. Soulen et al., Science 282, 85 (1998)

219. D.J. Monsma and S.S.P. Parkin, Appl. Phys. Lett. 77, 720 (2000); J.S. Moodera, J. Nassar, and G. Mathon, Ann. Rev. Mater. Sci. 29, 381 (1999)

220. J.H. Park et al., Nature 392, 794 (1998); Y.S. Dedkov et al., Phys. Rev. B 65, 064417-1 (2002)

221. J.C. Slonezewski, Phys. Rev. G 39, 6995 (1989)

222. W.H. Butler, X.-G. Zhang, T.C. Schulthess, J.M. MacLaren, Phys. Rev. B 63 054416 (2001); J. Mathon, A. Umerski, Phys. Rev. B 63 220403(R) (2001)

223. S.S. Parkin, C. Kaiser, A. Panchula, P.M. Rice, B. Hughes, M. Samant, S.-H. Yang, Nature Mater. 3, 862 (2004)

224. J.C. Slonczewski, J. Mag. Mag. Mat. 159, 67 (1996)

225. L. Berger, Phys. Rev. B54, 9353 (1996)

226. M. Tsoi, A.G.M. Jansen, J. Bass, W.-C. Chiang, M. Seck, V. Tsoi, P. Wyder, Phys. Rev. Lett. 980, 4281 (1998)

227. J.Z. Sun, Mag. Mag. Mat. 202, 157 (1999)

228. J.A. Katine, F.J. Albert, R.A. Buhrman, E.B. Meyers, D.C. Ralph, Phys. Rev. Lett. 84, 3149 (2000)

229. M.D. Stiles, A. Zangwill, J. Appl. Phys. 91, 6812 (2002)

230. J.Z. Sun, Phys. Rev. B 62, 570 (2000)

231. S.M. Rezende, F.M. de Aguiar, A. Azevedo, Phys. Rev. Lett. 94, 037202 (2005)

232. X. Zhu, J.-G. Shu, R.M. White, J. Appl. Phys. 95, 6630 (2004); J.-G. Zhu, X. Zhu, IEEE Trans. Mag. 40, 182 (2004)

233. See, for example, S. Zhang, Phys. Rev. B 51, 3632 (1995)

234. J.E. Hirsch, Phys. Rev. Lett. 83, 1834 (1999)

235. S. Zhang, Phys. Rev. Lett. 85, 393 (2000)

236. Y.K. Kato. R.C. Myers, A.C. Gossard, D.D. Awschalom, Science 306, 1910 (2004)

237. P.-G. deGennes, J. Phys. Radium 23, 630 (1962)

Chapter 10

238. P.A. Egelstaff (Ed.): Thermal Neutron Scattering (Academic Press, New York 1965)

239. W. Marshall, S.W. Lovesey: Theory of Thermal Neutron Scattering (Oxford University Press, Oxford 1971)

240. S.W. Lovesey, T. Springer (Eds.): Dynamics of Solids and Liquids by Neutron Scattering. Topics in Current Physics, Vol. 3 (Springer, Berlin Heidelberg New York 1977)

241. R.E. Rundle, C.G. Shull, E.O. Wollen, Acta Crystallogr. 5, 22 (1952)

242. N.W. Ashcroft, N.D. Mermin: Solid State Physics (Holt, Rinehort, and Winston, New York 1974) Appendix L

243. R.M. Moon, T. Riste, W.C. Koehler, Phys. Rev. 181, 920 (1969)

244. R.A. Erickson, Phys. Rev. 90, 779 (1953)

245. T. Izuyama, D. Kim, R. Kubo, J. Phys. Soc. Jpn. 18, 1025 (1963)

246. C.C. Becerra, Y. Shapira, N.F. Oliveira, T.S. Chang, Phys. Rev. Lett. 44, 1692 (1980)

247. R.M. Moon, J.W. Cable, Y. Shapira, J. Appl. Phys. 52, 2025 (1981)

248. C.G. Shull, Y. Yamada, J. Phys. Soc. Jpn. 17, Suppl. B-III, 1 (1962)

249. C.G. Shull, H.A. Mook, Phys. Rev. Lett. 16, 184 (1966)

250. G.C. Low, A. Okazaki, R.W.H. Stevenson, K.C. Tuberfield, J. Appl. Phys. 35, 998 (1964)

251. H. Möller, J. Houmann, A. Mackintosh, J. Appl. Phys. 39, 807 (1968)

252. D. Cribier, T. Jacrot, G. Parette, J. Phys. Soc. Jpn. 17, Suppl. B-III 67 (1962)

253. R.D. Lowde, C.G. Windsor, Adv. Phys. 19, 813 (1970)

254. C.G. Shull, M.K. Wilkinson, Phys. Rev. 97, 304 (1955)

255. W. Marshall, J. Phys. C1, 88 (1968)

256. R.A. Medina, J.W. Cable, Phys. Rev. B 15, 1539 (1977)

257. Y. Tokura, Y. Tomioka, J. Mag. Mag. Mat. 200, 1 (1999)

258. D.C. Worledge, Ph.D. Thesis, Stanford University (2000)

259. J.W. Lynn, C.P. Adams, Y.M. Mukovskii, A.A. Arsenov, D.A. Shulyatev, J. Appl. Phys. 89, 6846 (2001)

260. M.B. Salamon, M. Jaime, Rev. Mod. Phys. 73, 583 (2001)

261. A. Oosawa, M. Ishii, H. Tanaka, J. Phys. Condens. Matter 11, 265 (1999)

262. N. Cavadini et al., Phys. Rev. B 63, 172414 (2001)

263. O. Nohadani, et al., Phys. Rev. B 69, 220402 (2004)

264. M. Matsumoto, B. Normand, T.M. Rice, M. Sigrist, Phys. Rev. Lett. 89, 077203 (2002)

265. T. Nikuni, M. Oshikawa, A. Oosawa, H. Tanaka, Phys. Rev. Lett. 84, 5868 (2000)

266. For a short summary of Bose condensation see A.L. Fetter, J.D. Walecka, "Quantum Theory of Many-Particle Systems". (McGraw-Hill, New York, 1971) Chap. 1

267. C. Rüegg, N. Cavadini, A. Furrer, K. Krämer, H.U. Güdel, P. Vorderwisch, H. Mutka, Appl. Phys. A 74 [suppl.] S840 (2002)

268. L. Landau, E.M. Lifshitz: Statistical Physics, 3rd edn., Sect. 71

269. C. Rüegg, et al., Nature (London) 423, 62 (2003)

270. V.N. Glazkov, A.I. Smirnov, H. Tanaka, A. Oosawa, Phys. Rev. B 69, 184410 (2004)

271. M. Matsumoto, B. Normand, T.M. Rice, M. Sigrist, Phys. Rev. B 69, 054423 (2004)

272. C. Rüegg, et al., Phys. Rev. Lett. 93, 257201 (2004)

273. S.E. Sebastion, P.A. Sharma, M. Jaime, N. Harrison, V. Correa, L. Balicas, N. Kawashima, C.D. Batista, I.R. Fisher, Phys. Rev. B 72, 100404R (2005)

274. T. Radu et al., Phys. Rev. Lett. 95, 127202 (2005)

Part IV
Problems and Questions

11 General questions

11.1 Общие вопросы

11.1.1 Typical exam questions

Leuven.exam.2022 (???)

(потом пропишу решения и ответы)

What are the main differences between a typical paramagnetic system and a typical superparamagnetic system? Answer this general question by answering the questions (a-e) below.

a) Between these two systems:

i) dilute Fe ions in a nonmagnetic matrix

ii) Fe nanoparticles (few nm in diameter) in a non-magnetic matrix Which one is more likely to behave as a typical paramagnet and which as a typical superparamagnet?

b) Draw typical magnetization curves for each of them (qualitative, no calculations required):

i) $M(H)$ at 5 K and 300 K, for μ_0H between 0 and 5 T.

ii) $M(T)$ between 5 K and 300 K, in a low field H (μ_0H $\approx 1 \times 10^{-3}$ T), after cooling in field.

c) Under certain conditions, the magnetization curves of paramagnets and superparamagnets can be described (to a reasonable approximation) by the same theory/function, and only one parameter/quantity is responsible for the differences highlighted in your answer to (b). Identify this parameter/quantity and comment on the difference that it originates in the magnetization curves drawn in (b).

(you can explain it in about 3 lines of text; no equations are explicitly required, but simple expression(s) may help your discussion)

d) The Langevin function is a good approximation to describe superparamagnetic systems, but a poorer description to describe a typical paramagnet. Comment on the fundamental reason for that. (hint: quantum mechanics) (you can explain it about 4 lines of text; no equations are explicitly required, but simple expression(s) may help your discussion)

e) Furthermore, even typical superparamagnets often deviate from Langevin function behavior at low temperature and low field. Describe this effect and its origin. (you can explain it in about 4 lines of text)

Расскажите о Магнитные моменты. Диамагнетизм и парамагнетизм.

Расскажите о Ферромагнитный резонанс.

Расскажите о Взаимодействие магнитных моментов. Обменный механизм.

Расскажите о Спиновые волны.

Расскажите о теория Ландау-Гинзбурга в магнетизме и симметрийном анализе

(см. ЛЛ8, там много про это, у меня частично уже написано это тоже, но это большие темы.)

Расскажите о Переход спин-флоп.

Расскажите о Одномерные системы. Преобразование Джордана-Вигнера.

Расскажите о Анзац Бете.

(см интегрируемые системы, пока не доучивал, отдельный метод)

Расскажите о Спин-волновая теория I.

Расскажите о Экспериментальные методы изучения магнитных материалов.

Расскажите о Взаимодействие магнонов I. Связанное состояние магнонов.

Расскажите о Взаимодействие магнонов II.

Расскажите о Спиновые жидкостях

11.1.2 Typical exam problems

Определите восприимчивость кислорода (!?!?!!!!)

(додумаю, уже близок к этому!!!)

Молекула кислорода имеет восемь валентных электронов. Их суммарный орбитальный момент равен нулю. Спины шести электронов попарно направлены в противоположные стороны, а у двух спины параллельны. Таким образом, молекула кислорода обладает магнитным моментом, равным двум магнетонам Бора, и восприимчивость кислорода равна

$$\varkappa_{\mathrm{M}} = N\frac{\mu_{\mathrm{B}}^2 g^2 \Sigma(\Sigma-1)}{3kT} = \frac{8N\mu_{\mathrm{B}}^2}{3kT} = \frac{0,993}{T}.$$

11.1.3 Exam questions about Exotic magnetic models

Расскажите о Модель Китаева I.

Расскажите о Модель Китаева II.

Расскажите о Бозе-эйншейновская конденсация магнонов.

Расскажите о Орбитальное упорядочение

Расскажите о ЯМР в изучении магнитных свойств

Расскажите о Спинтроника

(????)

Расскажите о Распад магнонов

(см кинетику тоже)

Расскажите о магнетострикции: магнон-фононное взаимодействие

(см ЛЛ8 и ЛЛ9)

Расскажите о Мультиферроики

Расскажите о Стоунеровский магнетизм в металлах

Расскажите о Теория кристаллического поля

Расскажите о Устройство нейтронных спектрометров

Расскажите о Магнонный термический эффект Холла

Расскажите о Спектр спинонов в АФМ Гейзенберге 1Д

(см мой статфиз, дописал бы запросто, если бы занялся эти)

Расскажите о J1-J2 модель на квадратной (треугольной) решетке

(???)

Расскажите о Метод DMRG

(специфика)

Расскажите о Order from disorder effect

12 Problems

12.1 Problems about magnetic materials

(тут понятно, как они решаются, просто набью потом руку, см. запись по ЭД)

12.1.1 Problems about магнетизм типичных веществ

Leuven-practice-2024 (??)

1 Use Hund's rules to determine the values of S, L and J in the ground state of Sc^{2+}, Ho^{3+} and Tm^{3+}.

2 Calculate the diamagnetic orbital susceptibility of a gas of hydrogen atoms (H atoms, not H_2 molecules), with a number density of 10^{20} m^{-3}, in the ground state (assume the Bohr radius a_0). Calculate the paramagnetic spin susceptibility at 100 K (hint: α is small...). Compare the two.

3 Calculate the paramagnetic susceptibility of $CuSO_4 \cdot 5H_2O$, at 10 K and 300 K, assuming that only Cu^{2+} ions contribute to the magnetization $\left(\rho\left(CuSO_4 \cdot 5H_2O\right) = 2.286 \text{ g cm}^{-3}\right)$:

a) assuming no orbital quenching (Hund's rules)

b) assuming orbital quenching

c) compare these two values to the experimentally determined value $\left(1.76 \times 10^{-4} \text{ at } 300 \text{ K}\right)$

d) calculate the magnetic moment of a crystal with dimensions $2 \times 2 \times 2$ mm^3, at 300 K, under a field of 1 T (using the experimentally determined susceptibility)

(это же в конд средах будет и в записи про магнетизм, пока тут)

Spaldin.2.1(5 min) Units and magnetic moment

A cylindrical bar magnet 10 inches long and 1 inch in diameter has a magnetic moment of 10000 erg/ Oe.

(a) What is its magnetic moment in SI units?

(b) What is its magnetization in both cgs and SI units?

(c) What current would have to be passed through a 100-turn solenoid of the same dimensions to give it the same magnetic moment?

Solution

The solution is obvious, one just needs to be careful. Still, it is a nice practice, to do it at least once.

2.1(a)

$$1 \text{ erg} = 10^{-7} \text{ J}, \qquad 1 \text{ Oe} = 10^{-4} \text{ T},$$

$$10000\text{erg/ Oe} = \frac{10000 \cdot 10^{-7}}{10^{-4}} \text{ J/T} = 10 \text{ J/T}$$

2.1(b) $1 \text{ in} = 2.54 \text{ cm} = 2.54 \cdot 10^{-2}$ m. The volume of the cylinder is $\pi r^2 l = 128.704 \text{ cm}^3 = 1.28704 \cdot 10^{-4} \text{ m}^3$.

$$M_{\text{cgs}} := \frac{\boldsymbol{m}}{V} = \frac{10000\text{erg/ Oe}}{128.704 \text{ cm}^3} = 77.70\text{emu/cm}^3$$

since 1erg/ Oe = 1 emu.

$$M_{\text{SI}} = \frac{m}{V} = \frac{10 \text{ J T}^{-1}}{1.28704 \cdot 10^{-4} \text{ m}^3} = 77.7 \cdot 10^3 \frac{\text{kgm}^2 \text{ s}^{-2}}{\text{m}^3 \text{ kg s}^{-2} \text{ A}^{-1}} = 77.7 \cdot 10^3 \text{ A/m}$$

2.1(c) For a current loop, the magnetic moment $\boldsymbol{m} = IA$. For a solenoid with N loops, the magnetic moment is NIA. In SI units, $10 \text{ J/T} = 100 \cdot I \cdot \pi \cdot (0.0127)^2 \text{ m}^2$, so

$$I = 197.3 \text{ J/}\left(\text{Tm}^2\right) = 197.3 \text{ A}$$

(??? а что насчет того, что там расстояние разное от точки в центре до разного витка с током??? пренебрегаем? я пока забыл точную формулу, когда-то мб додумаю это)

Spaldin.3.1(2 min) Magnetic moment

Calculate the allowed values of the magnetic moment along the field axis of an atom which has $J = 1$ and $g = 2$.

Solution

The solution is obvious, one just needs to be careful. Still, it is a nice practice, to do it at least once.

$$\mu_{\text{atom}} = g\sqrt{J(J+1)}\mu_{\text{B}},$$
$$\mu_{\text{z}} = gM_J\mu_{\text{B}}.$$

For $J = 1$, $\sqrt{J(J+1)} = \sqrt{2}$ and $M_J = -1, 0$, or 1, so

$$\mu_{\text{atom}} = 2\sqrt{2}\mu_{\text{B}},$$

$$\mu_{\text{z}} \in \{-2\mu_{\text{B}}, 0, +2\mu_{\text{B}}\}.$$

Spaldin.3.2(5 min) Magnetism of iron

Calculate the electronic structure and magnetic properties of a transition-metal ion: Fe^{2+}.

(a) What is the electronic configuration of an Fe^{2+} ion? (Note that the transition metals give up their 4 s electrons before their 3 d electrons on ionization.)

(b) Use Hund's rules to determine the values of S, L, and J in the ground state of an Fe^{2+} ion.

(c) Calculate the Landé g-factor.

(d) Calculate the total magnetic moment of an Fe^{2+} ion and the magnetic moment along the field direction.

(e) Try calculating the total magnetic moment using the value of S determined in (b), but assuming that $L = 0$ (so $J = S$). In fact the measured value is $5.4\mu_{\text{B}}$.

Solution to (a)

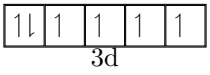
$$\text{Fe}: \qquad (1\text{ s})^2(2\text{ s})^2(2\text{p})^6(3\text{ s})^2(3\text{p})^6(4\text{ s})^2(3\text{ d})^6.$$

Because the transition elements give up their 4 s electrons before their 3 d electrons on ionization:

$$\text{Fe}^{2+}: \qquad (1\text{ s})^2(2\text{ s})^2(2\text{p})^6(3\text{ s})^2(3\text{p})^6(3\text{ d})^6$$

Solution to (b)

Using Hund's first rule we maximize S and L:



The total spin $S = 4 \cdot \frac{1}{2} = 2$.

m_l has values of $-2, -1, 0, 1$, and 2. Since $M_L \equiv -L, -L+1, \ldots, 0, \ldots, L-1, L$, the total orbital quantum number, L=2.

(! note that me don't have $L = \frac{1}{2} \cdot 2$ or $L = \sqrt{L(L+1)}$! because we need to use the logic above.)

Finally, from Hund's third rule, because the shell is more than half full, $J = L + S = 4$

Solution to (c)

$$g = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)} = 1 + \frac{20 + 6 - 6}{40} = 1.5.$$

Note that, since $S = 2$ and $L = 2$, the g value is exactly halfway between that for the $S = 0$ case ($g = 1$) and the spin-only case ($g = 2$).

Solution to (d)

$$\mu = g\sqrt{J(J+1)}\mu_{\text{B}} = 1.5\sqrt{4 \cdot 5} = 6.7\mu_{\text{B}}.$$

Since $J = 4, M_J = -4, -3, -2, -1, 0, 1, 2, 3, 4$, the $\mu_z \equiv gM_J\mu_{\text{B}}$ can take the values

$6\mu_{\text{B}}$, $4.5\mu_{\text{B}}, 3\mu_{\text{B}}, 1.5\mu_{\text{B}}, 0, -1.5\mu_{\text{B}}, -3\mu_{\text{B}}, -4.5\mu_{\text{B}}$, or $-6\mu_{\text{B}}$.

Solution to (e)

If $L = 0$, then $J = S = 2$, $g = 2$ and

$$\mu = 2\sqrt{6}\mu_{\mathrm{B}} = 4.9\mu_{\mathrm{B}}.$$

This is in good agreement with the measured value of $5.4\mu_{\mathrm{B}}$, whereas the moment we calculated using the total angular momentum ($6.7\mu_{\mathrm{B}}$) does not agree well with the experimental value. This is a manifestation of a phenomenon known as quenching of the orbital angular momentum (???).

Spaldin.4.1(5 min) Susceptibility of a carbon

The value of $\sqrt{\langle r^2 \rangle_{\mathrm{av}}}$ for carbon is known from X-ray diffraction measurements to be around 0.7Å. The density is 2220 kg/m^3. Calculate the value of susceptibility, give your answer in SI and cgs units.

The measured value of susceptibility is $\xi_{\mathrm{exper}} = -13.82 \cdot 10^{-6}$ per unit volume. The agreement for carbon is better than that for most diamagnets. Comment on possible sources of error in the derivation.

Solution

The solution is obvious, one just needs to be careful. Still, it is a nice practice, to do it at least once.

$$\chi_{\mathrm{SI}} = -\frac{N\mu_0 Ze^2}{6m_{\mathrm{e}}} \left\langle r^2 \right\rangle_{\mathrm{av}}.$$

Here $N = N_{\mathrm{A}}\rho/A$, where N_{A} is Avogadro's number (the number of atoms per mole), ρ is the density, and A is the molar mass, μ_0 is the permeability of free space, Z is the number of electrons per atom, e is the electronic charge, m_{e} is the mass of the electron, and $\left\langle r^2 \right\rangle_{\mathrm{av}}$ is the average distance squared of the electrons from the nucleus. For carbon, $Z = 6$ and $A = 12$ g/mol, so

$$\begin{aligned}\chi =& -\frac{6.022 \cdot 10^{23}\ \mathrm{mol}^{-1} \cdot 2220\ \mathrm{kg\ m}^{-3}}{12 \cdot 10^{-3}\ \mathrm{kg\ mol}^{-1}} \cdot \frac{1.256 \cdot 10^{-6}\ \mathrm{H\ A}^{-2} \cdot 6\left(1.60 \cdot 10^{-19}\right)^2 \mathrm{C}^2 \cdot \left(0.7 \cdot 10^{-10}\right)^2\ \mathrm{m}^2}{6 \cdot 9.109 \cdot 10^{-31}\ \mathrm{kg}} \\ =& -19.33 \cdot 10^{-6}\ \mathrm{m}^{-3}\mathrm{HA}^{-2}\ \mathrm{C}^2\ \mathrm{m}^2\ \mathrm{kg}^{-1} = -19.33 \cdot 10^{-6}.\end{aligned}$$

This is reasonably close to the experimental value of $-13.82 \cdot 10^{-6}$.

Now, in cgs units

$$\begin{aligned}\chi_{\mathrm{cgs}} =& -\frac{NZe^2}{6m_{\mathrm{e}}c^2} \left\langle r^2 \right\rangle_{\mathrm{av}} = -\frac{6.022 \cdot 10^{23}\ \mathrm{mol}^{-1} \cdot 2.22\ \mathrm{g\ cm}^{-3}}{12\ \mathrm{g\ mol}^{-1}} \frac{6 \cdot \left(4.8 \cdot 10^{-10}\right)^2\ \mathrm{esu}^2 \cdot \left(0.7 \cdot 10^{-8}\right)^2\ \mathrm{cm}^2}{6 \cdot 9.109 \cdot 10^{-28}\ \mathrm{g} \cdot \left(3 \cdot 10^{10}\right)^2\ \mathrm{cm\ s}^{-2}} \\ =& -1.5 \cdot 10^{-6}\mathrm{emu}/\left(\mathrm{cm}^3\ \mathrm{Oe}\right).\end{aligned}$$

(??? don't know, how to get here "Oe" at the end, not so interesting)

In general,

$$\chi_{\mathrm{cgs}} = \frac{\chi_{\mathrm{SI}}}{4\pi}$$

For most materials, the value calculated using the classical Langevin model only shows order-of-magnitude agreement with experiment. Possible sources of the discrepancies between theory and experiment include the following:

- application of Lenz's law (which was determined for electrical circuits) on the atomic scale (???)
- difficulty in calculating or measuring $\left\langle r^2 \right\rangle_{\mathrm{av}}$; in particular, χ depends on the choice of origin for computing $\left\langle r^2 \right\rangle_{\mathrm{av}}$
- our assumption that the electrons are orbiting their nucleus; therefore, we expect a poor description of itinerant conduction electrons
- our assumption that the system is spherically symmetric.

(??? how can we know, that Carbon is a diamagnet? 1) from example in the problem statement, 2) from remembering magnetic periodic table, 3) maybe we can calculate by hunds rule its ground state(???? I'll think about it later maybe). What if we try to calculate paramagnetic contribution????)

Spaldin.5.1 Brillouin approaches Langevin

Show that the Brillouin function $B_J(\alpha)$ approaches the Langevin function as $J \to \infty$. What are the limits of the Brillouin function if (a) $J \to \frac{1}{2}$ and if (b) $\alpha \to 0$?

Solution

The solution is obvious, one just needs to be careful. Still, it is a nice practice, to do it at least once.

$$B_J(\alpha) := \frac{2J+1}{2J}\coth\left(\frac{2J+1}{2J}\alpha\right) - \frac{1}{2J}\coth\left(\frac{\alpha}{2J}\right)$$

As $J \to \infty$, $2J+1 \to 2J$, $(2J+1)/2J \to 1$. with $\coth(x) \approx \frac{1}{x} + \frac{x}{3} - \cdots$, we have

$$-\frac{1}{2J}\coth\left(\frac{\alpha}{2J}\right) \approx -\frac{1}{2J}\frac{2J}{\alpha} - \frac{1}{2J}\frac{\alpha}{6J} \to -\frac{1}{\alpha}$$

$$B_J(\alpha) \overset{J\to\infty}{\to} \coth(\alpha) - \frac{1}{\alpha} = L(\alpha)$$

If $J \to \frac{1}{2}$, we have $(2J+1)/2J \to 2$, and $2J \to 1$. So, at $J = \frac{1}{2}$,

$$B_J(\alpha) = 2\coth(2\alpha) - \coth(\alpha) = 2\frac{e^{2\alpha}+e^{-2\alpha}}{e^{2\alpha}-e^{-2\alpha}} - \frac{e^{\alpha}+e^{-\alpha}}{e^{\alpha}-e^{-\alpha}} = \frac{e^{2\alpha}+e^{-2\alpha}-2}{\left(e^{\alpha}+e^{-\alpha}\right)\left(e^{\alpha}-e^{-\alpha}\right)} = \tanh(\alpha).$$

And if $\alpha \to 0$, second terms will give a result: $\coth[\frac{(2J+1)\alpha}{2J}] \to \frac{2J}{(2J+1)\alpha} + \frac{(2J+1)\alpha}{3\cdot 2J}$ and $\coth(\alpha/2J) \to \frac{2J}{\alpha} + \frac{\alpha}{3\cdot 2J}$, so

$$B_J(\alpha) \to \frac{2J+1}{2J}\cdot\frac{2J}{\alpha(2J+1)} - \frac{1}{2J}\cdot\frac{2J}{\alpha} + \left(\frac{2J+1}{2J}\right)^2\frac{\alpha}{3} - \left(\frac{1}{2J}\right)^2\frac{\alpha}{3} = \frac{\left[(2J+1)^2-1\right]\alpha}{12J^2} = \alpha\frac{(J+1)}{3J}.$$

Spaldin.5.2(5 min) Ideal gas susceptibility

Calculate the room-temperature paramagnetic susceptibility of an ideal gas, in which each atom has $J = 1$ and $g = 2$. (These are in fact the values of J and g for molecular oxygen. Remember the ideal gas law: $PV = nRT$.)

Solution

The solution is obvious, one just needs to understand correctly basic the formulas.

$$\chi = \frac{Ng^2J(J+1)\mu_0\mu_{\mathrm{B}}^2}{3k_{\mathrm{B}}T}.$$

Here N is a number atoms per m^3 and other values are known: $J = 1, g = 2, \mu_0 = 4\pi\cdot 10^{-7}\mathrm{H/m}, \mu_{\mathrm{B}} = 9.274\cdot 10^{-24}$ J/T, $k_{\mathrm{B}} = 1.380662\cdot 10^{-23}$ J/K, and $T = 273$ K:

$$\chi = \frac{N\cdot 8\cdot 4\pi\cdot 10^{-7}\cdot\left(9.274\cdot 10^{-24}\right)^2}{3\cdot 1.380662\cdot 10^{-23}\cdot 273}\frac{\mathrm{Hm^{-1}\ J^2\ T^{-2}}}{\mathrm{J\ K^{-1}\ K}} = 7.6465\cdot 10^{-32}N\frac{\mathrm{HJ}}{\mathrm{mT^2}} = 7.6465\cdot 10^{-32}\ \mathrm{N\ m^3}.$$

Now, from $PV = nRT = NkT$, (n is the number of moles of atoms), with $P = 1$ atm $= 10^5$ N m^{-2}, $T = 273$ K, and the volume $V = 1$ m^3 we get the number of atoms per m^3:

$$N = \frac{PV\cdot N_{\mathrm{A}}}{RT} = \frac{10^5}{1.38\cdot 10^{-23}\cdot 273}\frac{\mathrm{Nm^{-2}\ m^3\ mol^{-1}}}{\mathrm{J\ mol^{-1}\ K^{-1}\ K}} = 2.688\cdot 10^{25}.$$

Finally,

$$\chi = 2.056\cdot 10^{-6}.$$

(???? why can't we use $N = N_A\rho/A$ (????? maybe because we are not sure about A???))

Spaldin.5.3 Basic lattice spins

In this problem we will explore the properties of a model three-dimensional lattice of spins, each with spin $S = \frac{1}{2}$.

(a) What is the magnetic moment of each spin? What are the allowed values of the projection of the magnetic moment, $\boldsymbol{m}_i$, onto some chosen axis, say z?

(b) What are the possible values of the magnetic energy of each spin?

(c) Assuming that the spins are non-interacting, calculate the magnetization of the lattice of spins when a magnetic field, $\boldsymbol{H}$, is applied along the z axis.

(d) For a given value of field, $\boldsymbol{H}$, how does the magnetization, $\boldsymbol{M}$, depend on temperature? Explain the behavior of $\boldsymbol{M}$ for $T \to 0$. Taking the number of spins per unit volume to be $3.7 \cdot 10^{28}$ m^{-3}, calculate the numerical value of the saturation magnetization, $\boldsymbol{M}_{\mathrm{s}}$, at $T = 0$. Explain the behavior of $\boldsymbol{M}$ for $T \to \infty$.

(e) What does the relationship between $\boldsymbol{M}$ and $\boldsymbol{H}$ reduce to for weak fields ($\boldsymbol{H} \to 0$)? What is the expression for the susceptibility, χ, in this case, and how does it depend on temperature? Calculate the numerical value of χ at room temperature.

(f) Comment on the results which you have obtained for this spin system. What kind of magnetic behavior (antiferromagnetic, paramagnetic, diamagnetic, etc.) is displayed by this model system? Justify your conclusion. How would we need to modify the model in order to describe ferromagnetic behavior?

Solution to (a)

$$\mu = g_{\mathrm{e}}\mu_{\mathrm{B}}\sqrt{S(S+1)}$$
$$\mu_z = g_{\mathrm{e}}\mu_{\mathrm{B}}m_s$$

When $S = 1/2$ and $g = 2$, we have $\mu = \sqrt{3}\mu_{\mathrm{B}}$ and $\mu_z = \pm\mu_{\mathrm{B}}$.

Solution to (b)

$$E = -\boldsymbol{m} \cdot \boldsymbol{H},$$

the allowed magnetic energies are $\mp\mu_{\mathrm{B}}H$, $H = |\boldsymbol{H}|$.

Solution to (c)

$$Z = \Sigma_i e^{-E_i/k_{\mathrm{B}}T} = e^{\mu_{\mathrm{B}}H/k_{\mathrm{B}}T} + e^{-\mu_{\mathrm{B}}H/k_{\mathrm{B}}T} = 2\cosh\left(\mu_{\mathrm{B}}H/k_{\mathrm{B}}T\right).$$
$$\langle \boldsymbol{M} \rangle = \frac{1}{Z}\Sigma_i \boldsymbol{m}_i e^{-E_i/k_{\mathrm{B}}T} = \frac{\mu_{\mathrm{B}}}{Z}\left(e^{\mu_{\mathrm{B}}H/k_{\mathrm{B}}T} - e^{-\mu_{\mathrm{B}}H/k_{\mathrm{B}}T}\right) = \mu_{\mathrm{B}}\tanh\left(\frac{\mu_{\mathrm{B}}H}{k_{\mathrm{B}}T}\right).$$

The total magnetization:

$$\boldsymbol{M} = n\mu_{\mathrm{B}}\tanh\left(\frac{\mu_{\mathrm{B}}H}{k_{\mathrm{B}}T}\right),$$

where n is the number of spins per unit volume.

Solution to (d)

For a given field, the magnetization decreases from $\mu(T = 0) = n\mu_{\mathrm{B}}$ to $\mu(T \to \infty) = 0$. For $n = 3.7 \cdot 10^{28}$ m^{-3}:

$$\boldsymbol{M}_{\mathrm{s}}(T = 0) = 3.7 \cdot 10^{28}\ \mathrm{m}^{-3} \cdot 9.274 \cdot 10^{-24}\ \mathrm{J\ T}^{-1} = 3.43 \cdot 10^{5}\ \mathrm{A/m}.$$

At zero temperature the spins are perfectly aligned by an external field because there is no thermal energy available to randomize the spin directions. At infinitely high temperature there is enough thermal energy to randomize the spin directions (giving a net magnetization of zero) even in the presence of an external magnetic field.

Solution to (e)

As $x \to 0, \tanh(x) \to x$; so as $\boldsymbol{H} \to 0, \tanh\left(\mu_{\mathrm{B}}H/k_{\mathrm{B}}T\right) \to \mu_{\mathrm{B}}H/k_{\mathrm{B}}T$, $M \to n\mu_{\mathrm{B}}^2 H/k_{\mathrm{B}}T$.

$$\chi = \frac{\boldsymbol{M}}{\boldsymbol{H}} = \frac{n\mu_{\mathrm{B}}^2}{k_{\mathrm{B}}}\frac{1}{T},$$

i.e. the susceptibility is $\sim 1/T$, and diverges only as $T \to 0$. This is Curie's law. At room temperature $T = 300$ K,

$$\chi = \frac{3.7 \cdot 10^{28}\ \mathrm{m}^{-3} \cdot \left(9.274 \cdot 10^{-24}\right)^2\ \mathrm{J}^2\ \mathrm{T}^{-2}}{1.381 \cdot 10^{-23}\ \mathrm{J\ K}^{-1} \cdot 300\ \mathrm{K}} = 768.11\frac{\mathrm{J}^3\ \mathrm{T}^{-2}}{\mathrm{m}^3}$$

or, multiplying by μ_0 to convert to dimensionless units, $\chi = 9.65 \cdot 10^{-4}$.

Solution to (f)

The behavior described by this non-interacting spin system is paramagnetic. The system shows Curie-law behavior, and there is no phase transition to a magnetically ordered state. In order to describe ferromagnetic behavior we would have to add interactions to our model. The interactions would have to be such that the energy of the system was lower when neighboring spins were aligned parallel to each other, compared to when they were not aligned at all, or had some other (for example, antiparallel) alignment.

(??? maybe later I'll add a better answer)

Spaldin.6.1(3 min) Molecular field for nickel

The Curie temperature of nickel, $T_{\mathrm{C}} = 628.3$ K, and the saturation magnetization is $0.6\mu_{\mathrm{B}}$ per atom. Calculate the molecular field according to the Weiss theory.

Solution

The solution is obvious, one just needs to understand Curie-Wiess therory.

(???? я что-то устал уже, не особо въезжаю, почему так???)

Equating, at the origin, the slope of the magnetization described by the Langevin function (which is $\frac{1}{3}\cdot\mathrm{Nm}$) with the slope of the straight line representing magnetization by the molecular field, gives

$$\frac{k_{\mathrm{B}}T_{\mathrm{C}}}{\boldsymbol{m}\gamma} = \frac{1}{3}\cdot N\boldsymbol{m}.$$

So, if the Curie temperature is known, then the molecular field constant can be extracted:

$$\gamma = \frac{3k_{\mathrm{B}}T_{\mathrm{C}}}{Nm^2}.$$

Similarly the Weiss molecular field $\boldsymbol{H}_{\mathrm{W}} = \gamma\boldsymbol{M} = \gamma N\boldsymbol{m} = 3k_{\mathrm{B}}T_{\mathrm{C}}/\boldsymbol{m}$. For Ni, the magnetic moment per atom $\boldsymbol{m} = 0.6\mu_{\mathrm{B}}$, and the Curie temperature $T_{\mathrm{C}} = 628.3$ K. Therefore

$$\boldsymbol{H}_{\mathrm{W}} = \frac{3k_{\mathrm{B}}T_{\mathrm{C}}}{\boldsymbol{m}} = \frac{3\cdot 1.380662\cdot 10^{-23}\ \mathrm{J\ K^{-1}}\cdot 628.3\ \mathrm{K}}{0.6\cdot 9.274\cdot 10^{-24}\ \mathrm{J\ T^{-1}}} = 4676.89\ \mathrm{T}.$$

This is a very large field!

Spaldin.1.3

A classical electron is moving in a circular orbit of radius 1Å $\left(1\text{Å} = 10^{-10}\ \mathrm{m}\right)$ with angular momentum $\hbar$Js.

(a) Calculate the field generated by the electron, at a distance of 3Å from the center of the orbit, and along its axis.

(b) Calculate the magnetic dipole moment of the electron, giving your answer in SI and cgs units.

(c) Calculate the magnetic dipolar energy of the circulating electron when it is in the field generated by a second identical circulating electron at a distance of 3Å away along its axis. Assume that the magnetic moment of the first electron is aligned parallel to the field from the second electron.

Solution

1.3(a) We use the expression which we derived in Solution 1.2(a) above:

$$\boldsymbol{H}_{\mathrm{axial}} = \frac{Ia^2}{2\left(a^2+x^2\right)^{3/2}},$$

with $a = 1\text{Å} = 10^{-10}$ m and $x = 3\text{Å} = 3\times 10^{-10}$ m. To calculate the current, I, we use the fact that the angular momentum (which in general is given by $m_{\mathrm{e}}va$) is $\hbar$Js. So

$$v = \frac{\hbar}{m_{\mathrm{e}}a}\frac{\mathrm{Js}}{\mathrm{kgm}} = \frac{\hbar}{m_{\mathrm{e}}a}\frac{\mathrm{m}}{\mathrm{s}}$$

and the current is

$$\begin{aligned} I &= \frac{\text{charge}}{\text{time}} = \frac{e}{\text{distance / speed}} = e\frac{v}{2\pi a} = \frac{e}{2\pi a}\frac{\hbar}{m_{\mathrm{e}}a} \\ &= 2.952\times 10^{-4}\ \mathrm{A}. \end{aligned}$$

Then

$$\boldsymbol{H} = \frac{2.952\times 10^{-4}\times\left(10^{-10}\right)^2}{2\left[\left(10^{-10}\right)^2+\left(3\times 10^{-10}\right)^2\right]^{3/2}}\frac{\mathrm{Am}^2}{\mathrm{m}^3} = 46675.7\ \mathrm{A/m} = 586\mathrm{Oe}.$$

1.3(b)

The magnetic dipole moment, m, is given by

$$\begin{aligned}\boldsymbol{m} = IA = \frac{ev}{2\pi a}\pi a^2 = \frac{eva}{2} = \frac{a}{2}\frac{e\hbar}{m_e a} = \frac{e\hbar}{2m_e} \\ = 9.274 \times 10^{-24} \text{ A m}^2 \text{ or J/T.}\end{aligned}$$

This number is the Bohr magneton, μ_B, and is the natural unit of magnetic moment. In cgs units, the Bohr magneton is equal to $e\hbar/2m_e c = 0.927\times 10^{-20}$erg/Oe. (Remember that the value of the elementary charge, e, in cgs units is 4.80×10^{-10}esu, and the value of h is 6.62×10^{-27}ergs.)

1.3(c)

The magnetic dipolar energy is

$$\begin{aligned}E &= -\mu_0 \boldsymbol{m} \cdot \boldsymbol{H} \\ &= 1.256 \times 10^{-6} \text{ weber } /(\text{Am}) \times -9.274 \times 10^{-24}\text{Am}^2 \times 46675.7 \text{ A/m} \\ &= -5.44 \times 10^{-25} \text{ J.}\end{aligned}$$

Note that this number is very small, so it is unlikely that the parallel alignment of magnetic dipole moments in ferromagnetic materials results from a magnetic dipolar interaction.

Spaldin.6.2

For an circulating electron in a 1Å radius orbit, at a distance of 3Å from the center of the orbit we can calculate its field and magnetic moment. These values are typical for an atom in a transition-metal crystal.
(a) To what Curie temperature would this field correspond?
(b) We can also calculate the dipole moment of the electron. To what extent would your calculated magnetic moment be affected by an external field of, say, 50Oe?
(Use $E = -\boldsymbol{m} \cdot \boldsymbol{H}$ and compare the result with the thermal energy, $k_B T$, at room temperature.)

Solution

6.2(a) In Chapter 1, we calculated that the field generated by such an electron was 46675.7 A/m, and that the magnetic moment was μ_B. Assuming a "classical" electron, so that $T_C = Nm^2\gamma/3k_B$, and taking $\gamma = \boldsymbol{H}/\boldsymbol{M} = \boldsymbol{H}/N\boldsymbol{m}$,

$$\begin{aligned}T_C &= \frac{mH}{3k_B} \\ &= \frac{9.274 \cdot 10^{-24} \text{ J T}^{-1} \cdot 46675.7 \text{ A m}^{-1}}{3 \cdot 1.380662 \cdot 10^{-23} \text{ J K}^{-1}} \\ &= 10450.794 \text{ A K/(Tm)} \\ &= 0.0131 \text{ K}\end{aligned}$$

(multiplying by $\mu_0 = 1.25 \cdot 10^{-6}\text{Hm}^{-1}$, to convert to kelvin). Note that this is a very small number!
6.2(b) In a field of 50 Oe the magnetic dipole energy would be

$$\begin{aligned}E &= -\mu_0 \boldsymbol{m} \cdot \boldsymbol{H} \\ &= -9.274 \cdot 10^{-24} \text{ J T}^{-1} \cdot \left(50 \cdot \frac{1000}{4\pi}\right) \text{Am}^{-1} \cdot 1.25 \cdot 10^{-6}\text{Hm}^{-1} \\ &= -4.637 \cdot 10^{-26} \text{ J.}\end{aligned}$$

At 298 K the thermal energy $k_B T = 4.11 \cdot 10^{-21}$ J, which is five orders of magnitude larger than the magnetic energy! Therefore a field of around 50 Oe would have no effect on aligning electronic magnetic moments at room temperature. We can conclude that the effective internal "field" which aligns the magnetic moments of ferromagnets spontaneously is much larger than 50 Oe.

Spaldin.6.3 Review question

(a) Using Ampère's circuital law or the Biot-Savart law, make an order-of-magnitude estimate of the size of the magnetic field generated by the valence electrons in a Ni atom at a distance corresponding to the Ni − Ni spacing in a solid sample of Ni. (Assume that the field arises from the current generated by the circulation of unpaired electrons around the nucleus.)

(b) Use Hund's rules to determine the values of S, L, and J for an isolated Ni atom with electronic structure $(4\text{ s})^2(3\text{ d})^8$. What are the allowed values of the magnetic moment along the field axis for a Ni atom?

(c) Use your answers to parts (a) and (b) to estimate the difference in magnetic dipole energy between Ni atoms aligned parallel and antiparallel to each other.

(d) Given that the Curie temperature of Ni is 358°C, how does the magnetic dipole energy which you calculated in (c) compare with the actual strength of the ferromagnetic coupling between Ni atoms?

(e) Explain briefly (a few sentences) what the true origin of the ferromagnetic coupling is in Ni.

(f) The actual value of the magnetic dipole moment in metallic Ni is $0.54\mu_{\mathrm{B}}$. Why does this number not correspond to an integer number of electrons? (You'll probably need a diagram to explain this one!)

Solution to (a)

(a) For this problem it's much easier to use the Biot-Savart law, as we did in Solution 1.2(a). This gives us the following expression for the magnetic field at a distance x from a current-carrying circular loop, on the axis of the loop:

$$\begin{aligned}\boldsymbol{H} &= \frac{I}{2a}\sin^3\alpha \\ &= \frac{Ia^2}{2\left(a^2+x^2\right)^{3/2}}.\end{aligned}$$

To estimate the field generated by a Ni atom at its neighbor in a solid, let's pretend that the electrons in the Ni atom are orbiting around the nucleus with a radius $a = 1\text{Å}$, and that the neighboring Ni atom is a distance $x = 3\text{Å}$ from the first Ni atom.

To estimate the current I, we use the fact that the angular momentum of an electron (which in general is given by $m_e v a$) is of the order of $\hbar$J. So

$$v = \frac{\hbar}{m_e a}\frac{\text{Js}}{\text{kgm}} = \frac{\hbar}{m_e a}\frac{\text{m}}{\text{s}}$$

and the current is

$$\begin{aligned}I &= \frac{\text{charge}}{\text{time}} \\ &= \frac{e}{\text{distance/speed}} \\ &= e\frac{v}{2\pi a} \\ &= \frac{e}{2\pi a}\frac{\hbar}{m_e a} \\ &= 2.952\cdot 10^{-4}\text{ A}.\end{aligned}$$

Since there are two unpaired electrons in Ni we can double that number if we like, but since we're just looking for an order-of-magnitude estimate it doesn't really matter either way.

Solution to (b)

Solution to (c)

Solution to (d)

Solution to (e)

Solution to (f)

Spaldin.8.1 Parallel susceptibility

Show that Eq. $\chi_{\parallel} = \frac{2Nm^2B'(J,\alpha)}{2k_{\mathrm{B}}T+Nm^2\gamma B'(J,\alpha)}$, describing the susceptibility when the field is applied parallel to the magnetization direction, reduces to the Curie-Weiss expression (Eq. $\chi = \frac{\boldsymbol{M}}{\boldsymbol{H}} = \frac{C}{T+\theta}$) at high temperatures, and to zero at 0 K.

Solution

The susceptibility of an antiferromagnetic material in which the field is applied parallel to the magnetization direction is given by

$$\chi_{\parallel} = \frac{2Nm^2B'(J,\alpha)}{2k_{\mathrm{B}}T + Nm^2\gamma B'(J,\alpha)}.$$

Here $B'(J,\alpha)$ is the derivative with respect to α of the Brillouin function,

$$B_J(\alpha) = \frac{2J+1}{2J}\coth\left(\frac{2J+1}{2J}\alpha\right) - \frac{1}{2J}\coth\left(\frac{\alpha}{2J}\right),$$

and $\alpha = Jg\mu_{\mathrm{B}}H/k_{\mathrm{B}}T$.

At high temperature, α is very small; therefore we can expand the Brillouin function in a Taylor series about the origin to obtain

$$B_J(\alpha) = \frac{J+1}{3J}\alpha - \frac{\left[(J+1)^2 + J^2\right](J+1)}{90J^3}\alpha^3 + \cdots.$$

So, at small $\alpha, B'(J,\alpha) = (J+1)/3J$, which is a constant for a given J. Then

$$\begin{aligned}\chi_{\parallel} &= \frac{2Nm^2(J+1)/3J}{2k_{\mathrm{B}}T + Nm^2\gamma(J+1)/3J} \\ &= \frac{C}{T+\theta}\end{aligned}$$

- the Curie-Weiss law!

At low temperature, α is large and we cannot use the expansion of the Brillouin function given above. Instead, using the fact that $d(\coth\alpha)/d\alpha = -\operatorname{cosech}^2\alpha$, we differentiate the Brillouin function explicitly to obtain

$$B'(J,\alpha) = -\left(\frac{2J+1}{2J}\right)^2\operatorname{cosech}^2\left(\frac{2J+1}{2J}\alpha\right) + \left(\frac{1}{2J}\right)^2\operatorname{cosech}^2\left(\frac{\alpha}{2J}\right).$$

As $\alpha \to \infty, \operatorname{cosech}(\alpha) \to 0$, and $B'(J,\alpha) \to 0$. Therefore $\chi_{\parallel}$also tends to zero at low temperature.

Spaldin.8.2

Consider an antiferromagnetic material which has a susceptibility, χ_0, at its Néel temperature, T_{N}. Assuming that the exchange interactions between nearest-neighbor A and B ions are much larger than those between $A-A$ and $B-B$ pairs, calculate the values of the susceptibilities which would be measured under the application of fields perpendicular to the magnetization direction at $T = 0, T = T_{\mathrm{N}}/2$, and $T = 2T_{\mathrm{N}}$.

Solution

8.2 Since the $A-B$ interaction is much stronger than the $A-A$ and $B-B$ interactions, we can use the results which we derived in the notes using the Langevin-Weiss theory. We know that the expression for the susceptibility at and above the Néel temperature is

$$\chi = \frac{C}{T+\theta} = \frac{C}{T+T_{\mathrm{N}}}.$$

Leuven-exam-2022 (!!!)

a) Calculate the magnetic susceptibility of GaN (SI units), assuming that there are no free carriers and no local magnetic moments.

b) Now consider that 10% of the Ga atoms were replaced by Mn atoms (assume, for simplicity, that the number of atoms per unit volume does not change). Assume that Mn atoms are in a 3+ charge state $\left(3d^4\right)$, that you can apply the Russel-Saunders coupling scheme, and that there is no magnetic interaction between the Mn atoms. Calculate the paramagnetic susceptibility at 300 K (SI units) associated with the local Mn moments, still assuming that there are no free carriers, for two different scenarios:

i) without orbital quenching

ii) with orbital quenching

c) Calculate the total magnetic susceptibility of the system described in (b), at 300 K, for scenario (ii). Convert that value to mass susceptibility (in $\mathrm{m^3Kg^{-1}}$). Convert both quantities to the cgs unit system.

d) Now consider that the system in (b) orders ferromagnetically with a T_C of 10 K, with the moment per Mn atom being the same as in the paramagnetic regime (calculated in (b)). Estimate the magnitude of the molecular field in units of Tesla.

Some constants and parameters that you may need to use, but not necessarily all:

$$\rho(\mathrm{GaN}) = 6150 \text{ kg m}^{-3}$$
$$\text{Avogadro constant } = 6.02 \times 10^{23} \text{ mol}^{-1}$$
$$\text{relative atomic mass of Ga is 69.7 and of N is 14.0}$$
$$e = 1.6 \times 10^{-19}\mathrm{C}$$
$$\mu_0 = 4\pi \times 10^{-7} \text{ N A}^{-2}$$
$$m_e = 9.11 \times 10^{-31} \text{ kg}$$
$$Z(\mathrm{Ga}) = 31; Z(\text{ N}) = 7; Z(\mathrm{Mn}) = 25$$
$$\left\langle r^2 \right\rangle_{\mathrm{av}} (\mathrm{Ga}) = 1.69 \times 10^{-20} \text{ m}^2$$
$$\left\langle r^2 \right\rangle_{\mathrm{av}} (\mathrm{N}) = 4.22 \times 10^{-21} \text{ m}^2$$
$$\left\langle \boldsymbol{r}^2 \right\rangle_{\mathrm{av}} (\mathrm{Mn}) = 2.56 \times 10^{-20} \text{ m}^2$$
$$\mu_{\mathrm{B}} = 9.27 \times 10^{-24} \text{ J T}$$
$$k_{\mathrm{B}} = 1.38 \times 10^{-23} \text{ m}^2 \text{ kg s}^{-2} \text{ K}^{-1}$$

Solution

(!! пропишу, я очень похожую решал во время экзамена, просто не до этого было пока что)

12.2 Problems about magnetic field

12.2.1 Problems in magnetostatics

(порешаю задачи - пропишу)

13 Problems in spin models

13.0.1 Problems about AKLT or similar

Любошин-ПТФ-спецкурс-2022-5/7-1/3.

Рассмотрим стандартный антиферромагнитный одномерный гейзенберговский гамильтониан для спина S

$$H = |J| \sum_{k=1}^{N} \mathbf{S}_k \mathbf{S}_{k+1}$$

с периодическими граничными условиями $(N+1=1)$. а) Найдите спектр H для $N = 2, 3, 4$. б)* Найдите энергии основного и первого возбужденного состояний для $S = 1/2$ и $S = 1$ при $N = 5$, при необходимости - численно. в) Попробуйте экстраполировать полученные в пунктах а) и б) ответы для энергии основного состояния в расчете на одну связь E_0/N и для щели $\Delta = E_1 - E_0$ в область $N \gg 1$. Видно ли качественное различие в поведении Δ при $S = 1/2$ и $S = 1$?

Любошин-ПТФ-спецкурс-2022-5/7-2/3.

Пусть ψ_N - основное состояние модели AKLT$(S = 1)$

$$H = |J| \sum_{k=1}^{N} \mathbf{S}_k \mathbf{S}_{k+1} + \frac{1}{3} \left(\mathbf{S}_k \mathbf{S}_{k+1} \right)^2$$

с периодическими граничными условиями $(N+1=1)$. а) Выведите формулу (4) статьи AKLT для перекрытия волновых функций VBS-состояний с открытыми граничными условиями. б) Получите из формулы пункта а) основной результат статьи - выражение для спинового коррелятора в ψ_N,

$$\left\langle \psi_N \left| S_0^a S_k^b \right| \psi_N \right\rangle \to \frac{4}{3}(-1)^k 3^{-k} \delta^{ab}, \quad N \to \infty; \quad a, b \in \{x, y, z\}.$$

в) Совпадает ли ψ_N с основным состоянием модели из задачи 1 при $N = 2, 3, 4, 5$? Какова энергия в расчете на одну связь $e_N = \langle\psi_N|H|\psi_N\rangle / N$ состояния ψ_N как вариационной волновой функции модели Гейзенберга в этих случаях? Из результата пункта б) следует, что $e_N \to -\frac{4}{3}|J|$ при $N \to \infty$. Какую следует ожидать скорость сходимости по N?

Любошин-ПТФ-спецкурс-2022-5/7-**3/3.**

Рассмотрим модель Majumdar-Ghosh ($S = 1/2$)

$$H = |J| \sum_{k=1}^{N} \mathbf{S}_k \mathbf{S}_{k+1} + \frac{1}{2} \mathbf{S}_k \mathbf{S}_{k+2}$$

с периодическими граничными условиями ($N+1 = 1$). а) Покажите, что при четном N любое основное состояние будет линейной комбинацией двух полностью димеризованных состояний. б)* Как обстоит дело при нечетном N?

Part V
Other Topics

14 Other phenomena and properties

14.1 Other, General Questions

14.1.1 How Long Can a Magnet Hold a Load?

Там энергия как уходит на это??? Неужели от совершения работы спины переворачиваются обратно в неупорядоченное состояние? Как тогда они знают, что поле совершило работу?

Об этом всем тут подробно.

14.1.2 Critical Entropy of Quantum Heisenberg Magnets on Simple-Cubic Lattices by Stefan Wessel

Abstract

We analyze the temperature dependence of the entropy of the spin-1/2 Heisenberg model on the three-dimensional simple-cubic lattice, for both the case of antiferromagnetic and ferromagnetic nearest neighbor exchange interactions. Using optimized extended ensemble quantum Monte Carlo simulations, we extract the entropy at the critical temperature for magnetic order from a finite-size scaling analysis. For the antiferromagnetic case, the critical entropy density equals $0.341(5)k_B$, whereas for the ferromagnet, a larger value of $0.401(5)k_B$ is obtained. We compare our simulation results to estimates put forward recently in studies assessing means of realizing the antiferromagnetic Néel state in ultra-cold fermion gases in optical lattices.

Introduction

Ultra-cold atom gases provide a unique opportunity to study basic models of strongly interacting many-body systems in a highly controlled experimental framework [1]. After the pioneering realization of the superfluid-to-Mott insulator transition of ultra-cold bosonic atoms [2], recently different groups reported the realization of a Mott-insulating state also for ultra-cold ^{40}K fermionic atoms on three-dimensional simple-cubic lattices [3, 4]. A next major step would be the realization and identification of a low-temperature antiferromagnetically ordered Néel state in such systems. In fact, various recent studies explored the prospects of realizing the antiferromagnetic Néel state in an ultra-cold gas of fermionic atoms confined to an optical lattice [5, 7, 8, 6, 9, 10, 11, 12]. A quantity that is important in assessing the relevant temperature scale is the upper bound on the entropy allowed in order to transfer the atoms into the antiferromagnetically ordered state. For a three-dimensional fermionic Hubbard model on a simple-cubic lattice at half-filling, estimates of the entropy S as a function of temperature, as well as the ratio U/t between the onsite repulsion U and the hopping amplitude t, were obtained within a single-site dynamical mean-field theory (DMFT) approach [5]. As reported in Ref. [5], DMFT however overestimates the Néel temperature by about 50% in the intermediate coupling regime, as compared to direct quantum Monte Carlo simulations based on systems with up to 10^3 lattice sites [13]. Obtaining in a similar way the entropy from direct quantum Monte Carlo simulations is challenging, and typically involves integration of the specific heat over an extended temperature range.

However, in the limit of large U/t, the spin physics of the Hubbard model is well known to be described by a nearest-neighbor spin-1/2 quantum Heisenberg model, with an antiferromagnetic exchange coupling $J = 4t^2/U > 0$, obtained within perturbation theory around the strong coupling limit $t = 0$. This model is accessible to large scale quantum Monte Carlo simulations and moreover – as shown below – it is possible to provide precise values of the entropy in this large-U limit. In particular, one can obtain the value of the critical entropy density (i.e. the entropy per lattice site) S_C, below which antiferromagnetic order emerges. Nevertheless, thus far no systematic quantum Monte Carlo study concerning the entropy has been reported. On the other hand, different estimates of S_C for the Heisenberg antiferromagnet have been put forward recently. In Ref. [5], a Schwinger boson approach [14] was employed, leading to the estimate that S_C is about 50% of the mean-field value $S_C^{MF} = k_B \ln(2)$. A more recent study [6] reported a value of S_C obtained from a fluctuation-corrected mean-field approach, leading to a reduction in S_C of only 20% from the mean-field value S_C^{MF}. This rather large discrepancy among current estimates of S_C calls for a clarification based on numerically exact simulations of the Heisenberg model.

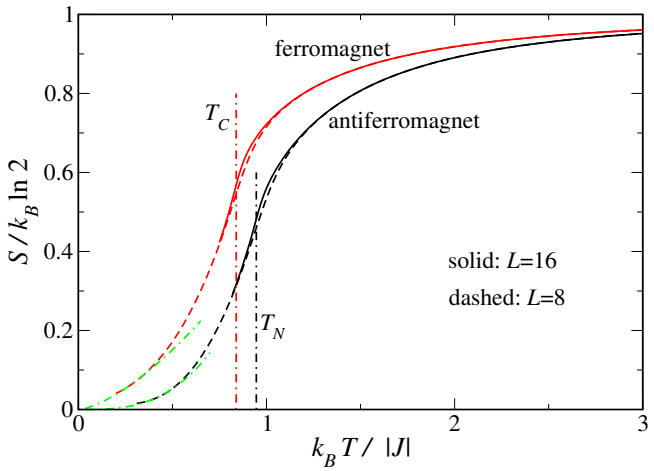


Figure 325: Temperature dependence of the entropy of the spin-1/2 quantum Heisenberg antiferromagnet and ferromagnet on the simple-cubic lattice. For both cases, results from quantum Monte Carlo simulations are shown for finite systems with $L = 8$ (dashed lines) and $L = 16$ (solid lines). Error bars for the shown data are below the line thickness. Vertical lines show the transition temperatures T_N and T_C for the antiferromagnet and the ferromagnet, respectively. Green dashed-dotted curves denote fits to the algebraic low-T scaling $S \propto T^3$ ($S \propto T^{3/2}$) for the antiferromagnet (ferromagnet).

Here, we present an unbiased estimate of the temperature dependence of the entropy and in particular the critical entropy density by combining extended ensemble quantum Monte Carlo simulations with a finite-size scaling analysis based on the known critical properties of the quantum Heisenberg model. For comparison, we consider in the following both the ferromagnetic and the antiferromagnetic spin-1/2 Heisenberg model on the three-dimensional simple-cubic lattice. These models are described by a Hamiltonian

$$H = J \sum_{i,j} \mathbf{S}_i \cdot \mathbf{S}_j, \tag{1}$$

where $\mathbf{S}_i$ denotes a spin-1/2 degree of freedom on lattice site i, and J is nearest-neighbor exchange interaction, with $J > 0$ ($J < 0$), for the antiferromagnetic (ferromagnetic) case. From previous studies, the locations of the transition temperatures to the ordered phases have

been determined as $k_B T_N/J = 0.946(1)$ [15] (Néel temperature) for the antiferromagnet, and $k_B T_C/|J| = 0.839(1)$ [16] (Curie temperature) for the ferromagnet, based on quantum Monte Carlo simulations.

Method

In order to extract the temperature dependence of the entropy, we use an optimized extended ensemble approach [17, 18], that is based on a generalization of the Wang-Landau [19] algorithm to the case of quantum Monte Carlo simulations [20, 16], performed within the stochastic series expansion representation [21] using multi-cluster deterministic loop updates [22]. Within this approach, one obtains Monte Carlo estimates of the expansion coefficients $g(n)$ of the high-temperature series expansion of the partition function Z in the inverse temperature $\beta = 1/(k_B T)$,

$$Z = \mathrm{Tr} e^{-\beta H} = \sum_{n \geq 0} g(n) \beta^n, \tag{2}$$

for a given system of $N = L^3$ lattice sites. Here, L denotes the linear size of the simulated finite cube, and we employ periodic boundary conditions in all directions. From the expansion coefficients $g(n)$, the free energy

$$F = -\frac{1}{\beta} \ln Z = -\frac{1}{\beta} \ln \sum_{n \geq 0} g(n) \beta^n, \tag{3}$$

the internal energy

$$E = \frac{1}{Z} \sum_{n \geq 0} n \, g(n) \beta^{n-1}, \tag{4}$$

and the entropy

$$S = \frac{E - F}{T} \tag{5}$$

are obtained as continuous functions of the temperature T. In practice, as discussed in Ref. [20], we obtain the expansion coefficients up to a upper cutoff Λ, that is chosen sufficiently large in order to reliably calculate the thermodynamic properties of the finite system based on the coefficients $g(n)$, $n = 0, ..., \Lambda$ down to a given temperature scale β_Λ. The required Λ scales linear in both β_Λ and the system size N. In our simulations, we consider finite systems up to $L = 20$, for which $\Lambda = 16000$ is required in order to obtain thermodynamic properties down to a temperature of about $k_B T \approx 0.7|J|$. Repeated simulations with independent random streams results in statistical error bars on the thermodynamic quantities.

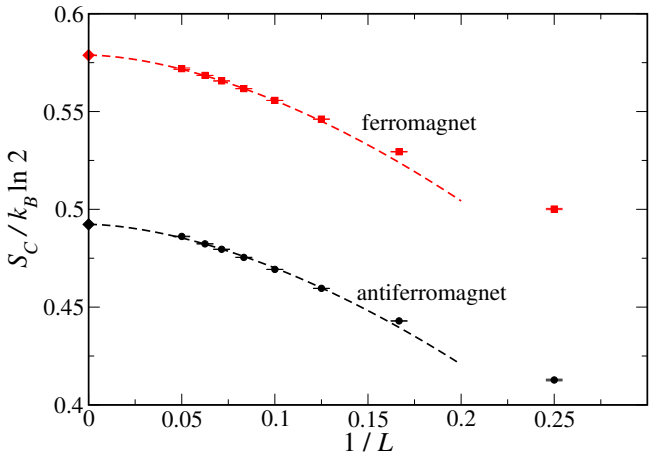


Figure 326: Finite-size scaling of the critical entropy density $S_C(L)$ with the inverse linear system size $1/L$ for the spin-1/2 quantum Heisenberg antiferromagnet and ferromagnet on the simple-cubic lattice. The quantum Monte Carlo data are shown by symbols, and the dashed lines represent fits of the finite-size data for $L > 6$ to the finite-size scaling form in Eq. (6). The results for S_C in the thermodynamic limit are indicated by symbols as well.

Results

In Fig. 1, we present the results for the entropy from our simulations for both the antiferromagnetic and the ferromagnetic Heisenberg model. For both cases, we show data for systems with $L = 8$ and $L = 16$. Outside a narrow range around the transition temperature, the entropy curves obtained for the two different system sizes compare well, and thus essentially represent the behavior of the entropy density in the thermodynamic limit. Also shown in Fig. 1 are fits of the numerical data to the low-T scaling $S \propto T^3$ and $S \propto T^{3/2}$ for the antiferromagnet and ferromagnet, respectively. This difference in the low-T scaling behavior relates to the linear (quadratic) dispersion of the magnetic spin-wave excitations in the antiferromagnetic (ferromagnetic) case [14].

However, in order to extract the critical entropy S_C in the thermodynamic limit, we need to perform a finite-size scaling analysis of the finite-size data $S_C(L)$ taken at the transition temperature. For this purpose, we use a finite-size scaling ansatz

$$S_C - S_C(L) \propto L^{1/\nu - d}, \tag{6}$$

where ν is the critical exponent of the correlation length, and d the dimensionality of the system. Starting from $S = -\partial F/\partial T$, this relation follows from the scaling behavior of the free energy $F \propto \xi^{-d}$ with the correlation length ξ, that scales as $\xi \propto |t|^{-\nu}$ when the reduced temperature $t \to 0$, by the usual finite-size scaling ansatz, that at criticality ξ relates to the linear system size L. For the current case of a three-dimensional system ($d = 3$) with $O(3)$ symmetry, $\nu = 0.7112(5)$ [23]. Figure 2 shows that the finite-size data indeed fits well to the above scaling form in the range of system sizes $L > 6$. From the extrapolated values, we obtain $S_C = 0.341(5)k_B = 0.492(7)S_C^{MF}$ for the antiferromagntic model, whereas for the ferromagnet, a larger value of $0.401(5)k_B = 0.578(7)S_C^{MF}$ results. We also performed an algebraic fit without biasing the exponent, which returned identical values of S_C within the statistical error bars.

While the ferromagnet has a critical temperature T_C that is about 11% *lower* than T_N, it orders already below an entropy that is about 17% *higher* than the critical entropy of the antiferromagnet. The higher critical entropy directly exhibits a reduced efficiency of thermal disorder in destroying long-range magnetic order. The lower transition temperature of the ferromagnetic is due to a significantly enhanced production of entropy upon heating up the

system, which follows from both Fig. 1, as well as from the mentioned difference in the low-T behavior of the entropy.

Conclusions

Based on numerically exact quantum Monte Carlo simulations, we obtained the temperature dependence of the entropy for both the antiferromagnetic and the ferromagnetic spin-1/2 Heisenberg model on the simple-cubic lattice. Focusing on the region close to the magnetic ordering transition, we employed a finite-size scaling analysis to extract the critical entropy density below which long-ranged magnetic order sets in. In case of the antiferromagnetic model, our result for the critical entropy is consistent with a 50% reduction from the mean-field value, that was estimated recently from a Schwinger boson calculation[5]. This clarifies among the different values of the critical entropy reported previously[5, 6]. In comparison, we found the ferromagnet to be more robust to thermal disorder, which at criticality is about 17% larger for the ferromagnet than for the antiferromagnet in terms of the critical entropy. We consider the current results for the temperature dependence of the magnetic entropy of value also for future calculations of the entropy for the full fermionic Hubbard model beyond the strongly interacting regime.

14.2 Topological magentic models

(будет полезно - напишу в 1й части, пока не думаю, что что-то полезное тут можно придумать.)

Spin coherent states

(Тут https://www.youtube.com/watch?v=_4pb48GbzNo&t=2046s послушал лекцию, но не готов еще всерьез много про эти темы думать. потом подумаю больше когда-то)

Spin-coherent path integrals

Topology in the Heisenberg antiferromagnet

15 From Nagaoka's ferromagnetism to flat-band ferromagnetism and beyond by Tasaki

From Nagaoka's ferromagnetism to flat-band ferromagnetism and beyond

15.1 Introduction

Abstract

It is believed that strong ferromagnetic interactions in some solids are generated by subtle interplay between quantum many-body effects and spin-independent Coulomb interactions between electrons. It is a challenging problem to verify this scenario in the Hubbard model, which is an idealized model for strongly interacting electrons in a solid.

Nagaoka's ferromagnetism is a well-known rigorous example of ferromagnetism in the Hubbard model. It deals with the limiting situation in which there is one fewer electron than in the half-filling and the on-site Coulomb interaction is infinitely large. There are relatively new rigorous examples of ferromagnetism in Hubbard models called flat-band ferromagnetism. Flat-band ferromagnetism takes place in carefully prepared models in which the lowest bands (in the single-electron spectra) are "flat." Usually, these two

approaches are regarded as two complimentary routes to ferromagnetism in the Hubbard model.

In the present paper we describe Nagaoka's ferromagnetism and flat-band ferromagnetism in detail, giving all the necessary background as well as complete (but elementary) mathematical proofs. By studying an intermediate model called the long-range hopping model, we also demonstrate that there is indeed a deep relation between these two seemingly different approaches to ferromagnetism.

We further discuss some attempts to go beyond these approaches. We briefly discuss recent rigorous example of ferromagnetism in the Hubbard model which has neither infinitely large parameters nor completely flat bands. We give preliminary discussion regarding possible experimental realizations of the (nearly-)flat-band ferromagnetism. Finally, we focus on some theoretical attempts to understand metallic ferromagnetism. We discuss three artificial one-dimensional models in which the existence of metallic ferromagnetism can be easily proved.

We have tried to make the present paper as self-contained as possible, keeping in mind readers who are new to the field. Although the present paper is written as a review, it contains some material which appears for the first time.

15.1.1 Hubbard model and the origin of ferromagnetism

The origin of strong ferromagnetic ordering observed in some materials has been a mystery in physical science for a long time. Since quantum mechanical many-electron systems without interactions universally exhibit paramagnetism, the origin of ferromagnetism should be sought in electron-electron interactions. In most solids, however, the dominant part of the interaction between electrons is the Coulomb interaction, which is perfectly spin-independent. We are thus faced with a very interesting and fundamental problem in theoretical physics to determine *whether spin-independent interactions in an itinerant electron system can be the origin of ferromagnetic ordering.* This problem is important not only because ferromagnetism is a very common (and useful) phenomenon, but because it focuses on the fundamental role of nonlinear interactions in many-body quantum mechanical systems.

It was Heisenberg [1] who first realized that ferromagnetism is an intrinsically quantum mechanical phenomenon. In Heisenberg's approach to ferromagnetism, one starts from the picture that each electron (relevant to magnetism) is almost localized in an atomic orbit. By treating the effect of the Coulomb interaction and overlap between nearby atomic orbits in a perturbative manner as in the Heitler-London theory, Heisenberg concluded that there appears an "exchange interaction" between nearby electronic spins which determines the magnetic properties of the system.

In a different approach to the problem of ferromagnetism, which was originated by Bloch [2], one starts from the quantum mechanical free electron gas, in which electrons are in plane-wave like states. One then treats the effect of the Coulomb interaction perturbatively, and tries to find instabilities against certain magnetic ordering. When combined with the Hartree-Fock approximation (or a mean-field theory), this approach leads to the picture that there is an instability with respect to ferromagnetism when the density of states at the fermi energy and the Coulomb interaction are sufficiently large.

In spite of a considerable number of attempts to improve these ideas, neither of these two approaches has yet produced a truly convincing explanation about the origin of ferromagnetism.

A modern version of the problem of the origin of ferromagnetism was formulated by Kanamori [3], Gutzwiller [4], and Hubbard [5] in the 1960's[1]. They studied simple tight-binding models of electrons with on-site Coulomb interaction. This model is usually

[1] A similar formulation was given earlier, for example, in [6].

called the 'Hubbard model.' When there is no electron-electron interaction, the model exhibits paramagnetism as an inevitable consequence of the Pauli exclusion principle. Among other things, Kanamori, Gutzwiller, and Hubbard asked *whether the paramagnetism found for a non-interacting system can be converted into ferromagnetism when there is a sufficiently large Coulomb interaction.* This is a concrete formulation of the fundamental problem that we alluded to in the opening of the previous subsection.

It is worth noting that the on-site Coulomb interaction itself is completely independent of electronic spins, and it does not favor any magnetic ordering. One does not find any terms in the Hubbard Hamiltonian which explicitly favor ferromagnetism (or any other ordering). Our theoretical goal is to show that magnetic ordering arises as a consequence of the subtle interplay between the kinetic motion of electrons and the short-ranged Coulomb interaction. It is interesting to compare this situation with that in spin systems, where one is often given a Hamiltonian which favors some kind of magnetic ordering, and the major task is to investigate if such ordering really takes place. We can say that the Hubbard model formulation goes deeper into fundamental mechanisms of magnetism than that of spin systems. It offers a challenging problem to theoretical physicists to derive magnetic interaction from models which do not explicitly contain such interactions.

From a more global point of view, the importance of the Hubbard model may be understood from the philosophy of "universality", which, in our opinion, is at the very heart of contemporary physics. We believe that nontrivial physical phenomena or mechanisms found in a suitable idealized model can also be found in other systems in the same "universality class" as the idealized model. We expect that the universality class is often large and rich enough to contain various realistic systems with complicated details which are ignored in the idealized model. As for strongly interacting electron systems, the Hubbard model is regarded as one of the most promising candidates for an idealized model to be used in our search of possible universality classes.

Perhaps the best justification of the Hubbard model as a standard model of itinerant electron systems comes from such theoretical considerations, rather than its accuracy in modeling narrow band electron systems.

15.1.2 Rigorous results

The problem of ferromagnetism in the Hubbard model was studied extensively using various heuristic methods. The Hartree-Fock approximation discussed above leads one to the so called Stoner criterion. It states that the Hubbard model exhibits ferromagnetism when the condition $UD_{\mathrm{F}} > 1$ is satisfied, where U is the the strength of the on-site Coulomb interaction and D_{F} is the density of states of the corresponding single-electron problem measured at the fermi level (of the corresponding non-interacting system). Although the criterion cannot be trusted literally[2], it guides us to look for ferromagnetism in models in which U is large and/or the density of states is large.

The first rigorous result about ferromagnetism in the Hubbard model, which is one of the main topics of the present paper, was provided by Nagaoka[3] [10]. It was proved that certain Hubbard models have ground states with saturated ferromagnetism when there is exactly one hole and the Coulomb repulsion U is infinite.

In 1989, Lieb proved an important general theorem for the Hubbard model on a bipartite lattice at half-filling [11]. As a corollary of this theorem, Lieb showed that a rather general

[2]There are many systems which fully satisfy the criterion but do not exhibit ferromagnetism. Flat-band Hubbard models with low electron densities [7, 8] are typical examples.

[3]Thouless [9] reached a similar but slightly weaker conclusion.

class of Hubbard model exhibits ferrimagnetism[4]. See also [12].

In 1991, Mielke [13, 14] found a new class of rigorous examples of ferromagnetism in the Hubbard model. He showed that the Hubbard models on a general class of line graphs have ferromagnetic ground states. A special feature of Mielke's models is that the corresponding single-electron Schrödinger equation has highly degenerate ground states. In other words, Mielke's models have flat (or dispersionless) bands. Mielke' original results were for the case in which the number of electrons corresponds to the half-filling of the lowest flat band, but later it was extended to different electron densities in two-dimensional models [15].

A similar but different class of examples of ferromagnetism in Hubbard models, which we will discuss in detail in the present paper, were proposed in [7, 8]. These models are defined on a class of decorated lattices with "cell structures", and are also characterized by flat bands at the bottom of the single-electron spectrum. In a class of models in two and higher dimensions, it was proved that the ferromagnetism is stable against fluctuations in the electron number [7, 8]. The ferromagnetism in Mielke's models and those in [7, 8] is now called "flat-band ferromagnetism", and regarded as one of reliable starting points for the problem of ferromagnetism.

There are also rigorous results for ferromagnetism in 'nearly-flat-band models' obtained by perturbing the flat-band models of [7, 8]. In a general situation, local stability of ferromagnetism is known [16, 17], and for a special class of models, global stability of ferromagnetism has been established [18, 19]. We shall also briefly discuss the latter results in the present paper.

Nagaoka's ferromagnetism takes place in models with $U \to \infty$, while (nearly-)flat-band ferromagnetism takes place in models characterized by large (or infinite) density of states $D_{\rm F}$. One might say that these two rigorous results realize the Stoner criterion $UD_{\rm F} > 1$ through complementary paths.

15.1.3 About the present paper

One of the main aims of the present paper is to give complete descriptions of Nagaoka's ferromagnetism and flat-band ferromagnetism of [7, 8]. In particular, the proof of the generalized version of Nagaoka's theorem is described in complete detail for the first time. Another important aim is to show that there is a close relation between these two approaches which are usually regarded as complementary. We demonstrate this fact by studying an artificial model which we call the 'long range hopping model.' Our hope is that a clarification of the relation between the two approaches will lead us to a more global view of ferromagnetism in the Hubbard model, and that this might lead us in the long run to a better understanding of the essence of the fascinating phenomenon of ferromagnetism.

We also discuss three topics related to these results. First, we briefly discuss important attempts to go beyond flat-band ferromagnetism by treating non-singular Hubbard models. One of the main achievements in this direction is a proof of the existence of ferromagnetism in a Hubbard model which has a finite Coulomb repulsion U and a finite density of states $D_{\rm F}$. Next, we give a preliminary discussion of some experimental results which may be relevant to flat-band ferromagnetism. The ferromagnetism seen in $La_4Ba_2Cu_2O_{10}$, which was discovered back in 1990, shows some striking similarities with flat-band ferromagnetism. Finally we describe some attempts to obtain theoretical examples of metallic ferromagnetism. There are strong indications that some one-dimensional Hubbard models exhibit metallic ferromagnetism, but there are no rigorous results yet. We discuss three artificial models in one-dimension which are easily shown to exhibit metallic ferromagnetism. The results for the third model, the limiting U-V model, are perhaps new.

[4]Ferrimagnetism is a kind of antiferromagnetism on a bipartite lattice such that the numbers of sites in two sublattices are different.

We have tried to make the present paper as self-contained as possible[5]. We carefully start from basic assumptions and definition of the Hubbard model, keeping in mind readers who are new to the field. We give complete proofs to all the theorems which are directly related to the main subjects of the paper. The proofs, however, may not be optimally organized from a mathematical point of view. Instead, we have tried to present "readable" proofs from which the readers can learn physical mechanisms underlying the theorems. (This comment applies even to the proof of purely mathematical theorems such as Theorem 15.11.)

The present paper is organized as follows. In Section 15.2, we give a complete definition of the Hubbard model. We start from the description of the single-electron problem, and proceed by defining fermion operators, many-body Hilbert space, and the Hubbard Hamiltonian. An expert can safely skip this entire section, provided that he/she takes a brief look at Sections 15.2.1 and 15.2.3 to note our notation in the coordinate-free formalism of fermion operators.

Section 15.3 is also introductory and standard. We define the spin angular momenta of the model, and give a precise definition (Definition 15.4) of what we mean by ferromagnetism. We also present some results which rule out the possibility of ferromagnetism under several conditions.

In Section 15.4, we present a complete description and a proof of Nagaoka's ferromagnetism in its most generalized form.

In Section 15.5, we introduce and discuss an artificial model that we call the 'long-range hopping model.' Ferromagnetism in this model is first regarded as a special case of Nagaoka's ferromagnetism, but a different point of view is presented. This new picture naturally leads us to flat-band ferromagnetism.

In Section 15.6, we introduce flat-band ferromagnetism as a natural generalization of the long-range hopping model. We also discuss some results for the nearly-flat-band models obtained by adding a perturbation to the flat-band models.

In Section 15.7, we focus on experimental results which may be relevant to flat-band ferromagnetism.

In Section 15.8, we present some results about the possibility of metallic ferromagnetism in the Hubbard model. Preliminary rigorous results for related models in one-dimension are discussed.

Technical material is contained in the Appendices. In Appendices 15.9.1 and 15.9.2, we discuss the gauge transformation and the hole-particle transformation, respectively. In Appendix 15.9.3, we summarize some useful properties of positive semidefinite operators. In Appendix 15.9.4, we present (mainly for mathematicians) an explicit construction of the Hilbert space and fermion operators. In Appendix 15.9.5, we explain (for readers without background in condensed matter physics) the elementary notion of band structures for a single-electron in a tight-binding description. In Appendix 15.9.6, we present technical calculations required in the proof of Theorem 15.6.

15.2 Definition of the Hubbard model

15.2.1 Tight-binding description of a single electron

Before introducing the Hubbard model, we describe the corresponding single-electron problem. Let lattice Λ be a collection of $N_{\rm s}$ sites. Lattice sites $x, y, \ldots \in \Lambda$ represent atomic sites in a solid. In the tight-binding description, which is a kind of low-energy effective theory, we declare that *electrons can live only on lattice sites*. In the *single-orbital model* that we study

[5]For related reviews of mathematically rigorous results in the Hubbard model, see [20, 21]. Unfortunately there are many interesting related topics that we do not even mention in the present paper. For recent reviews of related topics from complementary points of view, see [22, 23, 24, 25].

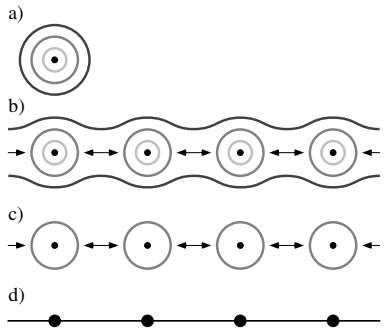


Figure 327: Highly schematic figures which explain the philosophy of tight-binding descriptions. a) A single atom which has multiple electrons in different orbits. b) When atoms come together to form a solid, electrons in the black orbits become itinerant, while those in the light gray orbits are still localized at the original atomic sites. Electrons in the gray orbits are mostly localized around the atomic sites, but tunnel to nearby gray orbits with non-negligible probabilities. c) We only consider the electrons in the gray orbits, which are expected to play essential roles in determining various aspects of low-energy physics of the system. d) If the gray orbit is non-degenerate, we get a lattice model in which electrons live on lattice sites and hop from one site to another.

here, we further assume that each atomic site carries a single non-degenerate[6] orbital state. Of course, actual atoms can have more than one orbit (or band). The philosophy behind the model building is that those electrons in other states do not play significant roles in determining the low-energy physics in which we are interested, and can be "forgotten" for the moment. See Figure 327.

Then, a quantum mechanical state of a single electron (with a fixed spin) is described by a vector state $\boldsymbol{\varphi} = (\varphi_x)_{x\in\Lambda}$ with $\varphi_x \in \mathbb{C}$. We denote by $\mathfrak{h} \cong \mathbb{C}^{N_{\rm s}} (\cong \ell^2(\Lambda;\mathbb{C}))$ the $N_{\rm s}$-dimensional Hilbert space formed by such $\boldsymbol{\varphi}$. The inner product of two states $\boldsymbol{\varphi}$ and $\boldsymbol{\psi} = (\psi_x)_{x\in\Lambda} \in \mathfrak{h}$ is

$$\langle \boldsymbol{\varphi}, \boldsymbol{\psi} \rangle = \sum_{x\in\Lambda} (\varphi_x)^* \psi_x. \tag{7}$$

For $x, y \in \Lambda$ with $x \neq y$, we denote by $t_{x,y}$ the quantum mechanical amplitude that an electron hops (i.e., tunnels) from the site y to x. We assume $t_{x,y}$ is real[7] and symmetric as $t_{x,y} = t_{y,x}$. For $x \in \Lambda$, we denote by $t_{x,x}$ the (real valued) potential energy[8] (which we usually write V_x) for the electron at site x. Then the Schrödinger equation for stationary states becomes

$$\varepsilon \varphi_x = \sum_{y\in\Lambda} t_{x,y}\, \varphi_y, \quad \text{for any } x \in \Lambda, \tag{8}$$

where ε is the energy eigenvalue[9]. By introducing the hopping matrix $\mathsf{T} = (t_{x,y})_{x,y\in\Lambda}$, which is a matrix on $\mathfrak{h}$, this can be written in a coordinate free vector form as

$$\varepsilon \boldsymbol{\varphi} = \mathsf{T}\boldsymbol{\varphi}. \tag{9}$$

[6] In some solids, the degeneracy in the original atomic orbit is lifted by crystalline anisotropy.

[7] In a system under magnetic field, $t_{x,y}$ is generally complex and satisfies $(t_{x,y})^* = t_{y,x}$.

[8] We do not want to refer to $V_x (= t_{x,x})$ as "site-dependent chemical potential." The notion of the chemical potential makes sense in the contexts of thermodynamics or statistical physics. This notion, in our opinion, should be distinguished from the purely quantum mechanical notion of the potential (although there is a relation between the two).

[9] It is also common to put a minus sign in front of $t_{x,y}$ in (8). We note that the hopping amplitude $t_{x,y}$ depends on the delicate overlap between orbital wave functions, and there are no simple principles with which we can determine their magnitudes or signs. Also see Appendix 15.9.1 for discussion of the arbitrariness of the sign of $t_{x,y}$.

Since T is real symmetric, the eigenvalue equation (9) has N_s eigenvalues $\varepsilon_1, \dots, \varepsilon_{N_\mathrm{s}}$. The corresponding eigenstate[10] $\boldsymbol{\psi}^{(j)} = (\psi_x^{(j)})_{x\in\Lambda}$ satisfies

$$\varepsilon_j \boldsymbol{\psi}^{(j)} = \mathsf{T}\boldsymbol{\psi}^{(j)}, \tag{10}$$

for each $j = 1, \dots, N_\mathrm{s}$. We can assume that the energy eigenvalues are ordered as $\varepsilon_j \le \varepsilon_{j+1}$, and eigenstates are orthonormal in the sense that $\left\langle \boldsymbol{\psi}^{(j)}, \boldsymbol{\psi}^{(j')} \right\rangle = \delta_{j,j'}$.

Let us consider a standard example. We take the one-dimensional lattice $\Lambda = \{1, 2, \dots, N_\mathrm{s}\}$, and set $t_{x,x+1} = t_{x+1,x} = -t$ for all $x \in \Lambda$, and $t_{x,y} = 0$ otherwise. We impose a periodic boundary condition and identify $N_\mathrm{s}+1$ with 1. The eigenvalues and eigenstates can be most naturally indexed as $\boldsymbol{\eta}^{(k)} = (\eta_x^{(k)})_{x\in\Lambda}$ and $\varepsilon(k)$, respectively, by the wave number $k = 2\pi n/N_\mathrm{s}$ where $n = 0, \pm 1, \pm 2, \dots, \pm\{(N_\mathrm{s}/2) - 1\}, N_\mathrm{s}/2$ (assuming N_s is even). Then we have

$$\eta_x^{(k)} = \frac{1}{\sqrt{N_\mathrm{s}}} e^{ikx}, \tag{11}$$

and $\varepsilon(k) = -2t\cos k$. Note that the eigenstates are described by plane waves.

15.2.2 Fermion operators and Hilbert space

We define the Hilbert space for many-electron problems using the fermion operator formalism[11]. For each lattice site $x \in \Lambda$ and spin index $\sigma = \uparrow, \downarrow$, we associate a fermion operator $c_{x,\sigma}$. One can freely take the conjugate[12], products, and linear combinations (with complex coefficients) of these operators (and the identity operator) to get new operators. We require that these operators satisfy the anticommutation relations[13]

$$\left\{c_{x,\sigma}^\dagger, c_{y,\tau}\right\} = \delta_{x,y}\delta_{\sigma,\tau}, \tag{12}$$

and

$$\left\{c_{x,\sigma}^\dagger, c_{y,\tau}^\dagger\right\} = \{c_{x,\sigma}, c_{y,\tau}\} = 0, \tag{13}$$

for any $x, y \in \Lambda$ and $\sigma, \tau = \uparrow, \downarrow$, where $\{A, B\} = AB + BA$. Note that (13) implies $(c_{x,\sigma})^2 = (c_{x,\sigma}^\dagger)^2 = 0$.

Physically, $c_{x,\sigma}$ and $c_{x,\sigma}^\dagger$ are interpreted as the operators which respectively annihilate and create an electron at site x with spin σ. The corresponding number operator is defined as

$$n_{x,\sigma} = c_{x,\sigma}^\dagger c_{x,\sigma}. \tag{14}$$

From (13), we find that number operators with different indices commute with each other. From (12) and (13), we see that $(n_{x,\sigma})^2 = c_{x,\sigma}^\dagger c_{x,\sigma} c_{x,\sigma}^\dagger c_{x,\sigma} = c_{x,\sigma}^\dagger (1 - c_{x,\sigma}^\dagger c_{x,\sigma}) c_{x,\sigma} = c_{x,\sigma}^\dagger c_{x,\sigma} = n_{x,\sigma}$. Thus we obtain $n_{x,\sigma}(1 - n_{x,\sigma}) = 0$, which means that $n_{x,\sigma}$ can only have eigenvalues 0 and 1. This is a mathematical realization of the Pauli exclusion principle. The number operator for site x is defined as $n_x = n_{x,\uparrow} + n_{x,\downarrow}$.

[10] In notation like $\psi_x^{(j)}$, the superscript j is the name of the state and the subscript x is the index for its component.

[11] The formalism is also called "second quantization." One should recall, however, that we are working with a many-body problem which is "quantized" only once. The explicit relation to the "first quantization" formalism can be found in Appendix 15.9.4. Readers with background in fields like functional analysis are suggested to take a look at this appendix.

[12] We denote the conjugate of an operator A by $A^\dagger$. We write $c_{x,\sigma}^\dagger$ instead of $(c_{x,\sigma})^\dagger$.

[13] The right-hand side of (12) means $\delta_{x,y}\delta_{\sigma,\tau}$ times the identity operator. Throughout the present paper, we refrain from writing the identity operator explicitly.

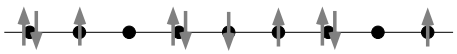


Figure 328: An allowed configuration in a tight-binding model with a single orbit per site. Each site in the lattice can be either empty, singly occupied by an electron with up or down spin, or doubly occupied by electrons with opposite spins. If we consider the Hubbard type interaction, we get an extra energy $U > 0$ whenever two electrons occupy a single site. (See Section 15.2.5.) The interaction energy for the above configuration is therefore $3U$.

We can now construct the Hilbert space for many-electron problems. We start from a single (vector) state Φ_{vac}, which physically corresponds to a (fictitious) state with no electrons in the lattice. This property is mathematically represented as

$$c_{x,\sigma}\Phi_{\text{vac}} = 0 \quad \text{for any } x \in \Lambda \text{ and } \sigma = \uparrow, \downarrow. \tag{15}$$

We fix a positive integer N_{e} such that $1 \le N_{\text{e}} \le 2N_{\text{s}}$, which is the total number of electrons in the system. It is useful to express the electron number in terms of the filling factor $\nu = N_{\text{e}}/(2N_{\text{s}})$, which takes a value in the range $0 < \nu \le 1$. Take N_{e} arbitrary sites $x_1, x_2, \dots, x_{N_{\text{e}}} \in \Lambda$ (with possible overlaps) and spin indices $\sigma_1, \sigma_2, \dots, \sigma_{N_{\text{e}}} = \uparrow, \downarrow$, and define the state

$$\Phi_{x_1,\sigma_1;x_2,\sigma_2;\dots;x_{N_{\text{e}}},\sigma_{N_{\text{e}}}} = c^\dagger_{x_1,\sigma_1} c^\dagger_{x_2,\sigma_2} \cdots c^\dagger_{x_{N_{\text{e}}},\sigma_{N_{\text{e}}}} \Phi_{\text{vac}}. \tag{16}$$

We interpret (16) as the state in which there is an electron at site x_i with spin σ_i for $i = 1, \dots, N_{\text{e}}$. We allow all the possible states of the form (16), but make identifications according to the anticommutation relations (13). In particular, the state (16) is vanishing whenever $(x_i, \sigma_i) = (x_j, \sigma_j)$ for some $i \ne j$. This is nothing but the Pauli exclusion principle[14]. In other words, each site in the lattice can be either empty, singly occupied by an electron with up or down spin, or doubly occupied by electrons with opposite spins (Figure 328). Also note that the anticommutation relation (13) leads to relations like

$$\Phi_{x_1,\sigma_1;x_2,\sigma_2} = -\Phi_{x_2,\sigma_2;x_1,\sigma_1}, \tag{17}$$

which is the well-known rule that a state changes its sign when one exchanges the *names* of two fermions.

The Hilbert space $\mathcal{H}_{N_{\text{e}}}$ for the given electron number N_{e} is generated by all the basis states of the form (16). The inner product[15], which we again denote as $\langle \cdot, \cdot \rangle$, is defined by $\langle \Phi_{\text{vac}}, \Phi_{\text{vac}} \rangle = 1$ and the anticommutation relations. We denote the norm of a state $\Phi \in \mathcal{H}_{N_{\text{e}}}$ as $\|\Phi\| = \langle \Phi, \Phi \rangle^{1/2}$.

We remark that there can be no *physical* situations which are represented by linear combinations of states with different electron numbers. To see this, we take two states $\Phi \in \mathcal{H}_{N_{\text{e}}}$ and $\Phi' \in \mathcal{H}_{N'_{\text{e}}}$ with $N_{\text{e}} \ne N'_{\text{e}}$. Since the total charge in the universe is conserved (as far as we know), the only (physically) possible way to linearly combine Φ and Φ' is to form a state like $\Xi = \Phi \otimes \Gamma_{N_{\text{e}}} + \Phi' \otimes \Gamma_{N'_{\text{e}}}$. Here $\Gamma_{N_{\text{e}}}$ and $\Gamma_{N'_{\text{e}}}$ are some states of the whole outside world with total charge N_{e} and N'_{e}, respectively. Note that the states $\Gamma_{N_{\text{e}}}$ and $\Gamma_{N'_{\text{e}}}$ are inevitably orthogonal. Therefore if A is any operator of the electron system, its expectation value in Ξ becomes

$$\begin{aligned} \langle A \rangle_\Xi &= \langle \Xi, (A \otimes \mathbf{1}), \Xi \rangle \\ &= \langle \Phi, A\Phi \rangle \langle \Gamma_{N_{\text{e}}}, \Gamma_{N_{\text{e}}} \rangle + \langle \Phi', A\Phi' \rangle \langle \Gamma_{N'_{\text{e}}}, \Gamma_{N'_{\text{e}}} \rangle \\ &= \omega \langle \Phi, A\Phi \rangle + \omega' \langle \Phi', A\Phi' \rangle, \end{aligned} \tag{18}$$

[14] If we set $N_{\text{e}} > 2N_{\text{s}}$, then the state (16) always vanishes.

[15] We do not use Dirac notation since we encounter many non-hermitian operators. Try rewriting the equation $\langle A\Phi, B\Psi \rangle = \langle \Phi, A^\dagger B\Psi \rangle$ in the Dirac notation.

where **1** is the identity operator for the outside world. This means that the state Ξ must be represented as a mixed state[16] if we restrict ourselves to the electron system[17].

15.2.3 Coordinate free notation

We introduce the standard coordinate-free notation for fermion operators. The systematic use of this notation simplifies some routine calculations. For a single-electron state $\boldsymbol{\varphi} = (\varphi_x)_{x\in\Lambda} \in \mathfrak{h}$, we define

$$C_\sigma^\dagger(\boldsymbol{\varphi}) = \sum_{x\in\Lambda} \varphi_x \, c_{x,\sigma}^\dagger, \quad C_\sigma(\boldsymbol{\varphi}) = \sum_{x\in\Lambda} (\varphi_x)^* \, c_{x,\sigma}. \tag{19}$$

These operators satisfy the generalized canonical anticommutation relations

$$\left\{ C_\sigma(\boldsymbol{\varphi}), C_{\sigma'}^\dagger(\boldsymbol{\psi}) \right\} = \delta_{\sigma,\sigma'} \, \langle \boldsymbol{\varphi}, \boldsymbol{\psi} \rangle \, , \tag{20}$$

and

$$\{ C_\sigma(\boldsymbol{\varphi}), C_{\sigma'}(\boldsymbol{\psi}) \} = \left\{ C_\sigma^\dagger(\boldsymbol{\varphi}), C_{\sigma'}^\dagger(\boldsymbol{\psi}) \right\} = 0, \tag{21}$$

for any $\boldsymbol{\varphi}, \boldsymbol{\psi} \in \mathfrak{h}$ and $\sigma, \sigma' = \uparrow, \downarrow$.

Again (21) implies $(C_\sigma^\dagger(\boldsymbol{\varphi}))^2 = (C_\sigma(\boldsymbol{\varphi}))^2 = 0$. We also have the following property.

Lemma 15.1 (Linear independence and the product of creation operators)**.** *Let $\boldsymbol{\varphi}^{(1)}, \ldots, \boldsymbol{\varphi}^{(n)}$ be n arbitrary states in $\mathfrak{h}$. The state*

$$C_\sigma^\dagger(\boldsymbol{\varphi}^{(1)}) \cdots C_\sigma^\dagger(\boldsymbol{\varphi}^{(n)}) \Phi_{\rm vac}, \tag{22}$$

which is usually called the Slater determinant state (see Appendix 15.9.4), is nonvanishing if and only if $\boldsymbol{\varphi}^{(1)}, \ldots, \boldsymbol{\varphi}^{(n)}$ are linearly independent.

Proof: Let $O = C_\sigma^\dagger(\boldsymbol{\varphi}^{(1)}) \cdots C_\sigma^\dagger(\boldsymbol{\varphi}^{(n)})$. Assume that the n states are linearly dependent. Then we can write $\boldsymbol{\varphi}^{(1)} = \sum_{j=2}^{n} \alpha_j \boldsymbol{\varphi}^{(j)}$ with some coefficients $\alpha_j \in \mathbb{C}$, and hence $C_\sigma^\dagger(\boldsymbol{\varphi}^{(1)}) = \sum_{j=2}^{n} \alpha_j C_\sigma^\dagger(\boldsymbol{\varphi}^{(j)})$. The product O vanishes because $(C_\sigma^\dagger(\boldsymbol{\varphi}^{(j)}))^2 = 0$.

Next we assume that the n states are linearly independent and show that $O\Phi_{\rm vac}$ is nonvanishing. By repeatedly using (20), (21) and $C_\sigma(\boldsymbol{\varphi})\Phi_{\rm vac} = 0$, we get

$$\begin{aligned} \| O\Phi_{\rm vac} \|^2 &= \left\langle \Phi_{\rm vac}, O^\dagger O \Phi_{\rm vac} \right\rangle \\ &= \left\langle \Phi_{\rm vac}, C_\sigma(\boldsymbol{\varphi}^{(n)}) \cdots C_\sigma(\boldsymbol{\varphi}^{(1)}) C_\sigma^\dagger(\boldsymbol{\varphi}^{(1)}) \cdots C_\sigma^\dagger(\boldsymbol{\varphi}^{(n)}) \Phi_{\rm vac} \right\rangle \\ &= \sum_{p:(1,\ldots,n)\to(p(1),\ldots,p(n))} (-1)^p \left\langle \boldsymbol{\varphi}^{(1)}, \boldsymbol{\varphi}^{(p(1))} \right\rangle \cdots \left\langle \boldsymbol{\varphi}^{(n)}, \boldsymbol{\varphi}^{(p(n))} \right\rangle , \end{aligned} \tag{23}$$

where p is summed over the $n!$ permutations of $(1, 2, \ldots, n)$, and $(-1)^p$ denotes the parity of p. The Gramm matrix G is an $n \times n$ matrix defined by $(G)_{i,j} = \left\langle \boldsymbol{\varphi}^{(i)}, \boldsymbol{\varphi}^{(j)} \right\rangle$. Then (23) implies

$$\| O\Phi_{\rm vac} \|^2 = \det(G). \tag{24}$$

[16] Note that the expectation value in the (unphysical) pure state $\Xi' = \alpha\Phi + \alpha'\Phi'$ becomes $\langle A \rangle_{\Xi'} = |\alpha|^2 \langle \Phi, A\Phi \rangle + |\alpha'|^2 \langle \Phi', A\Phi' \rangle + \{\alpha^*\alpha' \langle \Phi, A\Phi' \rangle + (\text{h.c.})\}$, which has extra cross terms.

[17] Although this remark might sound entirely trivial (and it indeed is), many beginners (and even some experts) are confused when they encounter BCS-type states which are linear combinations of states with different $N_{\rm e}$. One should always keep in mind that such linear combinations are introduced for purely mathematical (or theoretical) reasons (which are indeed deep and beautiful) and have nothing to do with realistic particle number fluctuations. When one has such a linear combination state, one *always* gets physically meaningful states with fixed electron numbers by taking suitable projections.

It is a well-known fact in linear algebra that G is regular (and hence $\det(G) \neq 0$) if and only if $\boldsymbol{\varphi}^{(1)}, \dots, \boldsymbol{\varphi}^{(n)}$ are linearly independent. □

The following lemma gives a basic characterization of the Slater determinant states (22). Although we do not use the lemma in the present paper, we state it here since it is enlightening. We encourage interested readers to prove the lemma.

Lemma 15.2 (Slater determinant depends on a subspace)**.** *Suppose that two sets* $\{\boldsymbol{\varphi}^{(1)}, \dots, \boldsymbol{\varphi}^{(n)}\}$ *and* $\{\boldsymbol{\psi}^{(1)}, \dots, \boldsymbol{\psi}^{(n)}\}$ *of states in* $\mathfrak{h}$ *span the same* n*-dimensional subspace of* $\mathfrak{h}$*. Then there is a nonvanishing constant* $c \in \mathbb{C}$*, and we have*

$$C_\sigma^\dagger(\boldsymbol{\varphi}^{(1)}) \cdots C_\sigma^\dagger(\boldsymbol{\varphi}^{(n)})\Phi_{\mathrm{vac}} = c\, C_\sigma^\dagger(\boldsymbol{\psi}^{(1)}) \cdots C_\sigma^\dagger(\boldsymbol{\psi}^{(n)})\Phi_{\mathrm{vac}}. \tag{25}$$

The following lemma is elementary, but of fundamental importance.

Lemma 15.3 (General basis of $\mathcal{H}_{N_\mathrm{e}}$)**.** *Let* $\boldsymbol{\varphi}^{(1)}, \dots, \boldsymbol{\varphi}^{(N_\mathrm{s})}$ *be* N_s *arbitrary states in* $\mathfrak{h}$ *which are linearly independent (but not necessarily orthogonal with each other). Then the states*

$$\Gamma_{S_\uparrow, S_\downarrow} = \left(\prod_{j \in S_\uparrow} C_\uparrow^\dagger(\boldsymbol{\varphi}^{(j)}) \right) \left(\prod_{j \in S_\downarrow} C_\downarrow^\dagger(\boldsymbol{\varphi}^{(j)}) \right) \Phi_{\mathrm{vac}} \tag{26}$$

for arbitrary subsets $S_\uparrow, S_\downarrow$ *of* $\{1, 2, \dots, N_\mathrm{s}\}$ *such that*[18] $|S_\uparrow| + |S_\downarrow| = N_\mathrm{e}$ *span the Hilbert space* $\mathcal{H}_{N_\mathrm{e}}$*.*

Proof: Since the N_s vectors are linearly independent, we have $\delta_{x,y} = \sum_{j=1}^{N_\mathrm{s}} a_{x,j} \varphi_y^{(j)}$ with some regular matrix $(a_{x,j})_{x \in \Lambda, j = 1, \dots, N_\mathrm{s}}$. Thus we have

$$c_{x,\sigma}^\dagger = \sum_{j=1}^{N_\mathrm{s}} a_{x,j} C_\sigma^\dagger(\varphi^{(j)}), \tag{27}$$

which means that the basis state (16) can be written as a linear combination of the states (26). □

Let $\mathsf{A} = (a_{x,y})_{x,y \in \Lambda}$ be a matrix on $\mathfrak{h}$. We define the corresponding bilinear form of fermion operators

$$B(\mathsf{A}) = \sum_{\substack{x,y \in \Lambda \\ \sigma = \uparrow, \downarrow}} c_{x,\sigma}^\dagger \, a_{x,y} \, c_{y,\sigma}, \tag{28}$$

which is usually called the "second quantization" of A. The easily verified commutation relation

$$\left[B(\mathsf{A}), C_\sigma^\dagger(\boldsymbol{\varphi}) \right] = C_\sigma^\dagger(\mathsf{A}\boldsymbol{\varphi}), \tag{29}$$

where $[A, B] = AB - BA$, sheds light on the relation between the single-particle quantum mechanics and the operator formalism. This will be useful later. A similar commutation relation for two bilinear forms

$$[B(\mathsf{A}), B(\mathsf{B})] = B([\mathsf{A}, \mathsf{B}]) \tag{30}$$

is also easy to prove.

We have so far treated the spin index σ separately from the coordinate x. It is of course possible to treat the combination (x, σ) as the coordinates of the system and consider the

[18] Throughout the present paper, we denote by $|S|$ the number of elements in the set S.

corresponding $2N_{\mathrm{s}}$-dimensional single-electron Hilbert space[19]. We can then restate all of the above results in the new language with suitable modifications. Let us only mention the relation corresponding to (30), because it will be useful later. Let $\tilde{\mathsf{A}}$ be a $2N_{\mathrm{s}} \times 2N_{\mathrm{s}}$ matrix indexed by (x, σ) with $x \in \Lambda$ and $\sigma = \uparrow, \downarrow$, and define $\tilde{B}(\tilde{\mathsf{A}}) = \sum_{x,y\in\Lambda} \sum_{\sigma,\tau=\uparrow,\downarrow} c^\dagger_{x,\sigma} (\tilde{\mathsf{A}})_{(x,\sigma),(y,\tau)}\, c_{y,\tau}$. Then for two such matrices $\tilde{\mathsf{A}}$ and $\tilde{\mathsf{B}}$, we have

$$[\tilde{B}(\tilde{\mathsf{A}}), \tilde{B}(\tilde{\mathsf{B}})] = \tilde{B}([\tilde{\mathsf{A}}, \tilde{\mathsf{B}}]), \tag{31}$$

where $[\tilde{\mathsf{A}}, \tilde{\mathsf{B}}]$ is the commutator as $2N_{\mathrm{s}} \times 2N_{\mathrm{s}}$ matrices.

15.2.4 Non-interacting system

Now that the Hilbert space has been prepared, we can introduce Hamiltonians. The Hamiltonian which describes the quantum mechanical hopping with the hopping matrix $\mathsf{T} = (t_{x,y})_{x,y\in\Lambda}$ is

$$H_{\mathrm{hop}} = B(\mathsf{T}) = \sum_{\substack{x,y\in\Lambda \\ \sigma=\uparrow,\downarrow}} t_{x,y}\, c^\dagger_{x,\sigma} c_{y,\sigma}. \tag{32}$$

To see that (32) describes the desired hopping, define the single-electron state $\Phi_{\boldsymbol{\varphi}} = C^\dagger_\sigma(\boldsymbol{\varphi})\Phi_{\mathrm{vac}}$ for $\boldsymbol{\varphi} \in \mathfrak{h}$. Then from the commutation relation (29), we immediately find that the Schrödinger equation $\varepsilon\Phi_{\boldsymbol{\varphi}} = H_{\mathrm{hop}}\Phi_{\boldsymbol{\varphi}}$ is equivalent to the single-electron Schrödinger equation (8) or (9).

When one has diagonalized the single-electron Schrödinger equation (8), it is easy to also diagonalize the many-body Hamiltonian (32). As in Section 15.2.1, we let $\boldsymbol{\psi}^{(j)}$ be the eigenstate satisfying (10) with energy ε_j, for $j = 1, 2, \ldots, N_{\mathrm{s}}$. Define the new fermion operator by

$$a^\dagger_{j,\sigma} = C^\dagger_\sigma(\boldsymbol{\psi}^{(j)}), \tag{33}$$

for $j = 1, 2, \ldots, N_{\mathrm{s}}$ and $\sigma = \uparrow, \downarrow$.

Let $S_\uparrow$ and $S_\downarrow$ be arbitrary subsets of $\{1, 2, \ldots, N_{\mathrm{s}}\}$ such that $|S_\uparrow| + |S_\downarrow| = N_{\mathrm{e}}$, and define

$$\Psi_{S_\uparrow,S_\downarrow} = \left(\prod_{j\in S_\uparrow} a^\dagger_{j,\uparrow}\right)\left(\prod_{j\in S_\downarrow} a^\dagger_{j,\downarrow}\right)\Phi_{\mathrm{vac}}. \tag{34}$$

Lemma 15.3 shows that these states form a basis of $\mathcal{H}_{N_{\mathrm{e}}}$. Note that the commutation relation (29) and the eigenstate equation (10) imply $[H_{\mathrm{hop}}, a^\dagger_{j,\sigma}] = [B(\mathsf{T}), C^\dagger_\sigma(\boldsymbol{\psi}^{(j)})] = C^\dagger_\sigma[\mathsf{T}\boldsymbol{\psi}^{(j)}] = \varepsilon_j a^\dagger_{j,\sigma}$. This and $H_{\mathrm{hop}}\Phi_{\mathrm{vac}} = 0$ immediately imply

$$H_{\mathrm{hop}}\Psi_{S_\uparrow,S_\downarrow} = \left(\sum_{j\in S_\uparrow} \varepsilon_j + \sum_{j\in S_\downarrow} \varepsilon_j\right)\Psi_{S_\uparrow,S_\downarrow}. \tag{35}$$

We have thus diagonalized H_{hop}.

By choosing subsets $S_\uparrow, S_\downarrow$ which minimize the energy eigenvalue $\sum_{j\in S_\uparrow} \varepsilon_j + \sum_{j\in S_\downarrow} \varepsilon_j$, we get ground state(s) of the present non-interacting model. In particular, if the corresponding single-electron energy eigenvalues are nondegenerate, i.e., $\varepsilon_j < \varepsilon_{j+1}$, and if N_{e} is even, the ground state of H_{hop} is unique and written as

$$\Phi_{\mathrm{GS}} = \left(\prod_{j=1}^{N_{\mathrm{e}}/2} a^\dagger_{j,\uparrow} a^\dagger_{j,\downarrow}\right)\Phi_{\mathrm{vac}}. \tag{36}$$

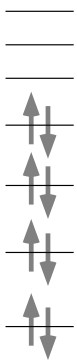


Figure 329: Schematic picture of the ground state of a non-interacting many-electron system. The lowest $N_{\rm e}/2$ single-electron energy levels are "filled" by both up spin and down spin electrons. The state naturally exhibits paramagnetism known as 'Pauli paramagnetism.'

This is nothing but the state obtained by "filling up" the low energy levels with up and down spin electrons, as one learns in elementary quantum mechanics (Figure 329).

In a single-electron eigenstate of the one-dimensional example in Section 15.2.1, the electron is in a plane wave state (11) with a definite wave number k. The same is true for any translation invariant model, as we see in Appendix 15.9.5. The fact that the Hamiltonian $H_{\rm hop}$ is diagonalized in the basis (34) implies that the electrons behave as "waves" in this non-interacting (Hubbard) model.

15.2.5 Interaction

It is believed that interactions between electrons in a solid mainly come from the electrostatic Coulomb force. Although the Coulomb force in vacuum is long-ranged, we consider *an extremely short-range interaction which acts only when two electrons occupy the same site (i.e., the same atomic orbit).* A crude justification of such a short-range (Coulomb) interaction comes from the observation that the Coulomb force should be most dominant when two electrons approach within the minimum possible distance. In a slightly more sophisticated justification, one argues that the long range Coulomb force is screened by electrons in different orbital states which we have decided to forget. Our point of view, however, is that models with artificial short-range interactions are worth studying because they are among *the minimum models which can be studied to elicit universal properties of strongly interacting electron systems.*

The Hamiltonian for the short-range Coulomb repulsion is

$$H_{\rm int} = U \sum_{x \in \Lambda} n_{x,\uparrow} n_{x,\downarrow}, \tag{37}$$

where $U \geq 0$ is the energy from the repulsive interaction. See Figure 328.

The expression (37) is already diagonalized, and it is trivial to write down eigenstates of $H_{\rm int}$. Let $L_\uparrow, L_\downarrow$ be arbitrary subsets of the lattice Λ such that $|L_\uparrow| + |L_\downarrow| = N_{\rm e}$, and define

$$\Psi_{L_\uparrow, L_\downarrow} = \left(\prod_{x \in L_\uparrow} c^\dagger_{x,\uparrow} \right) \left(\prod_{x \in L_\downarrow} c^\dagger_{x,\downarrow} \right) \Phi_{\rm vac}. \tag{38}$$

Then by using the commutation relation $[n_{x,\sigma}, c^\dagger_{x',\sigma'}] = \delta_{x,x'} \delta_{\sigma,\sigma'} c^\dagger_{x,\sigma}$, we find

$$H_{\rm int} \Psi_{L_\uparrow, L_\downarrow} = U |L_\uparrow \cap L_\downarrow| \Psi_{L_\uparrow, L_\downarrow}. \tag{39}$$

[19] We shall use such a formalism in Appendix 15.9.4.

The ground state for a given electron number N_e can be constructed by choosing subsets $L_\uparrow, L_\downarrow$ that minimize the energy eigenvalue $U|L_\uparrow \cap L_\downarrow|$. When $N_e \le N_s$, one can always choose $L_\uparrow$ and $L_\downarrow$ so that $L_\uparrow \cap L_\downarrow = \emptyset$ holds. In this case, the ground states are highly degenerate and have energy equal to 0.

It is clear (from the beginning) that the interaction Hamiltonian $H_{\rm int}$ is most naturally treated if we regard electrons as (classical) "particles" which live on lattice sites.

15.2.6 Hubbard model

The Hubbard model describes a tight-binding electron model in which *electrons hop around the lattice and interact with each other through short-range repulsive interactions*. The full Hamiltonian of the single-orbital Hubbard model[20] is simply

$$H = H_{\rm hop} + H_{\rm int}. \tag{40}$$

We have already seen that both $H_{\rm hop}$ and $H_{\rm int}$ can be easily diagonalized. We have observed, however, that electrons behave as "waves" in $H_{\rm hop}$, while they behave as "particles" in $H_{\rm int}$. How do they behave in a system whose Hamiltonian is a sum of these totally different Hamiltonians? This is indeed a fascinating problem which is deeply related to the wave-particle dualism in quantum physics. We might say that many of the important models in many-body problems, including the φ^4 quantum field theory and the Kondo problem, are minimum models which take into account both the wave-like nature and the particle-like nature (through point-like nonlinear interactions) of matter.

From a technical point of view, the wave-particle dualism implies that the Hamiltonians $H_{\rm int}$ are $H_{\rm hop}$ do not commute with each other. Even when each Hamiltonian is diagonalized, it is still highly nontrivial (or impossible) to find the properties of their sum. Of course, mathematical difficulty does not automatically guarantee that the model is worth studying. A truly exciting characteristic of the Hubbard model is that, though the Hamiltonians $H_{\rm hop}$ and $H_{\rm int}$ do not favor any nontrivial order, their sum $H = H_{\rm hop} + H_{\rm int}$ is believed to generate various types of nontrivial order including antiferromagnetism, ferromagnetism, and superconductivity. When we sum up the two innocent Hamiltonians $H_{\rm hop}$ and $H_{\rm int}$, competition between their wave-like and particle-like characters (or between linearity and nonlinearity) takes place, and one gets various interesting "physics". To confirm this fascinating scenario is a challenging problem in theoretical and mathematical physics.

15.3 Magnetism in the Hubbard model

15.3.1 Spin operators and $SU(2)$ invariance

We introduce the total spin operators[21] $\hat{\mathbf{S}}_{\rm tot} = (\hat{S}^{(1)}_{\rm tot}, \hat{S}^{(2)}_{\rm tot}, \hat{S}^{(3)}_{\rm tot})$ of the system by

$$\hat{S}^{(\alpha)}_{\rm tot} = \frac{1}{2}\sum_{x\in\Lambda}\sum_{\sigma,\tau=\uparrow,\downarrow} c^\dagger_{x,\sigma}(p^{(\alpha)})_{\sigma,\tau}\, c_{x,\sigma}, \tag{41}$$

for $\alpha = 1, 2$, and 3, where $p^{(\alpha)}$ are the Pauli matrices. It is clear from the general commutation relation (31) that[22] $\hat{S}^{(1)}_{\rm tot}, \hat{S}^{(2)}_{\rm tot}, \hat{S}^{(3)}_{\rm tot}$ satisfy the standard commutation relation for quantum

[20] This model is often called the single-band Hubbard model. We find this terminology confusing since a single-orbital model can have multiple bands depending on the structure of the hopping matrix T. See Appendix 15.9.5.

[21] We use the index $\alpha = 1, 2, 3$ to indicate the three axes, because the symbols x, y, z are already used as lattice sites.

[22] Note that $\hat{S}^{(\alpha)}_{\rm tot} = \tilde{B}(\tilde{p}^{(\alpha)})$ with $(\tilde{p}^{(\alpha)})_{(x,\sigma),(y,\tau)} = \delta_{x,y}(p^{(\alpha)})_{\sigma,\tau}$.

mechanical angular momentum operators. We define the raising and lowering operators by $S^{\pm}_{\rm tot} = \hat{S}^{(1)}_{\rm tot} \pm i\hat{S}^{(2)}_{\rm tot}$.

The operators $\hat{\mathbf{S}}_{\rm tot}$ are the generators of the global $SU(2)$ rotations in the spin space. We say an operator A is $SU(2)$ invariant if it commutes with $\hat{S}^{(\alpha)}_{\rm tot}$ for $\alpha = 1, 2$, and 3. Intuitively speaking, an $SU(2)$ invariant operator does not change if we change the "quantization axis" of spins in an arbitrary manner.

A typical $SU(2)$ invariant operator is the number operator $n_x = n_{x,\uparrow} + n_{x,\downarrow}$. The hopping Hamiltonian $H_{\rm hop}$ of (32) is also $SU(2)$ invariant. These facts can be checked easily by using the general commutation relations (31). The interaction Hamiltonian $H_{\rm int}$ of (37) is also $SU(2)$ invariant. This fact becomes evident if we note $n_{x,\uparrow}n_{x,\downarrow} = \{(n_x)^2 - n_x\}/2$.

Knowing that the Hamiltonian $H = H_{\rm hop} + H_{\rm int}$ is $SU(2)$ invariant, we can make use of the standard representation theory of quantum mechanical angular momenta. We note that the three operators H, $\hat{S}^{(3)}_{\rm tot}$, and $(\hat{\mathbf{S}}_{\rm tot})^2 = \sum_{\alpha=1,2,3} (\hat{S}^{(\alpha)}_{\rm tot})^2$ commute with each other. We will always consider simultaneous eigenstates of these operators. We denote by $S^{(3)}_{\rm tot}$ the eigenvalue of $\hat{S}^{(3)}_{\rm tot}$, and by $S_{\rm tot}(S_{\rm tot}+1)$ the eigenvalue of $(\hat{\mathbf{S}}_{\rm tot})^2$. We call $S_{\rm tot}$ the total spin of the state. Let

$$S_{\rm max} = \begin{cases} N_{\rm e}/2 & \text{if } N_{\rm e} \le N_{\rm s}; \\ (2N_{\rm s} - N_{\rm e})/2 & \text{if } N_{\rm e} \ge N_{\rm s}. \end{cases} \tag{42}$$

Then the allowed values of $S_{\rm tot}$ are $S_{\rm tot} = 0, 1, \dots, S_{\rm max}$ if $N_{\rm e}$ is even, and $S_{\rm tot} = 1/2, 3/2, \dots, S_{\rm max}$ if $N_{\rm e}$ is odd.

15.3.2 Ferromagnetism

Ferromagnetism is probably the most intuitive among various magnetic phenomena exhibited by solids. Here we shall be rather restrictive and only consider the strongest form of ferromagnetism, namely, saturated ferromagnetism in ground states.

Definition 15.4. *A Hubbard model is said to exhibit ferromagnetism if any ground state of H has the total spin $S_{\rm tot} = S_{\rm max}$.*

Recall that any state with $S_{\rm tot} = S_{\rm max}$ has its copies in the subspaces with $S^{(3)}_{\rm tot} = -S_{\rm max}, -S_{\rm max}+1, \dots, S_{\rm max}$. Therefore the ground states are at least $(2S_{\rm max}+1)$-fold degenerate when there is ferromagnetism (in the above strong sense). Let $\Phi_{\rm GS}$ be a ground state with $S^{(3)}_{\rm tot} = S_{\rm max}$. Since $\Phi_{\rm GS}$ contains only up spin electrons, we must have $H_{\rm int}\Phi_{\rm GS} = 0$. This means that $\Phi_{\rm GS}$ is the lowest energy state of $H_{\rm hop}$ within the subspace of $N_{\rm e}$ up spin electrons. Recalling the general eigenstates (34) of $H_{\rm hop}$, we find that the desired ferromagnetic ground state is

$$\Phi_{\rm GS} = \left(\prod_{j=1}^{N_{\rm e}} a^{\dagger}_{j,\uparrow} \right) \Phi_{\rm vac}, \tag{43}$$

with $a^{\dagger}_{j,\uparrow}$ defined in (33), which has the ground state energy

$$E_{\rm ferro} = \sum_{j=1}^{N_{\rm e}} \varepsilon_j. \tag{44}$$

When we have $\varepsilon_{N_{\rm e}} < \varepsilon_{N_{\rm e}+1}$, (43) is the unique ground state[23] in the subspace with $S^{(3)}_{\rm tot} = S_{\rm max}$. The ground state Φ_M with $S^{(3)}_{\rm tot} = M$ can be obtained from this ground state by the standard

[23] When $\varepsilon_{N_{\rm e}}$ is degenerate with $\varepsilon_n, \dots, \varepsilon_m$ (where $n \le N_{\rm e} < m$), we can replace the states $j = n, \dots, N_{\rm e}$ with arbitrary states from $j = n \dots, m$ to get other ground states. The ground states are degenerate even in the subspace with $S^{(3)}_{\rm tot} = S_{\rm max}$.

relation

$$\Phi_M = \frac{(S^-_{\rm tot})^{S_{\rm max}-M}\Phi_{\rm GS}}{\left\|(S^-_{\rm tot})^{S_{\rm max}-M}\Phi_{\rm GS}\right\|}. \tag{45}$$

The above elementary construction of the ground states (provided that the model exhibits ferromagnetism) is based on the fact that any state with $S_{\rm tot} = S_{\rm max}$ *does not feel* the on-site Coulomb repulsion. This reduces the problem to that of non-interacting spinless fermions. We stress that this is a very special feature of the Hubbard model, which has only on-site interactions.

15.3.3 Instability of ferromagnetism

To see that ferromagnetism is indeed a delicate phenomenon, we discuss some results which show that the Hubbard model with certain conditions does *not* exhibit ferromagnetism.

We first look at the non-interacting model with the Hamiltonian $H = H_{\rm hop}$ that we studied in Section 15.2.4. Assuming that the single-electron energy eigenvalues are non-degenerate, we found that the ground state is uniquely given by (36). By noting that the operator $a^\dagger_{j,\uparrow}a^\dagger_{j,\downarrow}$ is $SU(2)$ invariant and $\hat{S}^{(\alpha)}_{\rm tot}\Phi_{\rm vac} = 0$, we immediately find $\hat{S}^{(\alpha)}_{\rm tot}\Phi_{\rm GS} = 0$. Thus the ground state has the total spin[24] $S_{\rm tot} = 0$, and there is no ferromagnetism. This is nothing but the well-known Pauli paramagnetism. (See Figure 329.)

It is also trivial that the non-hopping model with $H = H_{\rm int}$ we studied in Section 15.2.5 does not exhibit ferromagnetism. In this case any state (38) for $L_\uparrow, L_\downarrow$ with minimum $|L_\uparrow \cap L_\downarrow|$ is a ground state. Thus the ground states are highly degenerate and exhibit a kind of paramagnetism.

Therefore neither $H_{\rm hop}$ nor $H_{\rm int}$ alone favor ferromagnetism. As we have stressed in Section 15.2.6, our hope is that, when these two Hamiltonians are added into a single Hubbard Hamiltonian $H = H_{\rm hop} + H_{\rm int}$, their "competition" generates totally new phenomena, including ferromagnetism.

Next we consider the situation in which the Coulomb interaction U is finite but small. We find here that there cannot be (saturated) ferromagnetism.

Theorem 15.5 (Impossibility of ferromagnetism for small U)**.** *Suppose $0 \le U < \varepsilon_{N_{\rm e}} - \varepsilon_1$. Then the ground state of the Hubbard model does not have $S_{\rm tot} = S_{\rm max}$, i.e., the model does not exhibit ferromagnetism. Note that the fermi energy $\varepsilon_{N_{\rm e}} - \varepsilon_1$ is usually independent of the system size for a fixed filling factor ν.*

Proof: One of the lowest energy states with $S_{\rm tot} = S_{\rm max}$ is given by (43) and has the energy (44). Consider a normalized trial state

$$\Psi = a^\dagger_{1,\downarrow}\left(\prod_{j=1}^{N_{\rm e}-1} a^\dagger_{j,\uparrow}\right)\Phi_{\rm vac}, \tag{46}$$

which is obtained from (43) by removing the up spin electron with the highest energy and then adding a down spin electron with the lowest energy. Noting the $SU(2)$ invariance of $a^\dagger_{1,\downarrow}a^\dagger_{1,\uparrow}$, one finds that Ψ has $S_{\rm tot} = S_{\rm max} - 1$. We want to evaluate the energy expectation value $\langle\Psi, H\Psi\rangle$. For the kinetic energy, we have $\langle\Psi, H_{\rm hop}\Psi\rangle = \varepsilon_1 + \sum_{j=1}^{N_{\rm e}-1}\varepsilon_j$. As for $H_{\rm int}$, we note that

[24] That $S_{\rm tot} = 0$ also follows from the uniqueness of the ground state.

the inequality[25] $n_{x,\uparrow} \le 1$ implies $H_{\rm int} \le U \sum_{x\in\Lambda} n_{x,\downarrow}$ to get $\langle \Psi, H_{\rm int}\Psi\rangle \le \left\langle \Psi, U\sum_{x\in\Lambda} n_{x,\downarrow}\Psi \right\rangle = U$. Therefore we have

$$\langle \Psi, H\Psi\rangle - E_{\rm ferro} \le \varepsilon_1 - \varepsilon_{N_{\rm e}} + U < 0, \tag{47}$$

where $E_{\rm ferro}$ is defined in (44), and the final bound follows from the condition of the theorem. From the variational principle, we see that $E_{\rm ferro}$ is not the ground state energy of H. □

Although the above theorem ensures that the ground state cannot be ferromagnetic, it does not provide any information about the nature of the true ground state of the model. To study the latter explicitly is in general a very difficult problem which (for the moment) is possible only in the simplest one-dimensional model.

Finally, we discuss the situation in which the interaction may be large but the density of electrons is very low. It is expected that the chance of electrons to collide with each other in this case becomes very small. It is likely that the model is close to an ideal gas, and there is no ferromagnetism.

This naive guess is easily justified for "healthy" models in dimensions three (or higher). The dimensionality of the lattice is taken into account by assuming that there are positive constants[26] c, n_0, ρ_0, and d, and the single electron energy eigenvalues satisfy

$$\varepsilon_n - \varepsilon_1 \ge c\left(\frac{n-n_0}{N_{\rm s}}\right)^{2/d}, \tag{48}$$

for any n such that $n/N_{\rm s} \le \rho_0$. Note that the right-hand side represents the n dependence of energy levels in a usual d-dimensional quantum mechanical system. Then we have the following theorem due to Pieri, Daul, Baeriswyl, Dzierzawa, and Fazekas [26].

Theorem 15.6 (Impossibility of ferromagnetism at low densities)**.** *Take a $H_{\rm hop}$ which has translation invariance (as in Appendix 15.9.5) and satisfies (48) with positive c, n_0, ρ_0, and $d > 2$. Then there exists a constant $\rho_1 > 0$, and the corresponding Hubbard model does not exhibit ferromagnetism for any $U \ge 0$ if $N_{\rm e}/N_{\rm s} \le \rho_1$ holds.*

Outline of proof: The naive trial state (46) does not work for large U. We follow [27], and consider the Roth state [28]

$$\widetilde{\Psi} = P_0\Psi, \tag{49}$$

where Ψ is defined in (46), and

$$P_0 = \prod_{x\in\Lambda}(1 - n_{x,\uparrow}n_{x,\downarrow}) \tag{50}$$

is the orthogonal projection (called the 'Gutzwiller projection') onto the space with no doubly occupied sites. Because of the projection, the state (49) minimizes the Coulomb interaction as $H_{\rm int}\Psi = 0$. Thus we only need to evaluate the expectation value of $H_{\rm hop}$. After a tedious but straightforward calculation whose details can be found in Appendix 15.9.6, we find

$$\frac{\left\langle \widetilde{\Psi}, H\widetilde{\Psi}\right\rangle}{\left\langle \widetilde{\Psi}, \widetilde{\Psi}\right\rangle} - E_{\rm ferro} \le \varepsilon_1 - \varepsilon_{N_{\rm e}} + c'\rho, \tag{51}$$

where $\rho = N_{\rm e}/N_{\rm s}$ is the electron density and $c' > 0$ is a constant. From the assumption (48), we find that the right-hand side becomes strictly negative for sufficiently small ρ provided that $d > 2$. □

[25] See Definition 15.22 for the definition of inequalities for operators.

[26] n_0 and d are the degeneracy of the single-electron ground states and the dimension of the system, respectively.

That we have a restriction on dimensionality in Theorem 15.6 is not merely technical. In a one-dimensional system, moving electrons must eventually collide with each other for an obvious geometric reason. Thus a one-dimensional model cannot be regarded as close to ideal no matter how low the electron density is. We do not know whether the inapplicability of the theorem to two-dimensional systems is physically meaningful or not.

15.3.4 Two more theorems for the absence of ferromagnetism

We briefly discuss (without proofs) two strong theorems which also rule out ferromagnetism.

The classical Lieb-Mattis theorem [29] states (among other things) that one can never have ferromagnetism in the one-dimensional Hubbard model with only nearest neighbor hoppings[27]. One-dimensional Hubbard models with next nearest neighbor hoppings may exhibit ferromagnetism as we explicitly see in Sections 15.6 and 15.8.

Theorem 15.7 (Lieb-Mattis theorem)**.** *Consider a Hubbard model with even N_{e} on a one-dimensional lattice $\Lambda = \{1, 2, \dots, N_{\mathrm{s}}\}$ with open boundary conditions. We assume that the hopping matrix elements satisfy $|t_{x,y}| < \infty$ when $x = y$, $0 < |t_{x,y}| < \infty$ when $|x - y| = 1$, and are vanishing otherwise. Then for any real U, the ground state of the model is unique and has $S_{\mathrm{tot}} = 0$.*

The next important theorem is due to Lieb [11].

Theorem 15.8 (Lieb's theorem, special case)**.** *Suppose that the lattice Λ is decomposed into two sublattices as $\Lambda = A \cup B$ with $|A| = |B|$, and we have $t_{x,y} = 0$ when $x, y \in A$ or $x, y \in B$. We also assume that the entire lattice is connected by nonvanishing $t_{x,y}$. When $N_{\mathrm{e}} = N_{\mathrm{s}}$, the ground state of the model is unique and has $S_{\mathrm{tot}} = 0$ for any $U \geq 0$.*

The electron number $N_{\mathrm{e}} = N_{\mathrm{s}}$ is usually referred to as "half-filling", since $2N_{\mathrm{s}}$ is the maximum possible number for N_{e}. The low energy properties of a Hubbard model at half-filling is believed to be described by the antiferromagnetic Heisenberg model. Lieb's theorem stated above gives a partial justification to this belief. This theorem applies to models on lattices with $|A| \neq |B|$ as well. In this case, it implies the existence of *ferrimagnetism*, as is explicitly stated in [12].

15.4 Nagaoka's ferromagnetism

15.4.1 Weak version of Nagaoka's theorem

We are now ready to discuss Nagaoka's ferromagnetism[28] [10]. Briefly speaking, Nagaoka's theorem establishes that some Hubbard models exhibit saturated ferromagnetism when the number of electrons is one less than the half-filling (i.e., $N_{\mathrm{e}} = N_{\mathrm{s}} - 1$) and the Coulomb repulsion U is infinitely large. Given the general fact that a half-filled system never exhibits ferromagnetism, this is a rather striking result, which demonstrates that strongly interacting electron systems can produce very rich and sometimes surprising phenomena.

When $U = \infty$ and $N_{\mathrm{e}} = N_{\mathrm{s}} - 1$, states with finite energies have no doubly occupied sites, and there is exactly one empty site which we call "hole." The basic mechanism of Nagaoka's ferromagnetism is that the hole hops around the lattice and generates a suitable linear combination of the basis states, in such a way that the resulting state exhibits ferromagnetism. See Figure 330. Thouless [9] also discussed the same mechanism of ferromagnetism in slightly more restricted situations. Also see [31].

[27] The present theorem appears in the Appendix of [29]. The main body of [29] treats interacting electron systems in continuous spaces.

[28] As for the presentation of the results and the proofs, we follow [30], where a generalization of Nagaoka's theorem was discussed. The proof briefly described in [30] is now presented in its full detail.

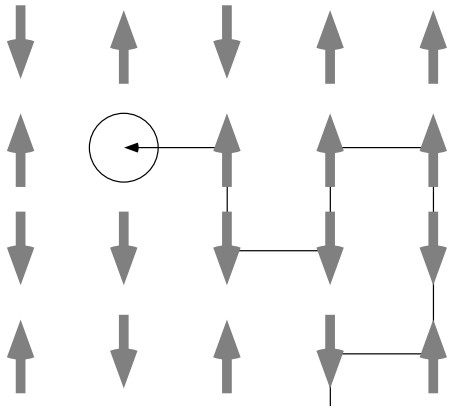


Figure 330: Schematic picture of the origin of Nagaoka's ferromagnetism. When the hole hops around the lattice, the spin configuration is changed. For a model with $t_{x,y} \geq 0$, the hole motion produces a precise linear combination of various spin configurations which leads to a ferromagnetic state.

We start from the following weaker result which is easy to state and prove.

Theorem 15.9 (Weak version of Nagaoka's theorem)**.** *Consider an arbitrary Hubbard model with $t_{x,y} \geq 0$ for any $x, y \in \Lambda$, $N_{\rm e} = N_{\rm s} - 1$, and $U = \infty$. Then among the ground states there are $(2S_{\rm tot} + 1)$ states with total spin $S_{\rm tot} = S_{\max} (= N_{\rm e}/2)$.*

Note that the theorem does not establish the existence of ferromagnetism since it does not state ferromagnetic states are the only ground states[29]. Stronger statement will be proved in the next section as Theorem 15.13.

Let us prove Theorem 15.9. We first consider the treatment of the $U \to \infty$ limit. Let us decompose the Hilbert space $\mathcal{H}_{N_{\rm e}}$ as

$$\mathcal{H}_{N_{\rm e}} = \mathcal{H}^{(0)}_{N_{\rm e}} \oplus \mathcal{H}'_{N_{\rm e}}, \tag{52}$$

where $\mathcal{H}^{(0)}_{N_{\rm e}}$ consists of the states[30] satisfying $H_{\rm int}\Phi = 0$. We denote by P_0 (see (50)) the orthogonal projection onto $\mathcal{H}^{(0)}_{N_{\rm e}}$. Then we have the following.

Lemma 15.10 (Characterization of the $U \to \infty$ limit)**.** *The $U \to \infty$ limit of the Hubbard model is equivalent to studying the Hamiltonian[31] $\widetilde{H} = P_0 H_{\rm hop}$ on the Hilbert space $\mathcal{H}^{(0)}_{N_{\rm e}}$.*

Proof: By continuity, the eigenstates of $H = H_{\rm hop} + H_{\rm int}$ can be classified into those with energies diverging as $U \to \infty$ and those with finite energies in the $U \to \infty$ limit. We are only interested in the latter. Let Φ be an eigenstate with energy E which has a finite limit. We can assume both Φ and E are continuously parameterized by U. Note that we have $P_0\Phi \to \Phi$ in the $U \to \infty$ limit, because otherwise E diverges in this limit. By applying the projection P_0 onto the Schrödinger equation $E\Phi = (H_{\rm hop} + H_{\rm int})\Phi$, we get $EP_0\Phi = P_0 H_{\rm hop}\Phi$ because $P_0 H_{\rm int}\Phi = 0$. In the $U \to \infty$ limit, this becomes $E\Phi = \widetilde{H}\Phi$ for $\Phi \in \mathcal{H}^{(0)}_{N_{\rm e}}$. □

[29] Note that the statement of the theorem is valid in the trivial model with $t_{x,y} = 0$ for all x, y, where all the possible spin states are degenerate. We must remark that some of the published "proofs" of Nagaoka's theorem and its extensions only prove such weak statements. See also footnote 38.

[30] We are assuming $N_{\rm e} \leq N_{\rm s}$. If $N_{\rm e} > N_{\rm s}$, the relevant condition is replaced by $H_{\rm int}\Phi = U(N_{\rm e} - N_{\rm s})\Phi$, and Lemma 15.10 is still valid.

[31] In the present case, we will eventually find that the ground states in the $U \to \infty$ limit are eigenstates of $H_{\rm hop}$. This is of course an accidental situation for the ferromagnetic states. Examples of exact ground states in the $U \to \infty$ limit which are not eigenstates of $H_{\rm hop}$ can be found in [32, 33].

We now prepare a basis for the Hilbert space $\mathcal{H}_{N_e}^{(0)}$. We use the basis states characterized by the position x of the hole and the spin configuration[32] $\boldsymbol{\sigma} = (\sigma_y)_{y\in\Lambda\mathbf{S}x} \in \mathcal{S}_{\Lambda\mathbf{S}x}$, defined as

$$\Phi_{x,\boldsymbol{\sigma}} = c_{x,\uparrow}\left(\prod_{y\in\Lambda} c^\dagger_{y,\sigma'_y}\right)\Phi_{\rm vac} = c_{x,\downarrow}\left(\prod_{y\in\Lambda} c^\dagger_{y,\sigma''_y}\right)\Phi_{\rm vac}. \tag{53}$$

Here the product is taken over all the sites in Λ with an arbitrary but fixed order. We gave two equivalent expressions for $\Phi_{x,\boldsymbol{\sigma}}$. The spin configurations $(\sigma'_y)_{y\in\Lambda}$ and $(\sigma''_y)_{y\in\Lambda}$ are essentially the same as $\boldsymbol{\sigma} = (\sigma_y)_{y\in\Lambda\mathbf{S}x}$, and defined by $\sigma'_y = \sigma''_y = \sigma_y$ for all $y \in \Lambda\mathbf{S}x$. As for the missing site x, we set $\sigma'_x = \uparrow$ and $\sigma''_x = \downarrow$. In the definition (53), we are simply supplying an electron at the hole site x, and then annihilating the electron by $c_{x,\uparrow}$ or $c_{x,\downarrow}$. This may sound meaningless, but we get fermion signs appropriate for our purpose in this way.

We wish to examine the action of $\widetilde{H} = P_0 H_{\rm hop}$ on the states (53). In order to get a finite contribution, an electron must hop into the hole site x from other sites. So we examine the action of $\sum_{\sigma=\uparrow,\downarrow} c^\dagger_{x,\sigma}c_{z,\sigma}$ on $\Phi_{x,\boldsymbol{\sigma}}$. Using the two equivalent expressions (53) for $\Phi_{x,\boldsymbol{\sigma}}$, we get

$$\begin{aligned}
\sum_{\sigma=\uparrow,\downarrow} c^\dagger_{x,\sigma}c_{z,\sigma}\Phi_{x,\boldsymbol{\sigma}} &= -c_{z,\uparrow}n_{x,\uparrow}\left(\prod_{y\in\Lambda} c^\dagger_{y,\sigma'_y}\right)\Phi_{\rm vac} - c_{z,\downarrow}n_{x,\downarrow}\left(\prod_{y\in\Lambda} c^\dagger_{y,\sigma''_y}\right)\Phi_{\rm vac} \\
&= -c_{z,\uparrow}\left(\prod_{y\in\Lambda} c^\dagger_{y,\sigma'_y}\right)\Phi_{\rm vac} - c_{z,\downarrow}\left(\prod_{y\in\Lambda} c^\dagger_{y,\sigma''_y}\right)\Phi_{\rm vac} \\
&= -\Phi_{z,\boldsymbol{\sigma}_{z\to x}},
\end{aligned} \tag{54}$$

where $\boldsymbol{\sigma}_{z\to x} \in \mathcal{S}_{\Lambda\mathbf{S}z}$ is the new spin configuration on $\Lambda\mathbf{S}z$ obtained from $\boldsymbol{\sigma}$ by moving σ_z to x. Note that in the second line of (54), only one of the two terms survive depending on the value of σ_z. Therefore the matrix elements of the effective Hamiltonian $\widetilde{H}$ are given by

$$\left\langle \Phi_{y,\boldsymbol{\tau}}, \widetilde{H}\Phi_{x,\boldsymbol{\sigma}} \right\rangle = \begin{cases} -t_{x,y} & \text{if } \boldsymbol{\tau} = \boldsymbol{\sigma}_{y\to x}; \\ 0 & \text{otherwise.} \end{cases} \tag{55}$$

Let $\Phi_{\rm GS}$ be a ground state of H. Since it has a finite energy, it is expanded as

$$\Phi_{\rm GS} = \sum_{x\in\Lambda}\sum_{\boldsymbol{\sigma}\in\mathcal{S}_{\Lambda\mathbf{S}x}} \varphi(x,\boldsymbol{\sigma})\Phi_{x,\boldsymbol{\sigma}}. \tag{56}$$

Since the matrix elements (55) of $\widetilde{H}$ are real, we can assume that the coefficients $\varphi(x,\boldsymbol{\sigma})$ are real[33]. We define

$$\xi_x = \left(\sum_{\boldsymbol{\sigma}\in\mathcal{S}_{\Lambda\mathbf{S}x}} (\varphi(x,\boldsymbol{\sigma}))^2\right)^{1/2}, \tag{57}$$

and a ferromagnetic state

$$\Phi_\uparrow = \sum_{x\in\Lambda} \xi_x \Phi_{x,(\uparrow)}, \tag{58}$$

where $(\uparrow)$ denotes the spin configuration with all spins up.

[32] By $\Lambda\mathbf{S}x$ we mean the lattice obtained by removing x from Λ.

[33] If this is not the case, we redefine $\{\varphi(x,\boldsymbol{\sigma}) + \varphi(x,\boldsymbol{\sigma})^*\}/2$ as $\varphi(x,\boldsymbol{\sigma})$. The corresponding $\Phi_{\rm GS}$ is also a ground state.

By using (56) and (55), we find

$$\begin{aligned}
\left\langle \Phi_{\rm GS}, \widetilde{H}\Phi_{\rm GS} \right\rangle &= \sum_{x,y\in\Lambda} \sum_{\substack{\boldsymbol{\sigma}\in\mathcal{S}_{\Lambda \mathbf{S}x} \\ \boldsymbol{\tau}\in\mathcal{S}_{\Lambda \mathbf{S}y}}} \varphi(y,\boldsymbol{\tau})\varphi(x,\boldsymbol{\sigma}) \left\langle \Phi_{y,\boldsymbol{\tau}}, \widetilde{H}\Phi_{x,\boldsymbol{\sigma}} \right\rangle \\
&= -\sum_{x,y\in\Lambda} t_{x,y} \sum_{\boldsymbol{\sigma}\in\mathcal{S}_{\Lambda \mathbf{S}x}} \varphi(y,\boldsymbol{\sigma}_{y\to x})\varphi(x,\boldsymbol{\sigma}) \\
&\geq -\sum_{x,y\in\Lambda} t_{x,y} \left(\sum_{\boldsymbol{\sigma}\in\mathcal{S}_{\Lambda \mathbf{S}x}} \{\varphi(y,\boldsymbol{\sigma}_{y\to x})\}^2 \right)^{1/2} \left(\sum_{\boldsymbol{\sigma}\in\mathcal{S}_{\Lambda \mathbf{S}x}} \{\varphi(x,\boldsymbol{\sigma})\}^2 \right)^{1/2} \\
&= -\sum_{x,y\in\Lambda} t_{x,y}\xi_y\xi_x \\
&= \langle \Phi_\uparrow, H\Phi_\uparrow \rangle \,, \qquad (59)
\end{aligned}$$

where we have used the Schwartz inequality (with the assumption $t_{x,y} \geq 0$) to get the third line. This bound shows that $\Phi_\uparrow$ is also a ground state. This completes the proof of Theorem 15.9.

15.4.2 Nagaoka's theorem

We now state a stronger and the most general version of Nagaoka's theorem. Under an additional condition (which can be easily verified in some typical cases), we will prove that the ferromagnetic states are the only possible ground states. To determine the additional conditions, it is better to start from mathematics.

Let us recall the Perron-Frobenius theorem, which is standard in linear algebra. The following is the simplest version of the theorem[34].

Theorem 15.11 (Perron-Frobenius theorem for a real symmetric matrix)**.** *Let $M = (m_{i,j})_{i,j=1,\ldots,N}$ be an $N \times N$ real symmetric matrix (i.e., $m_{i,j} = m_{j,i} \in \mathbb{R}$) with the properties that*
i) $m_{i,j} \leq 0$ for any $i \neq j$.
ii) All $i \neq j$ are connected via nonvanishing matrix elements of M. More precisely[35], for any $i \neq j$, we can take a sequence $(i_1, \ldots, i_K)$ such that $i_1 = i$, $i_K = j$, and $m_{i_k,i_{k+1}} \neq 0$ for all $k < K$.

Then the lowest eigenvalue of M is nondegenerate and the corresponding eigenvector $\mathbf{v} = (v_i)_{i=1,\ldots,N}$ can be taken to satisfy $v_i > 0$ for all i.

Proof: Let us present a standard elementary proof based on a variational argument. The essence of the argument is that a state without "nodes" has low energy. This idea is familiar in quantum mechanics[36].

1) We first prove that if an eigenvector $\mathbf{u} = (u_i)_{i=1,\ldots,N}$ of M satisfies $u_i \geq 0$ for any i, then it inevitably satisfies $u_i > 0$ for any i. To do this, we assume the converse, i.e., $u_i \geq 0$ for all i, and $u_j = 0$ for some j. Then from the eigenvalue equation, we see that $\sum_i m_{j,i}u_i = \mu u_j = 0$. Since $m_{j,i}u_i \leq 0$, this means $u_i = 0$ for all i with $m_{j,i} \neq 0$. Because we have the connectivity ii), we can repeat this argument until we see $u_i = 0$ for all i, which is a contradiction.

2) We then prove that a normalized eigenvector $\mathbf{v} = (v_i)_{i=1,\ldots,N}$ for the lowest eigenvalue μ_0 can be chosen to satisfy $v_i > 0$ for all i. Since all $m_{i,j}$ and μ_0 are real, we can assume that the

[34] The full version of the theorem applies to non-symmetric matrices as well. See, for example, page 130 of [34].
[35] Another way of stating the condition is that, for any $i \neq j$, there is K such that $(M^K)_{i,j} \neq 0$.
[36] See, for example, Section 20 of [35].

v_i are all real. With 1) in mind, we assume the converse, i.e., $v_j > 0$ and $v_k < 0$ hold for some j and k. Define $\mathbf{u} = (u_i)_{i=1,\dots,N}$ by $u_i = |v_i|$. Then from i), we see

$$\mu_0 = \sum_{i,j=1}^{N} v_i m_{i,j} v_j \geq \sum_{i,j=1}^{N} u_i m_{i,j} u_j. \tag{60}$$

Since μ_0 is the lowest eigenvalue, the right-hand side must also be equal to μ_0, which means that $\mathbf{u}$ is an eigenvector of M. Then the above 1) implies $u_i > 0$, and hence $v_i \neq 0$ for all i. Recalling the connectivity ii), the latter property implies we can find i and j such that $v_i m_{i,j} v_j > 0$. Then we have $u_i m_{i,j} u_j < 0$ for the same i and j, which means (60) is valid with $\geq$ replaced by $>$. This contradicts with the assumption that μ_0 is the lowest eigenvalue.

3) Finally suppose that the lowest eigenvalue of M is degenerate. Then we can find two mutually orthogonal eigenvectors $\mathbf{u}$ and $\mathbf{v}$. But from 2), we must have $\mathbf{u} \cdot \mathbf{v} \neq 0$, which is a contradiction. □

We wish to apply this theorem to the present problem of the $U = \infty$ Hubbard model with a single hole. As for the matrix M, we take the matrix representation (55) of the Hamiltonian for the basis states (53) with a fixed $S^{(3)}_{\rm tot} = \sum_{y \in \Lambda \mathbf{S} x} \sigma_y$. Because of (55) and the assumption $t_{x,y} \geq 0$, the condition i) of Theorem 15.11 is satisfied. The condition ii), on the other hand, is not always valid. This motivates us to consider the following *connectivity condition*.

Definition 15.12 (Connectivity condition)**.** *A Hubbard model with $U = \infty$ and $N_{\rm e} = N_{\rm s} - 1$ (or more precisely the hopping matrix T) is said to satisfy the* connectivity condition *if all the basis states $\Phi_{x,\boldsymbol{\sigma}}$ with common $S^{(3)}_{\rm tot} = \sum_{y \in \Lambda \mathbf{S} x} \sigma_y$ are connected with each other through nonvanishing matrix elements of H.*

As we shall see in the next section, the connectivity condition can be easily verified in the Hubbard model with nearest neighbor hoppings on most standard lattices including triangular, square, simple cubic, fcc, and bcc.

Consider a model which satisfies the connectivity condition. Then we can readily apply the Perron-Frobenius theorem to see that the ground state in each subspace with a fixed $S^{(3)}_{\rm tot}$ is unique. Then Theorem 15.9 implies that this ground state must be ferromagnetic. So we have proved the following, which is the full (and most generalized) version of Nagaoka's famous theorem.

Theorem 15.13 (Generalized Nagaoka theorem)**.** *Consider an arbitrary Hubbard model with $t_{x,y} \geq 0$ for any $x, y \in \Lambda$, $N_{\rm e} = N_{\rm s} - 1$, and $U = \infty$, and further assume that the model satisfies the connectivity condition. Then the ground states have total spin $S_{\rm tot} = S_{\max} (= N_{\rm e}/2)$, and are non-degenerate apart from the trivial $(2S_{\max} + 1)$-fold degeneracy.*

As stressed earlier, the knowledge that all the ground states are ferromagnetic is of fundamental importance. The theorem is of course no longer valid for the trivial model with $t_{x,y} = 0$ for all x and y. Since the theorem asserts the nondegeneracy of the ferromagnetic ground state for $U = \infty$, the continuity of energy eigenvalues in U implies that the statement of the theorem is valid also for sufficiently large but finite U. However, we have no meaningful estimates of how large U should be.

15.4.3 Connectivity condition

We still have to verify the connectivity condition for some systems to make Theorem 15.13 meaningful. It seems, however, that to write down a simple necessary and sufficient condition

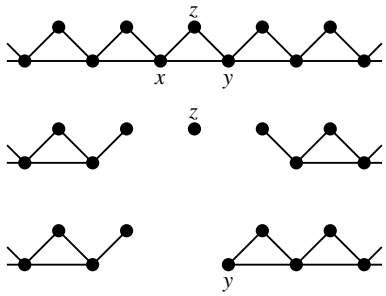


Figure 331: *Exchange bonds* in the "delta-chain." The horizontal bond $\{x, y\}$ is not an exchange bond since the site z is disconnected from the rest when x and y are removed. The bond $\{x, z\}$ is an exchange bond since the lattice remains connected after the removal of x and y provided that we use periodic boundary conditions.

for the connectivity condition is a nontrivial problem. We here follow Nagaoka's original spirit, and provide a constructive criterion (i.e., a sufficient condition) for the connectivity condition.

Let us introduce some terminology. By a *loop* of length m, we mean an ordered set $(x_1, \ldots, x_m)$ of sites such that $t_{x_i,x_{i+1}} \neq 0$ for all $i = 1, \ldots, m-1$, and $t_{x_m,x_1} \neq 0$. We say that a pair $\{x, y\}$ of lattice sites is an *exchange bond* if they belong to a loop of length three or four, and the whole lattice remains connected via nonvanishing $t_{x,y}$ when the sites x and y are removed. (See Figure 331.)

Then we have the following sufficient condition for the connectivity.

Lemma 15.14 (A sufficient condition for the connectivity)**.** *If the whole lattice is connected by exchange bonds, then the model satisfies the connectivity condition.*

This sufficient condition can be easily verified in various systems. In models defined on regular lattices like triangular, square, simple cubic, fcc, or bcc lattices with nonvanishing hopping amplitudes between nearest neighbor sites, it is obvious that all the nearest neighbor bonds are exchange bonds. Thus they trivially satisfy the condition. A less trivial example is the delta chain with periodic boundary conditions in Figure 331. Although the horizontal bonds are not exchange bonds, the whole lattice is connected via non-horizontal bonds, which are exchange bonds.

Proof of Lemma 15.14: Suppose that we are given an arbitrary configuration of $N_e = N_s - 1$ electrons on Λ. Our goal is to show that we can get an arbitrary configuration with the same $S^{(3)}_{tot}$ by moving the single hole along nonvanishing $t_{x,y}$.

Let $\{x, y\}$ be an exchange bond. We show below that we can exchange the spins at sites x and y without changing the configuration outside $\{x, y\}$. Since the whole lattice is connected via exchange bonds, this means we can generate any permutation of spin configurations by successive exchanges on the exchange bonds. This proves the connectivity condition.

We now prove the desired property of exchange bonds. Let $\{x, y\}$ be an exchange bond, and assume that x and y are occupied by electrons with opposite spins. We first bring the hole (by successive hops outside $\{x, y\}$) to a site other than x or y on the loop (of length three or four) that contains $\{x, y\}$. Next we let the hole move along the loop until the spins at x and y are exchanged. In a loop of length three, this is realized after the hole goes around the loop once, as in Figure 332. In a loop of length four[37], we have to move the hole along the loop once or twice, depending on the spin configuration, as in Fig 333. Finally, we bring the hole back to the original location. By following the same path as in the first step backwards, we recover exactly the same configuration, except on sites x and y. □

[37]The exchange on a length four loop is possible because electronic spins take (only) two values.

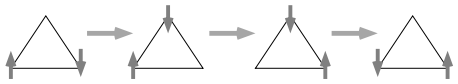


Figure 332: The two spins are exchanged when the hole hops around the loop once. Such an "exchange process along a triangle" appears repeatedly in various examples of ferromagnetism, and is regarded as a universal and fundamental mechanism leading to ferromagnetism [21].

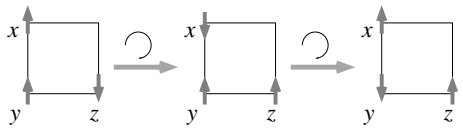


Figure 333: The spins at x and z are exchanged when the hole hops around the loop once. The spins at y and z are exchanged when the hole hops around the loop twice (or once in the opposite orientation).

15.4.4 Stability of Nagaoka's ferromagnetism

It is desirable to extend Nagaoka's ferromagnetism to systems with finite U and with finite densities of holes. Unfortunately, the Perron-Frobenius argument which works for the one-hole case fails even for models with two holes. There is a considerable number of rigorous works (including that in Nagaoka's original work [10]) which establish that saturated ferromagnetism *does not* take place in certain situations. See, for example, [36, 27, 37, 38, 39, 24]. Most of these works are essentially based on variational arguments where one constructs sophisticated variational states which have lower energies than the ferromagnetic state. (We have seen the most elementary versions of the argument in Section 15.3.3.) As far as we know, there are no rigorous results about the *stability* of Nagaoka's ferromagnetism[38] in the Hubbard model.

There is also a considerable number of numerical and theoretical works which indicate stability or instability of Nagaoka's ferromagnetism in various situations [42, 43, 44, 45].

In spite of all these efforts, it is still not known if straightforward extensions of Nagaoka's ferromagnetism to finite U and finite hole density are possible. It seems that a recent dominant opinion is that the possibility of extension depends strongly on lattice structures, where the triangular lattice is regarded as one of the most promising candidates. The importance of lattice structures is closely related to the approach from flat-band models that we discuss in Section 15.6

15.5 Ferromagnetism in the Hubbard model with long-range hopping

15.5.1 Main statement and preliminary results

There is at least one situation in which the extension of Nagaoka's ferromagnetism to finite U is possible. This is an artificial model with long range hopping that we shall now define.

We associate with each site x a constant $\lambda_x > 0$, and define the hopping matrix $\mathsf{T} = (t_{x,y})_{x,y\in\Lambda}$ with long range hopping amplitudes by[39]

$$t_{x,y} = t\lambda_x\lambda_y, \tag{61}$$

[38] Interesting lower bounds for the ground state energy of the $U = \infty$ model with multiple holes are proved in [40, 41]. We do not, however, interpret these bounds as "proofs of stability of Nagaoka's ferromagnetism" since the bounds do not rule out paramagnetism (e.g., in the trivial model with $t_{x,y} = 0$ for any x, y). See also footnote 29.

[39] In coordinate-free notation, this becomes $\mathsf{T} = t\boldsymbol{\lambda} \otimes \boldsymbol{\lambda}$, where $\otimes$ is understood as a Kronecker product, and $\boldsymbol{\lambda} = (\lambda_x)_{x\in\Lambda}$. As for this construction, the Dirac notation $\mathsf{T} = t|\boldsymbol{\lambda}\rangle\langle\boldsymbol{\lambda}|$ may be more informative.

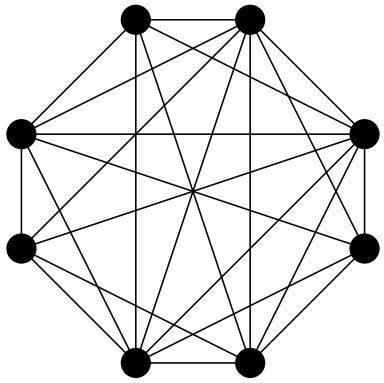


Figure 334: In the long-range hopping model, hopping between two arbitrary sites in the lattice is possible. The model serves as an intermediate step in our attempt to relate Nagaoka's ferromagnetism and flat-band ferromagnetism.

where $t > 0$ is a constant (which does not play any essential role). Note that an electron can hop from any site in the lattice to any other site (Figure 334). Then we have the following extension of Nagaoka's theorem.

Theorem 15.15 (Ferromagnetism in the long-range hopping model)**.** *Consider the Hubbard model with the hopping (61) with electron number* $N_{\rm e} = N_{\rm s} - 1$*. For any* $U > 0$*, the ground states have total spin* $S_{\rm tot} = S_{\rm max} (= N_{\rm e}/2)$*, and are non-degenerate apart from the trivial* $(2S_{\rm max} + 1)$*-fold degeneracy.*

At first glance, the theorem looks surprising and attractive since the precise statement of Nagaoka's theorem is extended to an arbitrary positive value of U. As we shall see below, this is indeed a consequence of a very special property of the present model.

With the hopping (61), the action of T onto any $\boldsymbol{\varphi} \in \mathfrak{h}$ becomes $\mathsf{T}\boldsymbol{\varphi} = t\boldsymbol{\lambda} \langle \boldsymbol{\lambda}, \boldsymbol{\varphi} \rangle$, where $\boldsymbol{\lambda} = (\lambda_x)_{x\in\Lambda} \in \mathfrak{h}$. Thus we find $\langle \boldsymbol{\varphi}, \mathsf{T}\boldsymbol{\varphi} \rangle = t \langle \boldsymbol{\varphi}, \boldsymbol{\lambda} \rangle \langle \boldsymbol{\lambda}, \boldsymbol{\varphi} \rangle = | \langle \boldsymbol{\lambda}, \boldsymbol{\varphi} \rangle |^2 \geq 0$, which means that $\mathsf{T} \geq 0$. On the other hand, the single-electron Schrödinger equation (8) becomes

$$\varepsilon \boldsymbol{\varphi} = t\boldsymbol{\lambda} \langle \boldsymbol{\lambda}, \boldsymbol{\varphi} \rangle . \tag{62}$$

This means that any $\boldsymbol{\varphi}$ orthogonal to $\boldsymbol{\lambda}$ is an eigenstate of (62) with $\varepsilon = 0$, which is the minimum possible eigenvalue. Clearly, there are $(N_{\rm s} - 1)$-fold degenerate eigenstates with $\varepsilon = 0$. The remaining eigenstate is $\boldsymbol{\lambda}$ itself, and it has the eigenvalue $\varepsilon = t \sum_{x\in\Lambda} (\lambda_x)^2 > 0$.

It is crucial for the electron number $N_{\rm e}$ to be set equal to the dimension $N_{\rm s} - 1$ of the above degeneracy. We see from (44) that the lowest energy among ferromagnetic states is $E_{\rm ferro} = \sum_{j=1}^{N_{\rm s}-1} \varepsilon_j = 0$.

On the other hand, it is clear from (35) that the lowest eigenvalue of $H_{\rm hop}$ is 0 if $N_{\rm e} \leq 2(N_{\rm s} - 1)$. Hence we can write $H_{\rm hop} \geq 0$ for $N_{\rm e} = N_{\rm s} - 1$. (See Definition 15.22 and Lemma 15.23.) We also have $H_{\rm int} \geq 0$ in general. Thus (by using Lemma 15.25) we get $H = H_{\rm hop} + H_{\rm int} \geq 0$. This, along with the fact that there is a ferromagnetic state with energy zero, implies that the ground state energy of H is 0. We have thus found that there are ferromagnetic states among the ground states.

Let $U > 0$, and let $\Phi_{\rm GS}$ be an arbitrary ground state. Then (from Lemma 15.27) $H_{\rm hop} \geq 0$, $H_{\rm int} \geq 0$, and $H\Phi_{\rm GS} = (H_{\rm hop} + H_{\rm int})\Phi_{\rm GS} = 0$ imply

$$H_{\rm hop}\Phi_{\rm GS} = 0, \quad H_{\rm int}\Phi_{\rm GS} = 0. \tag{63}$$

In other words, the ground state happens to be a simultaneous ground state[40] of $H_{\rm hop}$ and $H_{\rm int}$. This is of course a very special property of the present model with singular degeneracy in the single-electron eigenstates. Since $U \neq 0$, the second relation in (63) implies

$$\sum_{x\in\Lambda} n_{x,\uparrow} n_{x,\downarrow} = 0. \tag{64}$$

15.5.2 First proof

We give a proof of Theorem 15.15 which makes explicit use of Nagaoka's theorem. Since we have seen that there are ferromagnetic states among the ground states for $U \geq 0$, we only have to show that they are the only ground states for $U > 0$.

The ground states for any $U > 0$ are fully characterized by $H_{\rm hop}\Phi_{\rm GS} = 0$ and (64). Noting that (64) implies $H_{\rm int}\Phi_{\rm GS} = 0$ for any $U > 0$, we see that the ground states of a given $U > 0$ remain as ground states for any values of $U > 0$, and hence in the limit $U \to \infty$. Since the ground states for $U \to \infty$ are completely characterized by Nagaoka's theorem (Theorem 15.13), we have proved Theorem 15.15.

15.5.3 Second proof

We describe another proof of Theorem 15.15. This proof is certainly more involved than the first one, but provides us with an entirely different physical picture of the problem. In the new picture, ferromagnetism in this model is generated by an "exchange interaction" among the spins of electrons which are "frozen" in certain single-electron states. We recall that the first proof is based on Nagaoka's theorem, which suggests a dynamical picture that ferromagnetism is generated by the motion of the hole. It is interesting that the two totally different pictures apply equally to the present model. As we shall investigate in the next sections, the new picture will lead us to a new class of ferromagnetism now known as 'flat-band ferromagnetism.'

We start by constructing a (non-orthonormal) basis for the space of degenerate single-electron ground states. Fix an arbitrary site $x_0 \in \Lambda$, and denote by $\Lambda' = \Lambda \mathbf{S} x_0$ the lattice obtained by removing x_0 from Λ. For each $y \in \Lambda'$, we define a single-electron state $\boldsymbol{\varphi}^{(y)} = (\varphi_x^{(y)})_{x\in\Lambda} \in \mathfrak{h}$ by

$$\varphi_x^{(y)} = \begin{cases} \lambda_{x_0} & \text{if } x = y; \\ -\lambda_y & \text{if } x = x_0; \\ 0 & \text{otherwise.} \end{cases} \tag{65}$$

This state satisfies $\langle \boldsymbol{\lambda}, \boldsymbol{\varphi}^{(y)} \rangle = 0$, and hence it is a single-electron ground state. Since the states $\boldsymbol{\varphi}^{(y)}$ with $y \in \Lambda'$ are linearly independent, they span the entire space of the single-electron ground states. We define the corresponding fermion operators by $b_{y,\sigma}^\dagger = C_\sigma^\dagger(\boldsymbol{\varphi}^{(y)})$ for $y \in \Lambda'$ and $\sigma = \uparrow, \downarrow$, which satisfies $[H_{\rm hop}, b_{y,\sigma}^\dagger] = 0$ because of the general commutation rule (29) and the definition (32) of $H_{\rm hop}$.

Let us characterize an arbitrary ground state $\Phi_{\rm GS}$ for $U > 0$ in a constructive manner. Because $H_{\rm hop}\Phi_{\rm GS} = 0$, the ground state $\Phi_{\rm GS}$ should consist only of the (single-electron) zero

[40] Of course the existence of one simultaneous eigenstate does not imply $H_{\rm hop}$ and $H_{\rm int}$ are simultaneously diagonalizable. Recall that quantum mechanical angular momenta have simultaneous eigenstate with vanishing eigenvalue.

energy states. Therefore it is written as[41]

$$\Phi_{\rm GS} = \sum_{\substack{L_\uparrow, L_\downarrow \subset \Lambda' \\ \text{s.t. } |L_\uparrow|+|L_\downarrow|=N_{\rm e}}} f(L_\uparrow, L_\downarrow) \left(\prod_{y \in L_\uparrow} b^\dagger_{y,\uparrow} \right) \left(\prod_{y \in L_\downarrow} b^\dagger_{y,\downarrow} \right) \Phi_{\rm vac}, \tag{66}$$

where the $f(L_\uparrow, L_\downarrow)$ are coefficients, and $L_\uparrow$, $L_\downarrow$ are summed over all the subsets of Λ' such that $|L_\uparrow|+|L_\downarrow| = N_{\rm e} = N_{\rm s}-1$. Each state in the sum of (66) is nonvanishing because of Lemma 15.1.

Since $n_{x,\uparrow} n_{x,\downarrow} \geq 0$, the condition (64) (and Lemma 15.27) indeed implies a stronger statement that $n_{x,\uparrow} n_{x,\downarrow} \Phi_{\rm GS} = 0$ for any $x \in \Lambda$. By further rewriting this relation as $(c_{x,\downarrow} c_{x,\uparrow})^\dagger (c_{x,\downarrow} c_{x,\uparrow}) \Phi_{\rm GS} = 0$ (and using Lemma 15.28), we find that the ground state must satisfy

$$c_{x,\downarrow} c_{x,\uparrow} \Phi_{\rm GS} = 0, \tag{67}$$

for each $x \in \Lambda$.

We take $x \in \Lambda'$ and examine what (67) implies about (66). Using the anticommutation relation $\{c_{x,\sigma}, b^\dagger_{y,\sigma'}\} = \lambda_{x_0} \delta_{x,y} \delta_{\sigma,\sigma'}$, we obtain

$$\begin{aligned} c_{x,\downarrow} c_{x,\uparrow} \Phi_{\rm GS} \;=\; & \sum_{L_\uparrow, L_\downarrow} \chi[x \in L_\uparrow \cap L_\downarrow]\, f(L_\uparrow, L_\downarrow)\, {\rm sgn}[x, L_\uparrow, L_\downarrow] \times \\ & \times \left(\prod_{y \in L_\uparrow \mathbf{S} x} b^\dagger_{y,\uparrow} \right) \left(\prod_{y \in L_\downarrow \mathbf{S} x} b^\dagger_{y,\downarrow} \right) \Phi_{\rm vac}, \end{aligned} \tag{68}$$

where $\chi[\text{true}] = 1$, $\chi[\text{false}] = 0$, and the factor ${\rm sgn}[x, L_\uparrow, L_\downarrow] = \pm 1$ comes from the exchange of fermion operators. Since all the terms in the sum of (68) are linearly independent, the condition (67) implies $f(L_\uparrow, L_\downarrow) = 0$ for any $L_\uparrow, L_\downarrow$ such that $x \in L_\uparrow \cap L_\downarrow$. Using this for all $x \in \Lambda'$, we finally see that $f(L_\uparrow, L_\downarrow) = 0$ unless $L_\uparrow \cap L_\downarrow = \emptyset$. Since $|L_\uparrow| + |L_\downarrow| = |\Lambda'|$, the condition $L_\uparrow \cap L_\downarrow = \emptyset$ implies $L_\uparrow \cup L_\downarrow = \Lambda'$. This means that we can reorganize the sum (66) as

$$\Phi_{\rm GS} = \sum_{\boldsymbol{\sigma} \in \mathcal{S}_{\Lambda'}} g[\boldsymbol{\sigma}] \left(\prod_{y \in \Lambda'} b^\dagger_{y,\sigma_y} \right) \Phi_{\rm vac}, \tag{69}$$

where the sum is over all the spin configurations $\boldsymbol{\sigma} = (\sigma_y)_{y \in \Lambda}$ on Λ', and $g[\boldsymbol{\sigma}]$ is a new coefficient. The representation (69) suggests a new physical picture that the motion of the electrons is "frozen", and only spin degrees of freedom are left. Note, however, that this interpretation is somewhat arbitrary since the site x_0 is chosen arbitrarily.

Now we consider the condition (67) with $x = x_0$. By using (69) and the anticommutation relation $\{c_{x_0,\sigma}, b^\dagger_{y,\sigma'}\} = -\lambda_y \delta_{\sigma,\sigma'}$ for any $y \in \Lambda'$, we get

$$\begin{aligned} & c_{x_0,\downarrow} c_{x_0,\uparrow} \Phi_{\rm GS} \\ & = \sum_{\substack{x,y \in \Lambda' \\ \text{s.t. } x>y}} \sum_{\substack{\boldsymbol{\sigma} \in \mathcal{S}_{\Lambda'} \\ \text{s.t. } \sigma_x=\uparrow, \sigma_y=\downarrow}} {\rm sgn}[x,y]\, t \lambda_x \lambda_y \left(g[\boldsymbol{\sigma}] - g[\boldsymbol{\sigma}_{x \leftrightarrow y}] \right) \left(\prod_{z \in \Lambda' \mathbf{S} \{x,y\}} b^\dagger_{z,\sigma_z} \right) \Phi_{\rm vac}, \end{aligned} \tag{70}$$

where we have introduced arbitrary ordering in Λ', and ${\rm sgn}[x,y] = \pm 1$ is again the fermion sign[42]. The spin configuration $\boldsymbol{\sigma}_{x \leftrightarrow y}$ is obtained by switching σ_x and σ_y in the original configuration $\boldsymbol{\sigma}$.

[41] In a more careful proof, one uses Lemma 15.3 with $N_{\rm s}$ linearly independent states $\boldsymbol{\varphi}^{(y)}$ ($y \in \Lambda'$) and $\boldsymbol{\lambda}$ to construct a basis of the entire Hilbert space $\mathcal{H}_{N_{\rm e}}$, formally writes $\Phi_{\rm GS}$ as a linear combination of all the basis states, and finally uses the condition $H_{\rm hop} \Phi_{\rm GS} = 0$ to see that the terms including $C^\dagger_\sigma(\boldsymbol{\lambda})$ must be vanishing.

[42] $\Lambda \mathbf{S} \{y,z\}$ is the set obtained by removing y and z from Λ.

Since all the states in the sum (70) are linearly independent, the condition (67) is satisfied only when

$$g[\boldsymbol{\sigma}] = g[\boldsymbol{\sigma}_{x\leftrightarrow y}] \tag{71}$$

holds for all $\boldsymbol{\sigma}$ and all $x, y \in \Lambda'$. Note that we have deduced a kind of "exchange interaction" (71) among "localized spins" from the no-double-occupancy condition (67).

The exchange relation (71) implies that all the $g[\boldsymbol{\sigma}]$ with common $S^{(3)}_{\mathrm{tot}} = \sum_{y\in\Lambda'} \sigma_y$ assume exactly the same values[43]. This means that the lowest energy state in the space with a fixed $S^{(3)}_{\mathrm{tot}}$ is unique. Since we already know that there are ferromagnetic ground states, this completes the second proof of Theorem 15.15.

15.6 Flat-band ferromagnetism

15.6.1 Definition

We have seen that the Hubbard model with artificial long range hopping exhibits ferromagnetism for arbitrary $U > 0$. As we shall see in the present section, it is possible to assemble many (identical) copies of the long-range model to get Hubbard models with only short range hoppings which still exhibit ferromagnetism [7, 8]. Unlike Nagaoka's ferromagnetism, the electron density is away from half-filling, and U need only to be nonvanishing. On the other hand, these models have very singular band structure which is expressed by the name "flat-band models."

Let us define our model by means of a "cell construction[44]." Our lattice Λ can be written as

$$\Lambda = C_1 \cup \cdots \cup C_M, \tag{72}$$

where each C_j is called a cell. Each cell consists of a single *internal site* and n *external sites*, where $n \geq 2$ is a constant. The simplest cell is a triangle with $n = 2$. The models to be constructed will again involve "exchange processes along a triangle." This seems to be a universal and fundamental mechanism of ferromagnetism [21]. See Figure 332.

When assembling M cells to form Λ in (72), we identify some external sites from different cells[45] to regard them as a single site in Λ. See Figure 335. The lattice Λ is naturally decomposed as $\Lambda = \mathcal{I} \cup \mathcal{E}$, where $\mathcal{I}$ is the set of internal sites, and $\mathcal{E}$ is the set of external sites.

We define the hopping amplitudes for the j-th cell by

$$t^{(j)}_{x,y} = t\lambda^{(j)}_x \lambda^{(j)}_y, \tag{73}$$

where

$$\lambda^{(j)}_x = \begin{cases} \lambda & \text{if } x \text{ is the internal site of } C_j; \\ 1 & \text{if } x \text{ is one of the external sites of } C_j; \\ 0 & \text{otherwise.} \end{cases} \tag{74}$$

Here $t > 0$ and $\lambda > 0$ are constants. Note that this is the same as the hopping amplitudes (61) of the long range model. An essential difference, which makes the present model less artificial, is that we have this type of hopping only within each cell. The total hopping is defined as

$$t_{x,y} = \sum_{j=1}^{M} t^{(j)}_{x,y}. \tag{75}$$

[43] Although it is not necessary in the present proof, one can show directly that such a state has $S_{\mathrm{tot}} = S_{\max}$. See proof of Theorem 15.19.

[44] A class of exactly solvable Hubbard models with very similar cell structures (with paramagnetic ground states) was discovered by Brandt and Giesekus [32] before the present model for flat-band ferromagnetism was introduced. See also [33, 46] and references therein for further results concerning this model.

[45] There can be external sites left unidentified.

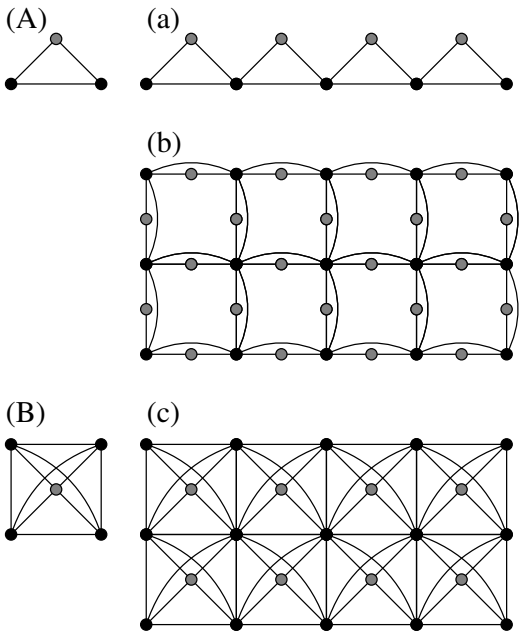


Figure 335: Examples of cells and lattices. Gray dots represent internal sites. From the triangular cell with three sites (A), one can form (a) the delta-chain by identifying two external sites, or (b) a decorated square lattice by identifying four. From the cell with five sites (B), one can form (c) another decorated square lattice. There are many similar examples in higher dimensions. The flat-band Hubbard model on these lattices exhibits ferromagnetism for any $U > 0$ when the filling factor $\nu = N_{\rm e}/(2N_{\rm s})$ is equal to (a) 1/4, (b) 1/6, or (c) 1/4.

We finally assume that the whole lattice is connected via nonvanishing $t_{x,y}$. We then define the Hubbard model on Λ with the hopping matrix given by (75).

Theorem 15.16 (Flat-band ferromagnetism)**.** *Consider the above Hubbard model with the electron number $N_{\rm e} = |\mathcal{E}|$. For any $U > 0$, the ground states have total spin $S_{\rm tot} = S_{\rm max} (= N_{\rm e}/2)$, and are non-degenerate apart from the trivial $(2S_{\rm max} + 1)$-fold degeneracy.*

Note that the existence of ferromagnetism has been established in models with finite U, finite-range hoppings, and electron densities away from half-filling. As one might guess, the straightforward (and general) methods using the Perron-Frobenius theorem (as in the proof of Nagaoka's theorem) do not work for this problem. The proof will make use of very special properties of the model. This ferromagnetism is closely related to Nagaoka's ferromagnetism (at least) via the ferromagnetism in the long-range hopping model, but it certainly belongs to a new class of ferromagnetism. This new class is now known as "flat-band ferromagnetism" for a reason which we shall explain in Section 15.6.3.

15.6.2 Examples

Clearly the cell construction leads to a wide variety of models[46]. Let us discuss a class of examples which is constructed from the most elementary triangular cell with $n = 2$.

We construct the d-dimensional $L \times \cdots \times L$ hypercubic lattice with periodic boundary conditions by assembling together the triangular cell (A) of Figure 335 regarding it as a basic bond (with length 1). The resulting lattice Λ is the decorated hypercubic lattice with $(d+1)L^d$ sites. The set $\mathcal{E}$ of external sites can be identified with the d-dimensional $L \times \cdots \times L$ hypercubic lattice with L^d sites. The internal sites in $\mathcal{I}$ are the decorating sites at the center of every bond. The lattices (a) and (b) in Figure 335 represent parts of Λ for $d = 1$ and 2, respectively.

[46]It is not necessary to have the same n or λ for all the cells. This further extends the possibility of models.

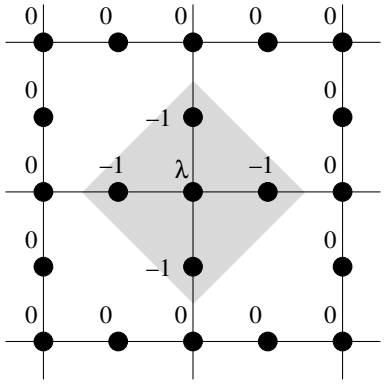


Figure 336: The components of the single-electron ground state $\boldsymbol{\varphi}^{(y)}$ in the model b) of Figure 335. The state is localized around the site y, which is at the center of the figure.

Then the hopping amplitude of (75) becomes

$$t_{x,y} = \begin{cases} \lambda t & \text{if } |x-y| = 1/2; \\ t & \text{if } x, y \in \mathcal{E} \text{ and } |x-y| = 1; \\ \lambda^2 t & \text{if } x = y \in \mathcal{I}; \\ 2dt & \text{if } x = y \in \mathcal{E}; \\ 0 & \text{otherwise.} \end{cases} \tag{76}$$

Note that there are next-nearest-neighbor hoppings between nearby external sites.

The electron number specified in Theorem 15.16 is $N_{\mathrm{e}} = L^d$. In terms of the filling factor, this corresponds to $\nu = N_{\mathrm{e}}/(2N_{\mathrm{s}}) = \{2(d+1)\}^{-1}$.

15.6.3 Single-electron problem

The present models have peculiar degeneracy in their single-electron ground states which are quite similar to that in the long range hopping model of the previous section.

We first note that the hopping matrix T defined by (75) satisfies $\mathsf{T} \geq 0$. This is because each $(t^{(j)}_{x,y})_{x,y\in\Lambda}$ determined by (73) is positive semidefinite (which fact is clear from the discussion in Section 15.5.1), and the sum of positive semidefinite matrices is positive semidefinite (as in Lemma 15.25).

We define for each $y \in \mathcal{E}$ the single-electron state $\boldsymbol{\varphi}^{(y)} = (\varphi^{(y)}_x)_{x\in\Lambda}$ by

$$\varphi^{(y)}_x = \begin{cases} \lambda & \text{if } x = y; \\ -1 & \text{if } x \text{ is the internal site of one of the cells containing } y; \\ 0 & \text{otherwise.} \end{cases} \tag{77}$$

See Figure 336. Observe that we have $\left\langle \boldsymbol{\lambda}^{(j)}, \boldsymbol{\varphi}^{(y)} \right\rangle = 0$ for any $j = 1, \dots, M$ and $y \in \mathcal{E}$, where $\boldsymbol{\lambda}^{(j)} = (\lambda^{(j)}_x)_{x\in\Lambda} \in \mathfrak{h}$, and hence

$$\mathsf{T}\boldsymbol{\varphi}^{(y)} = t \sum_{j=1}^{M} \boldsymbol{\lambda}^{(j)} \left\langle \boldsymbol{\lambda}^{(j)}, \boldsymbol{\varphi}^{(y)} \right\rangle = 0. \tag{78}$$

Since $\boldsymbol{\varphi}^{(y)}$ with $y \in \mathcal{E}$ are linearly independent, the single-electron Schrödinger equation (8) has $|\mathcal{E}|$-fold degenerate ground states[47] with $\varepsilon = 0$. It is needless to say that such a huge degeneracy

[47] It is not hard to prove that $\boldsymbol{\lambda}^{(j)}$ with $j = 1, \dots, M$ span the remaining space with $\varepsilon > 0$.

is accidental, and makes the present models very special. We also stress that such a degeneracy is lifted by an arbitrarily small generic perturbation applied to the hopping matrix.

In models with translational invariance, the degeneracy in the single-electron ground states corresponds to the lowest band being completely dispersionless, or "flat." For example, in the models on the d-dimensional hypercubic lattice of Section 15.6.2, we find that there are $d+1$ bands with dispersion relations

$$\varepsilon_j(k) = \begin{cases} 0 & \text{if } j = 1; \\ \lambda^2 t & \text{if } j = 2, \dots, d; \\ (\lambda^2 + 2d)t + t \sum_{\mu=1}^{d} \cos k_\mu & \text{if } j = d+1. \end{cases} \tag{79}$$

Note that the lowest band and the $(d-1)$ middle bands are completely flat, while the upper band is dispersive. For definitions of bands and dispersion relations as well as an explicit calculation for the $d = 1$ case, see Appendix 15.9.5.

15.6.4 Proof

We now prove Theorem 15.16. Interestingly, the proof is completely analogous to and not more difficult than the second proof of the (less attractive) Theorem 15.15 for the long range hopping model.

Since the single-electron ground states of the model are $|\mathcal{E}|$-fold degenerate, and the electron number is set to $N_{\rm e} = |\mathcal{E}|$, all the preliminary considerations in Sections 15.5.1 and 15.5.3 apply to the present model as well. Any ground state $\Phi_{\rm GS}$ satisfies the conditions (63), and (67) for any $x \in \Lambda$.

Again introducing $b^\dagger_{y,\sigma} = C^\dagger_\sigma(\boldsymbol{\varphi}^{(y)})$ for $y \in \mathcal{E}$, and using $H_{\rm hop}\Phi_{\rm GS} = 0$, we can represent any ground state as

$$\Phi_{\rm GS} = \sum_{\substack{L_\uparrow, L_\downarrow \subset \mathcal{E} \\ \text{s.t. } |L_\uparrow|+|L_\downarrow| = N_{\rm e}}} f(L_\uparrow, L_\downarrow) \left(\prod_{y \in L_\uparrow} b^\dagger_{y,\uparrow} \right) \left(\prod_{y \in L_\downarrow} b^\dagger_{y,\downarrow} \right) \Phi_{\rm vac}, \tag{80}$$

where $f(L_\uparrow, L_\downarrow)$ are coefficients, and $L_\uparrow$, $L_\downarrow$ are summed over all the subsets of $\mathcal{E}$ such that $|L_\uparrow| + |L_\downarrow| = N_{\rm e} = |\mathcal{E}|$.

Exactly as before, the condition (67) for $x \in \mathcal{E}$ shows that there can be no "double occupancies" in $b^\dagger$, i.e., $f(L_\uparrow, L_\downarrow) = 0$ whenever $L_\uparrow \cup L_\downarrow \neq \emptyset$. We again reorganize the sum to get a spin system representation

$$\Phi_{\rm GS} = \sum_{\boldsymbol{\sigma} \in \mathcal{S}_\mathcal{E}} g[\boldsymbol{\sigma}] \left(\prod_{y \in \mathcal{E}} b^\dagger_{y,\sigma(y)} \right) \Phi_{\rm vac}, \tag{81}$$

where the sum is over all the spin configuration $\boldsymbol{\sigma} = (\sigma_y)_{y \in \mathcal{E}}$ on $\mathcal{E}$, and $g[\boldsymbol{\sigma}]$ is a new coefficient.

Next we examine the implication of the condition (67) for $x \in \mathcal{I}$ on the ground state (81). Unlike in the long range hopping model, there are still many relations to use. Let $\mathcal{E}_x$ be the set of external sites contained in the cell which contains an internal site x. By using (81) and the anticommutation relation $\{c_{x,\sigma}, b^\dagger_{y,\sigma'}\} = -\delta_{\sigma,\sigma'}\chi[y \in \mathcal{E}_x]$ for any $x \in \mathcal{I}$ and $y \in \mathcal{E}$, we get

$$c_{x,\downarrow}c_{x,\uparrow}\Phi_{\rm GS} = \sum_{\substack{y,z \in \mathcal{E}_x \\ \text{s.t. } y > z}} \sum_{\substack{\boldsymbol{\sigma} \in \mathcal{S}_\mathcal{E} \\ \text{s.t. } \sigma_y = \uparrow, \sigma_z = \downarrow}} \operatorname{sgn}[x, y] (g[\boldsymbol{\sigma}] - g[\boldsymbol{\sigma}_{y \leftrightarrow z}]) \left(\prod_{u \in \mathcal{E} \mathbf{S}\{y,z\}} b^\dagger_{u,\sigma_u} \right) \Phi_{\rm vac}. \tag{82}$$

Since this quantity vanishes for all $x \in \mathcal{I}$, we finally find that

$$g[\boldsymbol{\sigma}] = g[\boldsymbol{\sigma}_{y \leftrightarrow z}] \tag{83}$$

for any $y, z \in \mathcal{E}$ which belong to a common cell. Note that the "exchange interaction" is short range because we are treating models with short range hoppings. Since the entire lattice is connected, (83) ensures that the lowest energy state is unique in each sector with a fixed $S^{(3)}_{\rm tot}$. This completes the proof of Theorem 15.16.

15.6.5 Mielke's flat-band ferromagnetism

A slightly different class of Hubbard models which have highly degenerate single-electron ground states and exhibit ferromagnetism was discovered by Mielke [14, 15] *before* the above models were introduced. Here we briefly summarize Mielke's beautiful construction based on graph theoretic notion. For a general theory of flat-band ferromagnetism obtained by Mielke, see [47].

We start from abstract notation. Let $G = (V, E)$ be an abstract graph, where V is the set of vertices (sites) $\alpha, \beta, \ldots \in V$, and E is the set of edges (bonds) which are nothing but pairs of vertices like $\{\alpha, \beta\}$. Given a graph G, one can construct the corresponding line graph $L(G) = (V_{\rm L}, E_{\rm L})$ by the following procedure. The set of vertices (sites) $V_{\rm L}$ (whose elements are denoted as $x, y, \ldots \in V_{\rm L}$) is taken to be identical to the set E. This means that we identify edges in G with the vertices (sites) in $L(G)$ as, for example, $x = \{\alpha, \beta\}$, $y = \{\alpha, \gamma\}$, etc. Next we declare that two vertices $x, y \in V_{\rm L}$ are adjacent to each other if the corresponding two edges in E share a common vertex in V. The vertices x and y in the above example are adjacent to each other since the corresponding edges in E have a common vertex α. $E_{\rm L}$ is the set of edges (bonds) in $L(G)$, which consists of all the adjacent pairs (like $\{x, y\}$) of vertices (sites) in $V_{\rm L}$. Finally we set $M(G) = |E| - |V| + 1$ if G is bipartite[48], and $M(G) = |E| - |V|$ if G is non-bipartite.

We define the Hubbard model on the line graph $L(G)$. With each site $x \in V_{\rm L}$, we associate the fermion operator $c_{x,\sigma}$, and consider the Hamiltonian

$$H = t \sum_{\substack{\{x,y\} \in E_{\rm L} \\ \sigma = \uparrow, \downarrow}} (c^\dagger_{x,\sigma} c_{y,\sigma} + c^\dagger_{y,\sigma} c_{x,\sigma}) + U \sum_{x \in V_{\rm L}} n_{x,\uparrow} n_{x,\downarrow}, \tag{84}$$

where $t > 0$ is a constant. Then the main result of [14] is the following.

Theorem 15.17 (Mielke's flat-band ferromagnetism)**.** *Suppose that the graph G is twofold connected*[49]*. Consider the above Hubbard model with $N_{\rm e} = M(G)$. Then for any $U > 0$, the ground states of the model have total spin $S_{\rm tot} = S_{\rm max} (= N_{\rm e}/2)$, and are nondegenerate apart from the trivial $(2S_{\rm max} + 1)$-fold degeneracy.*

This theorem applies to the Hubbard model defined on various line graphs, a typical one being the kagomé lattice of Figure 337.

15.6.6 Beyond flat-band ferromagnetism

Let us once again take a look at the mechanism by which ferromagnetism is generated in a flat-band model. Reflecting the bulk degeneracy in the single-electron ground states, the ground

[48] G is bipartite if it can be decomposed into two disjoint sublattices as $G = A \cup B$ with the property that any edge in E joins a vertex in A with a vertex in B.

[49] A graph is twofold connected if and only if one cannot make it disconnected by the removal of a single vertex.

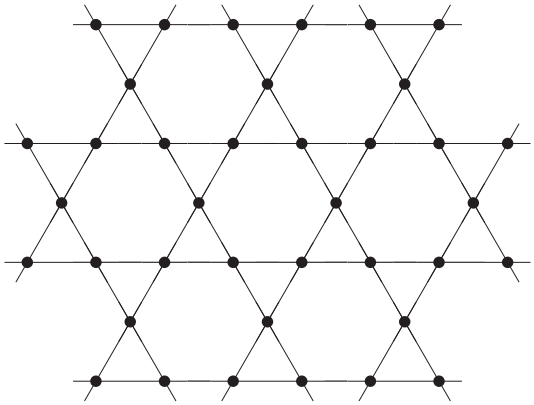


Figure 337: The kagomé lattice is the line graph of the hexagonal lattice. Mielke showed that the Hubbard model on the kagomé lattice exhibits ferromagnetism when the filling factor is $\nu = 1/6$ for any $U > 0$. This is the most beautiful example of flat-band ferromagnetism.

states of the non-interacting model with $H = H_{\rm hop}$ are highly degenerate. There are ground states for any possible values of total spin $S_{\rm tot}$, including the smallest $S_{\rm tot} = 0$ or $1/2$ and the maximum $S_{\rm tot} = S_{\rm max} (= N_{\rm e}/2)$. When we introduce the Hubbard interaction $H_{\rm int}$, the energy of the ferromagnetic ground states (of $H_{\rm hop}$) do not change since ferromagnetic states *do not feel* on-site repulsion. These observations are indeed trivial. A truly nontrivial point (which is proved by imposing some conditions on models) is that all the other ground states of $H_{\rm hop}$ receive extra energy from $H_{\rm int}$, and the ferromagnetic ground states become the only ground states of $H = H_{\rm hop} + H_{\rm int}$.

In this sense, the flat-band ferromagnetism certainly takes into account nontrivial interplay of $H_{\rm hop}$ and $H_{\rm int}$. However, there is no true "competition" between $H_{\rm hop}$ and $H_{\rm int}$. The ferromagnetic ground states are already present in the $U = 0$ model among highly degenerate ground states. The only role of the interaction is to lift the degeneracy, and "select" the ferromagnetic ground states as the only ground states. This is why flat-band ferromagnetism takes place for any $U > 0$.

As we have seen in Theorem 15.5, we never realize saturated ferromagnetism for values of U which are too small in a system without bulk degeneracy in the single-electron ground states. Of course, the bulk degeneracy in a flat-band model is far from being robust, and it is easily destroyed by adding to the hopping matrix T an arbitrarily small generic perturbation. Then, an essential question is whether ferromagnetism found in a flat-band model remains stable after adding a small perturbation to T which makes the lowest band "nearly flat." If the ferromagnetism were unstable against perturbations, we would have to conclude that flat-band ferromagnetism is a mere mathematical game.

It was conjectured [7, 8] that the flat-band ferromagnetism is stable against small perturbations to the hopping Hamiltonian. Among the main arguments used in that discussion was that an approximate low-energy effective Hamiltonian of the flat-band model has the precise form of the ferromagnetic Heisenberg spin system. Kusakabe and Aoki [48, 49] made the first systematic study about stability of flat-band ferromagnetism. They argued that the flat-band models possess spin-wave excitations which have healthy dispersions, and this fact guarantees the robustness of the flat-band ferromagnetism. They also found numerical evidence that ferromagnetism remains stable if we add sufficiently small perturbations to the hopping matrix, thus making the lowest flat band "nearly flat."

As for rigorous results, stability of ferromagnetism under a single-spin flip is proved in [16, 17] for the model obtained by adding an *arbitrary* small translation invariant perturbation to the hopping matrix of the translation invariant flat-band Hubbard model of Section 15.6.1. Although this only establishes the local stability of flat-band ferromagnetism, it is a remarkable

nonperturbative result which applies to a robust class of models. Given the global stability of ferromagnetism in the original flat-band model, the local stability of ferromagnetism in a nearly-flat-band model provides very strong evidence that the nearly-flat-band model indeed exhibits globally stable ferromagnetism.

Finally, the problem of stability of ferromagnetism was completely solved in [18, 19] for a special class of models. Let us explain the results briefly.

Take the same lattice Λ as in Section 15.6.2, i.e., the d-dimensional decorated hypercubic lattice. We define special perturbations to the hopping amplitudes as follows. For each $z \in \mathcal{E}$, we let

$$s_{x,y}^{(z)} = -s\mu_x^{(z)}\mu_y^{(z)}, \tag{85}$$

where

$$\mu_x^{(z)} = \begin{cases} \lambda & \text{if } x = z; \\ 1 & \text{if } x \text{ is an internal site adjacent to} z; \\ 0 & \text{otherwise.} \end{cases} \tag{86}$$

Here $s > 0$ is a new constant, and $\lambda > 0$ is the same as in Section 15.6.1.

Then the hopping matrix $\mathsf{T} = (t_{x,y})_{x,y\in\Lambda}$ of the perturbed models is defined by[50]

$$t_{x,y} = \sum_{j=1}^{M} t_{x,y}^{(j)} + \sum_{z\in\mathcal{E}} s_{x,y}^{(z)} + (\lambda^2 s - 2dt)\delta_{x,y}, \tag{87}$$

where $t_{x,y}^{(j)}$ is the same as in (73). This $t_{x,y}$ can be written more explicitly as

$$t_{x,y} = \begin{cases} \lambda(t+s) & \text{if } |x-y| = 1/2; \\ t & \text{if } x,y \in \mathcal{E} \text{ and } |x-y| = 1; \\ -s & \text{if } x,y \in \mathcal{I} \text{ and } |x-y| = 1 \text{ or } 1/\sqrt{2}; \\ (\lambda^2 - 2d)t + (\lambda^2 - 2)s & \text{if } x = y \in \mathcal{I}; \\ 0 & \text{otherwise.} \end{cases} \tag{88}$$

As can be seen also from Figure 338, this model contains nearest-neighbor hoppings and some next-nearest-neighbor hoppings (but not more than that). The amplitudes for the different hoppings and the on-site potential must satisfy special relations because there are only three controllable parameters, $t > 0$, $s > 0$ and $\lambda > 0$.

The single-electron energy bands of the model can be easily obtained. There are $(d+1)$ bands with dispersion relations

$$\varepsilon_j(k) = \begin{cases} -2d(t+s) - 2s\sum_{\mu=1}^{d}\cos k_\mu & \text{if } j = 1; \\ \lambda^2(t+s) - 2dt & \text{if } j = 2,\cdots,d; \\ \lambda^2(t+s) + 2t\sum_{\mu=1}^{d}\cos k_\mu & \text{if } j = d+1. \end{cases} \tag{89}$$

If $d > 1$, the middle bands with $j = 2,\cdots,d$ have constant values (i.e., are flat and degenerate), reflecting the geometry of the decorated hypercubic lattice[51]. For us it is crucial that the most important lowest band has a healthy non-constant dispersion.

[50] The term with the Kronecker delta is included only to make the expression (88) simpler and is not essential.

[51] It is possible to design short range perturbations to T which make these bands non-flat and non-degenerate, while allowing Theorem 15.18 to hold.

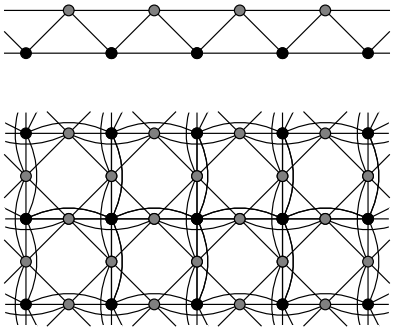


Figure 338: The d-dimensional decorated hypercubic lattice with the nearest neighbor and the next nearest neighbor hoppings for $d = 1$ and 2. These models are obtained by adding special perturbations to the models (a) and (b) in Figure 335, respectively. In a range of parameters, we can prove that the corresponding Hubbard model with sufficiently large but finite U exhibits ferromagnetism even when the lowest band is not flat.

We consider the Hubbard model on Λ with the above $t_{x,y}$ and electron number $N_{\rm e} = |\mathcal{E}| = L^d$. If $s = 0$, the model exhibits the flat-band ferromagnetism of Theorem 15.16 for any $U > 0$. When $s > 0$ and the lowest band is no longer flat, the model exhibits Pauli paramagnetism for $U = 0$. Moreover, from the elementary variational estimate of Theorem 15.5, we find that the model does not exhibit ferromagnetism when $U < 4ds$. We can say that there is a true "competition" between $H_{\rm hop}$ and $H_{\rm int}$ in this model. If the model exhibits ferromagnetism, it must be in a non-perturbative region with sufficiently large U. The following theorem establishes such a non-perturbative statement in the case that the lowest band is sufficiently flat.

Theorem 15.18 (Ferromagnetism in non-singular Hubbard models)**.** *Consider the above Hubbard model with electron number $N_{\rm e} = |\mathcal{E}| = L^d$, and let*[52] *$\lambda > 0$. If s/t is sufficiently small, and U/t is sufficiently large, the ground states have total spin $S_{\rm tot} = S_{\rm max} (= N_{\rm e}/2)$, and are non-degenerate apart from the trivial $(2S_{\rm max} + 1)$-fold degeneracy.*

For given $d \geq 1$ and $\lambda > 0$, the range of s/t and U/t in which the theorem holds can be determined by finite calculations which can be (in principle) executed by a computer. This means that one can construct a computer-aided proof if one desires. We have only performed very elementary calculations for $d = 1$ with a personal computer. When we set $\lambda = \sqrt{2}$ (in which case $t_{x,x} = 0$ for all x), the theorem is valid, for example[53], if $s/t \leq 0.4$ and $U/t \geq 40$.

Although the model is rather artificial, this is the first rigorous example of saturated ferromagnetism in a non-singular Hubbard model in which we must deal with the competition between $H_{\rm int}$ and $H_{\rm hop}$. If we further have $s/t \ll 1$, $U/t \gg 1$ and $\lambda \gg 1$, it has also been proved that low-lying excitation above the ground state has a normal dispersion relation of spin-wave excitations [18, 17]. Starting from a Hubbard model of itinerant electrons, the existence of a "healthy" ferromagnetism has been established rigorously.

Recently, Penc, Shiba, Mila, and Tsukagoshi [50] made a systematic study of related one-dimensional Hubbard models and found various pieces of evidence suggesting that the ferromagnetism in these models are indeed robust.

[52] In [18], we presented a complete proof for $d = 1$ which, however, assumed $\lambda > \lambda_{\rm c} > 0$. The improved proof in [19] applies to general d and only requires $\lambda > 0$.

[53] For sufficiently large U, the theorem is valid for $t = 1.6s$, in which case the lower band occupies more than 1/4 of the whole range of the single-electron energy spectrum (including the gap). In this case the lower band may not even be "nearly-flat".

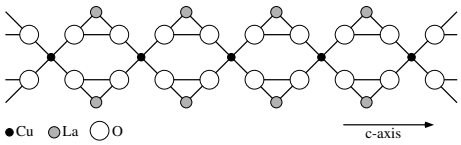


Figure 339: A chain of Cu, O, and La formed along the c-axis of La422. One finds a triangular structure formed by La and O, which reminds us of the (nearly-)flat-band models of Section 15.6.

15.7 Possible experimental realizations of (nearly-)flat-band ferromagnetism

15.7.1 Ferromagnetism in $La_4Ba_2Cu_2O_{10}$

In 1990 (when the only example of saturated ferromagnetism in the Hubbard model was that of Nagaoka), Mizuno, Masuda, Hirabayashi, Tanaka, Hasegawa and Mizutani [51, 52] reported that the tetragonal cuprate $La_4Ba_2Cu_2O_{10}$ (also called La_2BaCuO_5), which we shall abbreviate as La422, exhibits ferromagnetism[54]. A series of experiments [51, 52, 53] has revealed that La422 is an ideal insulating ferromagnet with Curie temperature 5.2 K, where most of the magnetic moment comes from the spin 1/2 moments of the Cu^{2+} ions.

The origin of the ferromagnetic coupling between Cu-ions within the ab-plane of La422 is expected to be described by the GKA rules [52, 54], which involve exchange interactions between excited orbital states and ground orbital states. Ferromagnetic coupling along the c-axis seems to require more careful treatment. Along the c-axis, La422 forms a characteristic chain structure consisting of Cu, O, and La. This is shown in Figure 339. Interestingly enough, we see La and O forming a triangular structure similar to those in the (nearly-)flat-band models of Section 15.6. We stress that there are experimental results which strongly indicate that the existence of La is essential for the occurrence of ferromagnetism. One such result is the observation that the cuprate Nd_2BaCuO_5, which has the same crystal structure as La422, exhibits antiferromagnetic ordering [55, 56]. It is pointed out that the essential difference is that Nd has a magnetic moment, while La does not. This point is supported by the result [57] that ferromagnetism remains stable upon partial replacement of La in La422 by Eu, which has no magnetic moment. Another such result is from NMR measurement in La422 that there is a large hyperfine field at La sites [58]. This suggests that the magnetic moments are not sharply localized at Cu-sites, but partially exist on La-sites as well (exactly as in the (nearly-)flat-band models).

On the other hand, a band calculation based on the Local Density Approximation (LDA) [59] revealed that two bands of La422 near the fermi level are almost flat[55].

These observations about the lattice structure and the band structure may be regarded as clues that the ferromagnetism in La422 is related to the ferromagnetism in the Hubbard model with (nearly-)flat-band. As far as the present author knows, this possible relation was first realized by Hirabayashi [60].

In fact it is not hard to construct a flat-band Hubbard model on the lattice of Figure 339 where hopping amplitudes are nonvanishing only on the bonds in the figure. We can prove that the model exhibits ferromagnetism when the number of electrons is identical to the number of Cu-sites in the lattice.

This theoretical result, however, seems to be almost irrelevant to the ferromagnetism in

[54] Ferromagnetism was discovered while investigating the superconductor $LaBa_2Cu_3O_{7-y}$.

[55] In LDA calculations, the Coulomb interaction between electrons is supposed to be taken into account in a certain self-consistent manner. Therefore the flatness of two bands in the LDA calculation does not directly indicate that the corresponding tight-binding model has a flat band.

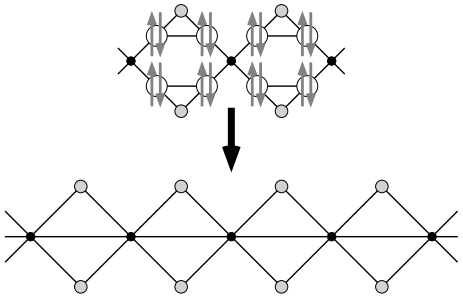


Figure 340: In the second order perturbation in $1/\epsilon_O$, the model of Figure 339 reduces to a Hubbard model on the Cu-La chain in the figure with the electron number equal to the number of the Cu-sites. The new model falls into the class of (nearly-)flat-band ferromagnetism for suitable hopping amplitudes. This observation provides a crude link between the ferromagnetism in La422 and the (nearly-)flat-band ferromagnetism.

La422. Analysis of the orbital structures of La422 [54] suggests that a simplified[56] but reasonable tight-binding description of the system is given by the lattice structure of Figure 339 with one non-degenerate orbit per each site, and with the electron number equal to the number of Cu-sites plus the twice the number of O-sites. (In other words, originally Cu orbits are singly occupied, O orbits are doubly occupied, and La orbits are empty.) Then the resulting ferromagnetism is expected to be a "partial ferromagnetism" in which only spins from Cu-sites align ferromagnetically. Unfortunately, we still do not know how to treat such partial ferromagnetism theoretically.

Nevertheless we believe that there is a deep relation between the ferromagnetism in La422 and that in Hubbard models with (nearly-)flat-bands. It is a challenging problem to further investigate this relation both from the above described tight-binding model[57] and from a continuum model.

If one wishes to see a more direct connection between the ferromagnetism in La422 and the (nearly-)flat-band ferromagnetism, one may try reducing the above Hubbard model to a simpler one by relying on some approximate argument. For example, one can consider the limit in which the on-site potential ϵ_O for O-sites (which is $t_{x,x}$, with x being an O-site) is negative, and its absolute value is much larger than other $|t_{x,y}|$. In this limit, all the O-sites are doubly occupied, and electrons (including those on Cu- and La-sites) cannot move on the lattice. If we consider a perturbation expansion in $1/\epsilon_O$ up to second order, there arise effective hoppings between Cu-sites and La-sites, and between nearby Cu-sites. The model then reduces to an effective model with Cu-sites and La-sites, as in Figure 340, with the electron number equal to the number of the Cu-sites. For suitable values of hopping amplitudes, the model falls into the class of (nearly-)flat-band ferromagnetism of Section 15.6. This observation provides a crude link between the ferromagnetism in La422 and the (nearly-)flat-band ferromagnetism.

15.7.2 Possible ferromagnetism in quantum wires

Recent developments in nanotechnology are making it possible to design various atomic scale structures which should show nontrivial quantum behavior. In [61], Watanabe, Ichimura, Onogi, Ono, Hashizume, and Wada made a theoretical study of the electronic properties of Ga adsorbates around dangling-bond wires on an H-terminated Si surface, and pointed out the possibility of ferromagnetism. The effective tight-binding model (which is not quite faithful to the actual

[56] This model is a slight simplification of the tight-binding model discussed in [54].

[57] Apart from numerical experiments (which are always possible), systematic perturbative calculations of "effective Hamiltonian" along the line of [50, 54] are also possible and may be enlightening.

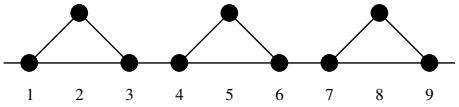


Figure 341: The effective tight-binding model for the "quantum wire." It is possible to design Hubbard model with the same structure which shows flat-band ferromagnetism [62].

lattice structure) used in [61] has the lattice and hopping structure shown in Figure 341 and has electron number[58] $N_{\rm e} = N_{\rm s}/3$.

It is possible to construct Hubbard models with exactly the same hopping structure so that the models have flat lowest bands [62]. Let the lattice be $\Lambda = \{1, 2, \ldots, N_{\rm s}\}$, where $N_{\rm s}$ is a multiple of three, and impose periodic boundary conditions. The Hamiltonian is then

$$\begin{aligned} H &= t\sum_{j=1}^{N_{\rm s}/3}\sum_{\sigma=\uparrow,\downarrow}(\lambda c_{3j-1,\sigma}+c_{3j-2,\sigma}+c_{3j,\sigma})^{\dagger}(\lambda c_{3j-1,\sigma}+c_{3j-2,\sigma}+c_{3j,\sigma}) \\ &\quad +s\sum_{j=1}^{N_{\rm s}/3}\sum_{\sigma=\uparrow,\downarrow}(c_{3j,\sigma}+c_{3j+1,\sigma})^{\dagger}(c_{3j,\sigma}+c_{3j+1,\sigma}) \\ &\quad +U\sum_{j=1}^{N_{\rm s}/3}(n_{3j-2,\uparrow}n_{3j-2,\downarrow}+n_{3j,\uparrow}n_{3j,\downarrow})+U'\sum_{j=1}^{N_{\rm s}/3}n_{3j-1,\uparrow}n_{3j-1,\downarrow}, \end{aligned} \tag{90}$$

with $t, s, U, U' > 0$ and $\lambda \neq 0$. This model has a flat lowest band with energy 0. For electron number $N_{\rm e} = N_{\rm s}/3$, one can prove that the model exhibits ferromagnetism [62]. Since the discussion about the robustness of the flat-band ferromagnetism applies to the present models as well, the present ferromagnetism is also expected to be stable against perturbations.

15.8 Towards metallic ferromagnetism

15.8.1 Conjectures and some evidence

Metallic ferromagnetism is a fascinating phenomenon in which electrons exhibiting ferromagnetism contribute to electric conductivity as well. Whether such simple models as the single-orbital Hubbard model can describe metallic ferromagnetism is an unsolved and intriguing problem.

Nagaoka's ferromagnetism is certainly motivated by metallic ferromagnetism, and we believe it reveals some important aspects of a possible mechanism of metallic ferromagnetism. But for the models where the theorem is proved, the only dynamical freedom comes form the motion of the single "hole." We cannot expect the single hole to contribute to appreciable electric current in a bulk system.

In the flat-band models of Section 15.6.1 and the related nearly-flat band models of Section 15.6.6, the existence of ferromagnetism is proved for special electron numbers. These electron numbers correspond to the half-filling of the lowest bands, but since the ground states are ferromagnetic, the lowest bands become effectively fully filled. Then the systems should behave as (Mott) insulators.

In the flat-band models of Section 15.6.2 defined in dimensions greater than one, the existence of ferromagnetism is proved [7, 8] for (not too small) electron numbers less than $|\mathcal{E}|$.

[58] To be precise, 5/6-filling with $N_{\rm e} = 5N_{\rm s}/3$ was considered in [61]. This electron number is changed into $N_{\rm e} = N_{\rm s}/3$ by using the hole-particle transformation. (See Appendix 15.9.2.) One should note that the hole-particle transformation does not only change the signs of the hopping amplitudes, but it also introduces a difference in on-site potentials when $U \neq U'$ in the Hamiltonian (90).

There is a similar rigorous result for Mielke's model in two dimensions[15]. We expect the models to describe a kind of metallic ferromagnetism, but the situation is not clear because of pathological degeneracy in the (many-body) ground states. In any case, the flat-band models, which are quite useful in describing the origin of exchange interaction in certain systems, are too singular to discuss electric conductivity.

Promising candidates of simple models exhibiting metallic ferromagnetism are the nearly-flat-band models (obtained by adding perturbations to the flat-band models) at filling factor $\nu = N_{\rm e}/(2N_{\rm s})$ different from $\nu_0 = |\mathcal{E}|/(2N_{\rm s})$. For the case $\nu < \nu_0$, the approximate projection method (similar to those described in [8, 17], but based on a local Wannier basis) indicates that the low energy effective theory of these models is represented by the ferromagnetic t-J models in any dimension [19]. (The definition of the ferromagnetic t-J model can be found in Section 15.8.2.) This observation leads us to the conjecture that these models exhibit metallic ferromagnetism at least when $\nu < \nu_0$ is not too small and the model is sufficiently close to a flat-band model.

Penc, Shiba, Mila, and Tsukagoshi [50] made a detailed study of related problems in one dimension, and found both theoretical and numerical evidences that the model exhibits metallic ferromagnetism for arbitrary filling factor within the range $0 < \nu < \nu_0 = 1/4$. That the ferromagnetism is stable for an arbitrarily small density reflects the special character of one-dimensional systems as noted in Section 15.3.3. Sakamoto and Kubo [63] also found strong numerical evidence that related one-dimensional models exhibit metallic ferromagnetism for $0 < \nu < \nu_0$.

Watanabe and Miyashita [64, 65] treated the one-dimensional flat-band Hubbard model of Section 15.6.2 for a filling factor ν in the range $1/2 > \nu > \nu_0 = 1/4$. Since the lower band is totally filled for such electron densities, the flatness of the lower band is regarded as irrelevant. Note that the case with $N_{\rm e} = N_{\rm s} - 1$ is nothing but Nagaoka's ferromagnetism; we have come back to Nagaoka's ferromagnetism following a different path! Watanabe and Miyashita found numerical evidence that the model exhibits metallic ferromagnetism for all densities they considered.

A different (but in some sense related) candidate for a model with metallic ferromagnetism is the so-called t-t' model. It is defined on the one-dimensional lattice $\Lambda = \{1, \ldots, N_{\rm s}\}$, and its hopping amplitudes are defined by $t_{x,x+1} = t_{x+1,x} = -t$, $t_{x,x+2} = t_{x+2,x} = t'$ for all $x \in \Lambda$ and $t_{x,y} = 0$ otherwise. From a first order perturbation theory, it was suggested [66, 67] that the model with $U = \infty$ exhibits ferromagnetism if $t' > 0$. By considering a continuum limit theory, Müller-Hartmann [68] argued that the model exhibits metallic ferromagnetism when $4t' > |t| > 0$, $U = \infty$, and the electron density is sufficiently low. There is numerical evidence that the t-t' model exhibits ferromagnetism [26, 69].

Kohno [70] discussed the possibility of metallic ferromagnetism in the Hubbard model on a ladder. Both perturbation theory and numerical calculations suggest that metallic ferromagnetism appears for electron numbers satisfying $N_{\rm s}/2 < N_{\rm e} < N_{\rm s}$. It is interesting that, in this model, itinerant electrons in upper bands play essential roles in generating ferromagnetism.

Unfortunately, none of the above conjectures have yet been confirmed rigorously[59].

[59] In a fermion system on a finite lattice, a formal perturbation series alway converges because the operators are finite dimensional. Then one might think that the first order perturbation theories of [66, 67, 70] imply weak rigorous results that a finite model exhibits ferromagnetism for sufficiently small t'/t (or $t_{\parallel}/t_{\perp}$). However, this is not the case since there are no estimates of the energy difference between the ground state and the first excited state, and there remains a possibility that the higher order perturbations change the nature of the ground state for any small values of the expansion parameter. The situation is different from that of Nagaoka's ferromagnetism, where the result for $U = \infty$ automatically implies the same result for sufficiently large U (in a finite system). For an example of rigorous (and general) perturbation theory for quantum many-body systems, see [71].

15.8.2 Ferromagnetic t-J model

In the present and following sections, we discuss some rigorous results concerning (metallic) ferromagnetism in one-dimensional models with $U \to \infty$ related to the Hubbard model. Thanks to the special nature of one-dimensional models, all the results can be obtained by straightforward applications of the Perron-Frobenius theorem. As far as we know, such an application of the Perron-Frobenius theorem to metallic ferromagnetism in one dimension was developed in the pioneering work of Kubo [72], where he studied the double exchange model. Although the results we shall discuss appear to be somewhat similar to the conjectures in the previous section, we still have no idea if they shed light on the truly fascinating problem of metallic ferromagnetism in the Hubbard model.

We first discuss the ferromagnetic t-J model. This is not an interesting model (in one dimension), but helps us in illustrating the basic structure of the proof that we use for other models. Unlike the Hubbard model, this model contains spin-spin interactions which explicitly favor ferromagnetism. The problem is whether ferromagnetism is realized in the presence of hopping. As we shall see (and is well-known), the problem is trivial in one dimension.

Consider the one-dimensional lattice $\Lambda = \{1, \dots, N_{\rm s}\}$. The Hamiltonian of the ferromagnetic t-J model is

$$H = -t \sum_{x=1}^{N_{\rm s}} \sum_{\sigma=\uparrow,\downarrow} (c^\dagger_{x,\sigma} c_{x+1,\sigma} + \text{h.c.}) - J \sum_{x=1}^{N_{\rm s}} (\hat{\mathbf{S}}_x \cdot \hat{\mathbf{S}}_{x+1} - \frac{n_x n_{x+1}}{4}) + U \sum_{x=1}^{N_{\rm s}} n_{x,\uparrow} n_{x,\downarrow}, \tag{91}$$

with $t > 0$, $J > 0$. Here h.c. represents the hermitian conjugate. We let $U \to \infty$ to inhibit double occupancies. We use periodic boundary conditions and identify the site $N_{\rm s}+1$ with 1. The spin operator at site x is defined as $\hat{\mathbf{S}}_x = \sum_{\sigma,\tau} c^\dagger_{x,\sigma} (p^{(\alpha)})_{\sigma,\tau} c_{x,\sigma}$. (See (41).)

Theorem 15.19 (Ferromagnetism in the one-dimensional ferromagnetic t-J model)**.** *Assume that the electron number $N_{\rm e}$ satisfies $N_{\rm e} \le N_{\rm s}$ and is odd*[60]*. Then the ground states of the present t-J model have total spin $S_{\rm tot} = S_{\max} (= N_{\rm e}/2)$, and are non-degenerate apart from the trivial $(2S_{\max}+1)$-fold degeneracy.*

Proof: The proof is based on an elementary observation which one may call a rigorous "spin-charge separation" argument. (See [50] and references therein.) To take into account the $U \to \infty$ limit (with the help of Lemma 15.10), we use as our basis the collection of the states (16) with $x_1 < x_2 < \cdots < x_{N_{\rm e}}$ and $\sigma_j = \uparrow, \downarrow$ such that $S^{(3)}_{\rm tot} = \sum_{j=1}^{N_{\rm e}} \sigma_j = 1/2$. The theory of angular momenta ensures that any eigenstate of H has its copy in the sector with the lowest $|S^{(3)}_{\rm tot}|$. This sector is spanned by this basis. We first claim that all the matrix elements of H in the present basis are nonpositive. As for the hopping term, this is trivial since there is no exchange in electron ordering and one does not have to worry about fermion signs[61]. As for the exchange term, this is easily verified by using the identity $\hat{\mathbf{S}}_x \cdot \hat{\mathbf{S}}_{x+1} = (\hat{S}^+_x \hat{S}^-_{x+1} + \hat{S}^-_x \hat{S}^+_{x+1})/2 + \hat{S}^{(3)}_x \hat{S}^{(3)}_{x+1}$. It is also verified that all the basis states are connected via nonvanishing matrix elements of H. Note that the existence of the exchange term is essential for this property. Therefore we can readily apply the Perron-Frobenius theorem (Theorem 15.11) to conclude that the ground

[60] We can allow any $N_{\rm e} \le N_{\rm s}$ if we use open boundary conditions. Peculiar dependence of the nature of the ground states on the parity of the electron number and boundary conditions is related to the appearance of spiral states. See [72, 73, 50, 65].

[61] Hops between sites 1 and $N_{\rm e}$ are exceptional, but these do not produce any signs if $N_{\rm e}$ is odd. When $N_{\rm e}$ is even, such hops generate a "wrong" sign for the Perron-Frobenius theorem. This "frustration" can be regarded as the origin of the spiral states [72, 73, 50, 65].

state $\Phi_{\rm GS}$ is unique (in this sector), and it is a linear combination of all the basis states with positive coefficients.

This fact indeed implies that for $\Phi_{\rm GS}$, $S_{\rm tot} = S_{\rm max}$. To see this, we first note that uniqueness of the ground state implies that $\Phi_{\rm GS}$ is an eigenstate of $(\hat{\mathbf{S}}_{\rm tot})^2$. We then take a reference state $\Phi = \left(\hat{S}^-_{\rm tot}\right)^{(N_{\rm e}-1)/2} \left(\prod_{x=1}^{N_{\rm e}} c^\dagger_{x,\uparrow}\right) \Phi_{\rm vac}$, which obviously has $S^{(3)}_{\rm tot} = 1/2$ and $S_{\rm tot} = S_{\rm max}$. Clearly Φ is written as a linear combination of basis states (16) with nonnegative coefficients. Thus we see $\langle \Phi, \Phi_{\rm GS}\rangle \neq 0$, which implies that $S_{\rm tot} = S_{\rm max}$ also holds for $\Phi_{\rm GS}$. □

Once knowing that the ground states have $S_{\rm tot} = S_{\rm max}$, we can look at the ground state with $S^{(3)}_{\rm tot} = S_{\rm max}$. Then it is apparent that the term $(\hat{\mathbf{S}}_x \cdot \hat{\mathbf{S}}_{x+1} - n_x n_{x+1}/4)$ always vanishes, and hence the ground state is identical to that of $H_{\rm hop}$. Then we have the exact expression for the ground state

$$\Phi_{\rm GS} = \left(\prod_{k \text{ s.t. } |k| \le k_{\rm F}} C^\dagger_\uparrow(\boldsymbol{\eta}^{(k)}) \right) \Phi_{\rm vac}, \tag{92}$$

where $\boldsymbol{\eta}^{(k)}$ is the plane wave state (11), and $k_{\rm F} = \pi(N_{\rm e}-1)/(2L)$ is the fermi momentum. The ground state allows gapless (charge) excitations, and hence describes a metallic system.

Unfortunately, the above theorem and proof do not shed *any* light on the corresponding problems in higher dimensions, which are much more important and interesting. The proof for the one-dimensional case relies heavily on the fact that essentially no nontrivial spin exchanges take place when $U = \infty$ and $J = 0$. The system suddenly becomes ferromagnetic when we turn on arbitrarily small $J > 0$.

In the corresponding problem in higher dimensional lattices (or even on slightly more complicated one-dimensional lattices, like a ladder), there are highly nontrivial exchange processes even for $J = 0$. There is no guarantee that one gets ferromagnetism when introducing an explicit ferromagnetic interaction, $J > 0$. We of course believe that the ferromagnetic t-J models in higher dimensions also exhibit ferromagnetism for sufficiently large J/t and $\rho = N_{\rm e}/N_{\rm s}$, but we have no idea how one can prove such a statement. Although "proving ferromagnetism in the ferromagnetic t-J models" might sound like a tautology at first glance, it is indeed a very difficult and deep problem, whose solution should shed light on various aspects of strongly interacting itinerant electron systems.

15.8.3 Hubbard model with correlated hopping

In this section, we discuss an artificial model which (like the Hubbard model, and unlike the t-J model) have no interactions explicitly favoring ferromagnetism, but in which we can still prove the appearance of (probably metallic) ferromagnetism in a wide range of parameters and electron density.

Take the one-dimensional lattice $\Lambda = \{1, \ldots, N_{\rm s}\}$ with even $N_{\rm s}$, and impose periodic boundary conditions by identifying $N_{\rm s} + x$ with x. We consider the Hamiltonian

$$H = -t \sum_{x=1}^{N_{\rm s}} \sum_{\sigma=\uparrow,\downarrow} (c^\dagger_{x,\sigma} c_{x+1,\sigma} + \text{h.c.}) + t' \sum_{j=1}^{N_{\rm s}/2} \sum_{\sigma=\uparrow,\downarrow} (c^\dagger_{2j,\sigma} c_{2j+2,\sigma} n_{2j+1} + \text{h.c.}) + U \sum_{x=1}^{N_{\rm s}} n_{x,\uparrow} n_{x,\downarrow}, \tag{93}$$

where we set $t > 0$, $t' > 0$. We will take the $U \to \infty$ limit.

This resembles the Hubbard model with next nearest neighbor hopping, but the hopping term with t' has an extra number operator $n_{2j+1} = n_{2j+1,\uparrow} + n_{2j+1,\downarrow}$. Thus a hop between $2j$ and $2j+2$ is possible only when $2j+1$ is occupied by an electron. Note, however, that such density-dependent hopping terms themselves are completely spin-independent, and do not prefer ferromagnetism. Nevertheless we have the following.

Theorem 15.20 (Ferromagnetism in the model with correlated hopping)**.** *Consider the model (93) in the limit $U \to \infty$ with any odd electron number N_e. Then the ground states have total spin $S_{tot} = S_{max}(= N_e/2)$, and are non-degenerate apart from the trivial $(2S_{max}+1)$-fold degeneracy.*

Proof: Exactly as in the proof of Theorem 15.19, we take account of the $U \to \infty$ limit by using the basis states (16) with $x_1 < x_2 < \cdots < x_{N_e}$ and $\sigma_j = \uparrow, \downarrow$ such that $S^{(3)}_{tot} = \sum_{j=1}^{N_e} \sigma_j = 1/2$. It is again trivial to check that the nearest-neighbor hopping terms in (93) have nonpositive matrix elements. By definition, an electron hops between even sites x and $x+2$ only when $x+1$ is occupied. Then, such a hop is always associated with a change of ordering in electrons. (See Figure 342.) This yields a minus sign, resulting again in nonpositive matrix elements.

It remains to prove that all the basis states (with $S^{(3)}_{tot} = 1/2$) are connected with each other by nonvanishing matrix elements. However, this is already proved in the one hole case in Section 15.4.3, and extensions to the multiple holes are trivial. Thus we can apply the Perron-Frobenius theorem exactly as in the proof of Theorem 15.19. □

We believe that the ferromagnetic ground states of the present model have a metallic character. But this is not as obvious as in the case of the t-J model. The model remains strongly interacting in the sector with $S^{(3)}_{tot} = S_{max}$, and it is not easy to investigate the properties of the ground states.

It is clear that the above theorem extends to much more general models with additional density-density interactions and on-site potential provided that there are nearest neighbor and next nearest neighbor (possibly site-dependent) hoppings with the correct signs. An interesting extension is obtained by replacing the second sum in (93) over even sites with a sum over all the sites as

$$H_1 = -t \sum_{x=1}^{N_s} \sum_{\sigma=\uparrow,\downarrow} (c^\dagger_{x,\sigma} c_{x+1,\sigma} + \text{h.c.}) + t' \sum_{x=1}^{N_s} \sum_{\sigma=\uparrow,\downarrow} (c^\dagger_{x,\sigma} c_{x+2,\sigma} n_{x+1} + \text{h.c.}) + U \sum_{x=1}^{N_s} n_{x,\uparrow} n_{x,\downarrow}. \quad (94)$$

This defines a translation invariant model which resembles the t-t' model. Of course the same proof as above shows that the model exhibits ferromagnetism for $U = \infty$.

To see the relation between the model (94) and the t-t' model, consider the Hamiltonian

$$H_2 = H_1 + t'' \sum_{x=1}^{N_s} \sum_{\sigma=\uparrow,\downarrow} \{c^\dagger_{x,\sigma} c_{x+2,\sigma} (1 - n_{x+1}) + \text{h.c.}\}. \quad (95)$$

When $t'' = t'$, this becomes the Hamiltonian of the t-t' model, while it defines a controllable model (94) when $t'' = 0$. One finds that the t''-term in (95) (along with other terms) generates an antiferromagnetic interaction. Therefore in order to control the possible ferromagnetism in the t-t' model, one has to deal with the competition between the t'-term and the t''-term in (95), which seems to be a nontrivial task[62].

15.8.4 Limiting U-V model

We consider another toy model (which perhaps is slightly less artificial than the one in the previous section) obtained by adding a strong nearest-neighbor repulsion to the Hubbard model. Again we can prove the existence of (metallic) ferromagnetism easily.

[62] One does not face this competition in first order perturbation theory [66, 67].

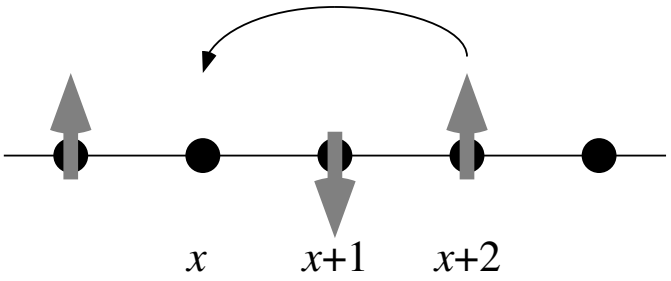


Figure 342: The essential electron exchange process in the correlated hopping model and the limiting U-V model. In the limiting U-V model with $N_{\rm e} > N_{\rm s}/2$, sites $x-1$ and $x+1$ must be occupied when x is empty. Then the ordering of electrons always changes when an electron hops from $x+2$ to x. In the model with correlated hopping, the same fact is guaranteed by the definition.

Take the one-dimensional lattice $\Lambda = \{1, \ldots, N_{\rm s}\}$ with even $N_{\rm s}$, and impose periodic boundary conditions. We consider a model with electron number $N_{\rm e}$ in the range $N_{\rm s} > N_{\rm e} > N_{\rm s}/2$, and with the Hamiltonian

$$\begin{aligned} H &= -t\sum_{x=1}^{N_{\rm s}}\sum_{\sigma=\uparrow,\downarrow}(c^\dagger_{x,\sigma}c_{x+1,\sigma} + {\rm h.c.}) + t'\sum_{j=1}^{N_{\rm s}/2}\sum_{\sigma=\uparrow,\downarrow}(c^\dagger_{2j,\sigma}c_{2j+2,\sigma} + {\rm h.c.}) \\ &\quad + U\sum_{x=1}^{N_{\rm s}} n_{x,\uparrow}n_{x,\downarrow} + V\left\{\left(\sum_{x=1}^{N_{\rm s}} n_x n_{x+1}\right) - (2N_{\rm e} - N_{\rm s})\right\}, \end{aligned} \tag{96}$$

where we require $t > 0$ and $t' > 0$. We will consider the $U, V \to \infty$ limits. The new nearest neighbor repulsion term is normalized so that the minimum energy for this term is zero. Note that the hopping amplitudes in the above model can be gauge transformed[63] to satisfy the condition for Nagaoka's theorem.

As in the Hubbard model, none of the terms in (96) explicitly favor ferromagnetism. Nevertheless we can prove that they together generate ferromagnetism.

Theorem 15.21 (Ferromagnetism in the limiting U-V model)**.** *Consider the model (96) in the limits $U \to \infty$ and $V \to \infty$ with odd $N_{\rm e}$ such that $N_{\rm s} > N_{\rm e} > N_{\rm s}/2$. Then the ground states have total spin $S_{\rm tot} = S_{\max}(= N_{\rm e}/2)$, and are non-degenerate apart from the trivial $(2S_{\max}+1)$-fold degeneracy.*

Proof: We take into account the $V \to \infty$ limit exactly as in Lemma 15.10. Now our basis states are (16) with $x_1 < x_2 < \cdots < x_{N_{\rm e}}$ and $\sigma_j = \uparrow, \downarrow$, where $x_1, \ldots, x_{N_{\rm e}}$ must further satisfy the condition that the nearest neighbor repulsion term in (96) is minimized. When $N_{\rm e} > N_{\rm s}/2$, this condition implies that there can be no neighboring holes in the relevant basis states. This guarantees that an electron hops between even sites x and $x+2$ only when $x+1$ is occupied. (See Figure 342.) The remainder of the proof is the same as that of Theorem 15.20. □

Again, this result can be extended to various more complicated models.

Note that the nature of the model is quite different for electron numbers $N_{\rm e} \le N_{\rm s}/2$ (in which case we drop $-(2N_{\rm e} - N_{\rm s})$ from the Hamiltonian (96)). In the limits $U \to \infty$ and $V \to \infty$, configurations with neighboring electrons are inhibited, and hence exchange processes never take place. In this case the model should exhibit paramagnetism. This is consistent with a numerical work [74] where a shift in the ferromagnetic region as a result of finite V is reported for $N_{\rm e} \le N_{\rm s}/2$. We expect that a finite value of V enlarges the region of ferromagnetism for $N_{\rm s}/2 < N_{\rm e} < N_{\rm s}$.

[63] Take Λ' in Appendix 15.9.1 as the set of even sites.

15.9 Adds

15.9.1 Gauge transformation

The signs of the hopping amplitude $t_{x,y}$ can be partially changed by means of gauge transformations. Let Λ' be an arbitrary subset of Λ. We define the new operators $\tilde{c}_{x,\sigma}$ by $\tilde{c}_{x,\sigma} = -c_{x,\sigma}$ if $x \in \Lambda'$ and $\tilde{c}_{x,\sigma} = c_{x,\sigma}$ if $x \notin \Lambda'$. Since $\tilde{c}_{x,\sigma}$ also satisfy the canonical anticommutation relations (12) and (13), we can use these operators to describe the system. Then the hopping Hamiltonian (32) is written as

$$H_{\rm hop} = \sum_{\substack{x,y\in\Lambda \\ \sigma=\uparrow,\downarrow}} t'_{x,y}\tilde{c}^\dagger_{x,\sigma}\tilde{c}_{y,\sigma}, \tag{97}$$

with $t'_{x,y} = t_{x,y}$ if both x and y are in Λ' or if neither x nor y is in Λ', and $t'_{x,y} = -t_{x,y}$ if exactly one of x and y is in Λ'.

The model is said to be bipartite when there is a subset Λ' with the property that $t_{x,y} = 0$ if both x and y are in Λ' or if neither x nor y is in Λ'. In a bipartite system, we can use the above gauge transformation to change the signs of all the hopping amplitudes. A typical example is the model on the simple cubic lattice with only nearest neighbor hoppings.

15.9.2 Hole-particle transformation

Let us discuss a simple and standard transformation which maps a Hubbard model onto a different Hubbard model with (usually) different electron number.

We define the new operators $\tilde{c}_{x,\sigma}$ by $\tilde{c}_{x,\sigma} = c^\dagger_{x,\sigma}$. This simply means that we switch the creation and the annihilation operators. Since the $\tilde{c}_{x,\sigma}$ also satisfy the canonical anticommutation relations (12) and (13), we can use these operators to describe the system. From (12), the new number operator $\tilde{n}_{x,\sigma} = \tilde{c}^\dagger_{x,\sigma}\tilde{c}_{x,\sigma}$ is related to the original number operator by $\tilde{n}_{x,\sigma} = 1 - n_{x,\sigma}$. Therefore the new total electron number $\tilde{N}_{\rm e}$ (which is an eigenvalue of $\sum_{x\in\Lambda,\sigma=\uparrow,\downarrow} \tilde{n}_{x,\sigma}$) is related to the original electron number by $\tilde{N}_{\rm e} = 2N_{\rm s} - N_{\rm e}$.

The hopping Hamiltonian (32) is written as

$$H_{\rm hop} = \sum_{\substack{x,y\in\Lambda \\ \sigma=\uparrow,\downarrow}} (-t_{x,y})\tilde{c}^\dagger_{x,\sigma}\tilde{c}_{y,\sigma} + 2\sum_{x\in\Lambda} t_{x,x}, \tag{98}$$

and the interaction Hamiltonian (37) as

$$\begin{aligned} H_{\rm int} &= U\sum_{x\in\Lambda}(1-\tilde{n}_{x,\uparrow})(1-\tilde{n}_{x,\downarrow}) \\ &= U\sum_{x\in\Lambda}\tilde{n}_{x,\uparrow}\tilde{n}_{x,\downarrow} - U\sum_{\substack{x\in\Lambda \\ \sigma=\uparrow,\downarrow}}\tilde{n}_{x,\sigma} + N_{\rm s}U. \end{aligned} \tag{99}$$

Note that (99) has exactly the same form as the original interaction Hamiltonian (37) apart from the constant term and the term proportional to the total electron number, which simply shift the total energy.

15.9.3 Positive semidefinite operators

We summarize the definition and elementary properties of positive semidefinite operators (or matrices) which we used in the main body of the paper. They should be well-known to readers with a mathematical background.

Let $\mathcal{H}$ be a finite dimensional Hilbert space with inner product $\langle \cdot, \cdot \rangle$. In the paper, $\mathcal{H}$ may be $\mathcal{H}_{N_e}$ or $\mathfrak{h}$.

Definition 15.22 (Positive semidefiniteness)**.** *For an operator (or matrix) on $\mathcal{H}$, we write $A \geq 0$ and say A is positive semidefinite if A is self-adjoint (or hermitian) and we have $\langle \Phi, A\Phi \rangle \geq 0$ for any $\Phi \in \mathcal{H}$. For two self-adjoint operators A and B, we write $A \geq B$ if $A - B \geq 0$.*

The following statement is easily proved by diagonalizing A.

Lemma 15.23 (Positive semidefiniteness and eigenvalues)**.** *A self-adjoint operator (or a hermitian matrix) A is positive semidefinite if and only if all the eigenvalues of A are nonnegative.*

The following lemma provides a standard way of constructing a positive semidefinite operator.

Lemma 15.24. *Let B be an arbitrary operator (or matrix) on $\mathcal{H}$. Then $A = B^\dagger B$ is positive semidefinite.*

Proof: Observe that for any Φ, we have $\left\langle \Phi, B^\dagger B\Phi \right\rangle = \langle B\Phi, B\Phi \rangle \geq 0$. □

Conversely, any $A \geq 0$ can be expressed as $A = B^2$ with some $B \geq 0$. This is most easily verified by diagonalizing A.

The following lemma is used repeatedly in the main body of the paper.

Lemma 15.25 (The sum of positive semidefinite operators)**.** *If $A \geq 0$ and $B \geq 0$, we have $A + B \geq 0$.*

Proof: Assume that $A + B$ has an eigenstate Φ with a negative eigenvalue. Then we get $0 > \langle \Phi, (A + B)\Phi \rangle = \langle \Phi, A\Phi \rangle + \langle \Phi, B\Phi \rangle \geq 0$, which is a contradiction. □

The following is trivial if we expand Φ into eigenstates of A.

Lemma 15.26. *Let $A \geq 0$. Then $\langle \Phi, A\Phi \rangle = 0$ is equivalent to $A\Phi = 0$.*

The following lemma is also useful.

Lemma 15.27. *Let $A_i \geq 0$ for $i = 1, \ldots, n$. Then $\sum_{i=1}^{n} A_i\Phi = 0$ implies $A_i\Phi = 0$ for each $i = 1, \ldots, n$.*

Proof: Since $\sum_{i=1}^{n} A_i\Phi = 0$, we have $0 = \left\langle \Phi, \sum_{i=1}^{n} A_i\Phi \right\rangle = \sum_{i=1}^{n} \langle \Phi, A_i\Phi \rangle$. By noting that $\langle \Phi, A_i\Phi \rangle \geq 0$, this means $\langle \Phi, A_i\Phi \rangle = 0$ for each $i = 1, \ldots, n$. We then use Lemma 15.26. □

We finally state the following lemma, which sometimes provides us with a powerful information.

Lemma 15.28. *Assume that $A \geq 0$ is expressed as $A = B^\dagger B$ where B is not necessarily self-adjoint. Then $A\Phi = 0$ implies $B\Phi = 0$.*

Proof: Since $B^\dagger B\Phi = 0$, we have $0 = \left\langle \Phi, B^\dagger B\Phi \right\rangle = \langle B\Phi, B\Phi \rangle$ which means $B\Phi = 0$. □

15.9.4 Explicit construction of the Hilbert space and fermion operators

In section 15.2.2, we first introduced the algebra of fermion operators, and then defined the Hilbert space by operating (a representation of) the algebra onto a single state $\Phi_{\rm vac}$. There is no problem in making the discussion mathematically rigorous since both the algebra and the Hilbert space are finite dimensional. However, readers familiar with fields like functional analysis might feel more comfortable if the Hilbert space is first defined explicitly and then the operators are defined. Here we explicitly define the Hilbert space and the fermion operators following the standard approach[64].

Let $\tilde{\Lambda} = \Lambda \times \{\uparrow, \downarrow\}$ be the configuration space for a single electron. We denote its elements as[65] $u, u_1, u_2, \cdots \in \tilde{\Lambda}$. The Hilbert space for a single electron is $\ell^2(\tilde{\Lambda}; \mathbb{C})$. For $n = 0, 1, \ldots, 2N_{\rm s}$, we define the n-electron Hilbert space $\mathcal{H}_n$ by $\mathcal{H}_0 = \mathbb{C}$, and

$$\mathcal{H}_n = P_- \underbrace{\ell^2(\tilde{\Lambda}; \mathbb{C}) \otimes \cdots \otimes \ell^2(\tilde{\Lambda}; \mathbb{C})}_{n}, \tag{100}$$

where P_- is the projection operator onto functions which are antisymmetric under exchanges of any two variables. More precisely, it is defined as

$$P_- \psi(u_1, \ldots, u_n) = \frac{1}{n!} \sum_{p:(1,\ldots,n)\to(p(1),\ldots,p(n))} (-1)^p \psi(u_{p(1)}, \ldots, u_{p(n)}), \tag{101}$$

where p is summed over all permutations of $(1, \ldots, n)$, and $(-1)^p$ denotes the parity of p.

We now define the fermion operators. For $\psi \in \mathcal{H}_0$, we let $c_u \psi = 0$. For $\psi \in \mathcal{H}_n$ with $n \geq 1$, we let

$$(c_u \psi)(u_1, \ldots, u_{n-1}) = \sqrt{n}\, \psi(u, u_1, \ldots, u_{n-1}), \tag{102}$$

where $c_u \psi \in \mathcal{H}_{n-1}$. On $\psi \in \mathcal{H}_n$ with $n < 2N_{\rm s}$, the adjoint operator $c_u^\dagger$ acts as

$$(c_u^\dagger \psi)(u_1, \ldots, u_{n+1}) = \frac{1}{\sqrt{n+1}} \sum_{j=1}^{n+1} (-1)^{j+1} \delta_{u,u_j} \psi(u_1, \ldots, u_{j-1}, u_{j+1}, \ldots, u_{n+1}), \tag{103}$$

where $c_u^\dagger \psi \in \mathcal{H}_{n+1}$. For $\psi \in \mathcal{H}_{2N_{\rm s}}$, we let $c_u^\dagger \psi = 0$.

It is natural to regard the operators $c_u : \mathcal{H}_n \to \mathcal{H}_{n-1}$ and $c_u^\dagger : \mathcal{H}_n \to \mathcal{H}_{n+1}$ as acting on the Fock space

$$\mathcal{F} = \bigoplus_{n=0}^{2N_{\rm s}} \mathcal{H}_n. \tag{104}$$

Finally, we take a basis state (say 1) of the Hilbert space $\mathcal{H}_0 = \mathbb{C}$, and identify it with $\Phi_{\rm vac}$. Then it is not difficult to check that states of the form (16) form a basis of the Hilbert space $\mathcal{H}_{N_{\rm e}}$.

As examples, let us calculate $f = (c_v^\dagger \Phi_{\rm vac}) \in \mathcal{H}_1$ and $g = (c_w^\dagger f) = (c_w^\dagger c_v^\dagger \Phi_{\rm vac}) \in \mathcal{H}_2$, where we understand that $\Phi_{\rm vac} = 1$. From (103), we have $f(u_1) = (c_v^\dagger 1)(u_1) = \delta_{v,u_1}$. Similarly we get

$$\begin{aligned} g(u_1, u_2) &= (c_w^\dagger f)(u_1, u_2) \\ &= \frac{1}{\sqrt{2}} \{\delta_{w,u_1} f(u_2) - \delta_{w,u_2} f(u_1)\} \\ &= \frac{1}{\sqrt{2}} (\delta_{w,u_1}\delta_{v,u_2} - \delta_{w,u_2}\delta_{v,u_1}) . \end{aligned} \tag{105}$$

[64] For more details, see, for example, Section 5.2 of [75].

[65] Of course $u = (x, \sigma)$ with $x \in \Lambda$ and $\sigma = \uparrow, \downarrow$.

Note that g is antisymmetric under the exchange of the names of the two particles. This is a special case of more general states

$$(c^\dagger_{v_1}\cdots c^\dagger_{v_n}\Phi_{\rm vac})(u_1,\ldots,u_n)=\frac{1}{\sqrt{n!}}\begin{vmatrix}\delta_{v_1,u_1} & \delta_{v_2,u_1} & \cdots & \delta_{v_n,u_1}\\ \delta_{v_1,u_2} & \delta_{v_2,u_2} & \cdots & \delta_{v_n,u_2}\\ \vdots & \vdots & \ddots & \vdots\\ \delta_{v_1,u_n} & \delta_{v_2,u_n} & \cdots & \delta_{v_n,u_n}\end{vmatrix}, \tag{106}$$

which are known as the Slater determinant states.

15.9.5 Band structures in single-electron energy spectra

Here we explain (for readers without background in condensed matter physics) the notion of band structure in a single-electron problem. The readers will find the notion quite elementary[66], especially in a tight-binding model.

We realize the lattice Λ as a subset of the d-dimensional $L\times\cdots\times L$ torus

$$T_L=\left\{(x_1,\ldots,x_d)\,\middle|\,0\le x_j<L\right\}=\mathbb{R}^d/\!\sim, \tag{107}$$

where the identification $\sim$ is defined by $x\sim y\iff x-y\in L\int e^d$. Here L is a positive integer. Let the unit cell $\mathcal{U}\subset[0,1)^d$ be a set of finite points (sites). The number of sites $|\mathcal{U}|$ in the unit cell will determine the number of bands. Assume that the lattice Λ can be written as

$$\Lambda=\left\{x+z\,\middle|\,x\in\mathcal{U},z\in\{0,1,\ldots,L-1\}^d\right\}. \tag{108}$$

We then assume that the hopping matrix $\mathsf{T}=(t_{x,y})_{x,y\in\Lambda}$ has translational invariance[67]

$$t_{x,y}=t_{x+z,y+z}, \tag{109}$$

for any $x,y\in\Lambda$ and $z\in\int e^d$. We note that perfect translational invariance is never possible in real physical systems which always have boundaries[68]. We often take artificial periodic boundary conditions because we want to concentrate on universal behavior taking place in the bulk of the system.

We are interested in the eigenvalue problem $\varepsilon\boldsymbol{\varphi}=\mathsf{T}\boldsymbol{\varphi}$ or

$$\varepsilon\varphi_x=\sum_{y\in\Lambda}t_{x,y}\,\varphi_y, \tag{110}$$

for a single electron. The translational invariance (109) suggests that we look for eigenstates in the form

$$\varphi_x=e^{ik\cdot x}v_{u(x)}, \tag{111}$$

where $k\in\mathcal{K}$ is a wave vector, $u(x)$ is the unique element in $\mathcal{U}$ such that $x-u(x)\in\int e^d$, and $\mathbf{v}=(v_u)_{u\in\mathcal{U}}$ is a $|\mathcal{U}|$-dimensional vector. The set of wave vectors $\mathcal{K}$ consists of $k=((2\pi/L)n_1,\ldots,(2\pi/L)n_d)$ with $n_j=0,\pm1,\ldots,\pm\{(L/2)-1\},L/2$ (where we assume L to be even). By substituting (111) into (110), we get

$$\varepsilon\,v_u=\sum_{u'\in\mathcal{U}}\tau^{(k)}_{u,u'}\,v_{u'}, \tag{112}$$

[66] In the band theory of electrons in solids, one usually talks about band structures in an "effective single-electron problem" (defined through complicated self-consistency arguments) in interacting many-electron problems. To explain this theory is beyond the scope of the present article (and the ability of the author).

[67] We of course identify $x+z$ and $y+z$ as elements in T_L by using the identification $\sim$ if necessary.

[68] The presence of boundaries does not change the band structure drastically, but introduces extra eigenstates which mainly live on the boundaries. Analysis of single-particle eigenstates may not be easy.

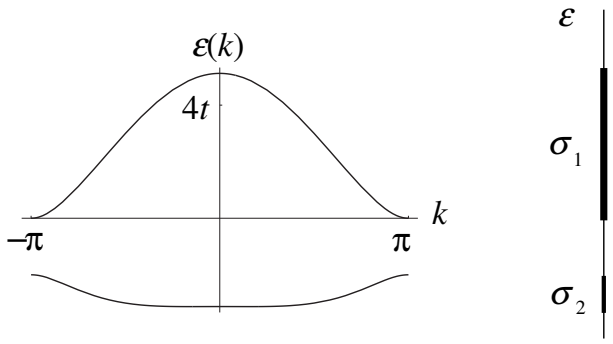


Figure 343: The dispersion relations and the corresponding band structure of the simple one-dimensional model with next-nearest-neighbor hoppings with $t = t'/2 > 0$ and $V = 0$. The spectrum consists of the two bands separated by a finite band gap.

with

$$\tau_{u,u'}^{(k)} = \sum_{y\in\Lambda \text{ s.t. } u(y)=u'} t_{u,y}\, e^{ik\cdot(y-u)}. \tag{113}$$

For each $k \in \mathcal{K}$, (112) has $|\mathcal{U}|$ eigenvalues that we denote as $\varepsilon_j(k)$ with $j = 1, \dots, |\mathcal{U}|$. We can choose $\varepsilon_j(k)$ so that $\varepsilon_j(k)$ is an analytic function of $k \in (-\pi, \pi]^d$ for each j. Consequently, the spectrum of the hopping matrix T is decomposed as

$$\sigma(\mathsf{T}) = \bigcup_{j=1}^{|\mathcal{U}|} \sigma_j(\mathsf{T}), \tag{114}$$

with $\sigma_j(\mathsf{T}) = \left\{ \varepsilon_j(k) \,\middle|\, k \in \mathcal{K} \right\}$. Each $\sigma_j(\mathsf{T})$ is called an *energy band*. The function $\varepsilon_j(k)$ is called the *dispersion relation* of the j-th band.

Let us discuss a simple example with two bands. Let $d = 1$, and take $\mathcal{U} = \{0, 1/2\}$. The resulting lattice is $\Lambda = \{0, 1/2, 1, 3/2, \dots, L - (1/2)\}$. We define the hopping matrix $\mathsf{T} = (t_{x,y})_{x,y\in\Lambda}$ by $t_{x,y} = t'$ if $|x - y| = 1/2$, $t_{x,y} = t$ if $x, y \in \int e$ and $|x - y| = 1$, $t_{x,x} = V$ if $x + 1/2 \in \int e$, and $t_{x,y} = 0$ otherwise. Then the eigenvalue equation (112) becomes

$$\varepsilon \begin{pmatrix} v_0 \\ v_{1/2} \end{pmatrix} = \begin{pmatrix} 2t\cos k & 2t'\cos(k/2) \\ 2t'\cos(k/2) & V \end{pmatrix} \begin{pmatrix} v_0 \\ v_{1/2} \end{pmatrix}. \tag{115}$$

The two eigenvalues of (115) define two dispersion relations

$$\varepsilon_{1,2}(k) = \frac{1}{2}\left\{ V + 2t\cos k \mp \sqrt{(V - 2t\cos k)^2 + (4t'\cos(k/2))^2} \right\}. \tag{116}$$

See Figure 343.

Consider special cases where the parameters t' and V can be written as $t' = \lambda t$ and $V = (\lambda^2 - 2)t$ with the parameter $\lambda > 0$. Then (116) becomes $\varepsilon_1(k) = -2t$ and $\varepsilon_2(k) = \lambda^2 t + 2t\cos k$. The model becomes the flat-band model[69] discussed in Section 15.6.3.

15.9.6 Proof of Theorem 15.6

We describe estimates required for the proof of Theorem 15.6. Although the calculations are straightforward, they illustrate some standard techniques we encounter in many-fermion problems and may serve as a good exercise for beginners.

[69] We recover the dispersion relation at the end of Section 15.6.3 if we shift the energy by $2t$.

First note that the sate Ψ defined in (46) satisfies $n_{x,\downarrow}n_{y,\downarrow}\Psi = 0$ for any $x \neq y$ since there is only one $\downarrow$ electron. Then the definition (49) of the variational state $\widetilde{\Psi}$ simplifies as

$$\widetilde{\Psi} = P_0\Psi = (1-\hat{\nu})\Psi, \tag{117}$$

with $\hat{\nu} = \sum_{x\in\Lambda} n_{x,\uparrow}n_{x,\downarrow}$. To evaluate the energy expectation value of (51), we observe that

$$\left\langle \widetilde{\Psi}, \widetilde{\Psi} \right\rangle = \langle P_0\Psi, P_0\Psi\rangle = \langle \Psi, P_0\Psi\rangle = 1 - \langle \Psi, \hat{\nu}\Psi\rangle\,, \tag{118}$$

and

$$\begin{aligned}\left\langle \widetilde{\Psi}, H_{\rm hop}\widetilde{\Psi} \right\rangle &= \langle \Psi, (1-\hat{\nu})H_{\rm hop}(1-\hat{\nu})\Psi\rangle \\ &= \langle \Psi, H_{\rm hop}\Psi\rangle - \langle \Psi, \hat{\nu}H_{\rm hop}\Psi\rangle - \langle \Psi, H_{\rm hop}\hat{\nu}\Psi\rangle + \langle \Psi, \hat{\nu}H_{\rm hop}\hat{\nu}\Psi\rangle \\ &= E_\Psi(1 - 2\langle \Psi, \hat{\nu}\Psi\rangle) + \langle \Psi, \hat{\nu}H_{\rm hop}\hat{\nu}\Psi\rangle\,, \end{aligned} \tag{119}$$

where we used $H_{\rm hop}\Psi = E_\Psi\Psi$ and $\langle \Psi, \Psi\rangle = 1$. Thus we need to estimate the quantities $\langle \Psi, \hat{\nu}\Psi\rangle$ and $\langle \Psi, \hat{\nu}H_{\rm hop}\hat{\nu}\Psi\rangle$.

The best way to evaluate these quantities is to go into a Fourier representation and express everything in terms of the $a_{j,\sigma}$ operators of (33). By using the completeness relation $\sum_{j=1}^{N_{\rm s}} (\psi_{x'}^{(j)})^* \psi_x^{(j)} = \delta_{x,x'}$, we find the inverse of (33) is

$$c_{x,\sigma} = \sum_{j=1}^{N_{\rm s}} \psi_x^{(j)} a_{j,\sigma}. \tag{120}$$

This leads us to the representations

$$\hat{\nu} = \sum_{p,q,r,s=1}^{N_{\rm s}} V_{p,q,r,s} a_{p,\uparrow}^\dagger a_{q,\uparrow} a_{r,\downarrow}^\dagger a_{s,\downarrow}, \tag{121}$$

with

$$V_{p,q,r,s} = \sum_{x\in\Lambda} (\psi_x^{(p)})^* \psi_x^{(q)} (\psi_x^{(r)})^* \psi_x^{(s)}, \tag{122}$$

and

$$H_{\rm hop} = \sum_{j=1}^{N_{\rm s}} \sum_{\sigma=\uparrow,\downarrow} \varepsilon_j\, a_{j,\sigma}^\dagger a_{j,\sigma}. \tag{123}$$

All we have to do now is substitute (121) and (123) into the desired quantities, and use the anticommutation relations $\{a_{j,\sigma}, a_{k,\tau}\} = 0$, $\{a_{j,\sigma}^\dagger, a_{k,\tau}\} = \delta_{j,k}\delta_{\sigma,\tau}$, the definition (46) of Ψ, and the fact that $a_{j,\sigma}\Phi_{\rm vac} = 0$. Then we obtain

$$\begin{aligned}\langle \Psi, \hat{\nu}\Psi\rangle &= \sum_{p=1}^{N_{\rm e}-1} V_{p,p,1,1} \\ &= \sum_{x\in\Lambda} \left(|\psi_x^{(1)}|^2 \sum_{p=1}^{N_{\rm e}-1} |\psi_x^{(p)}|^2 \right), \end{aligned} \tag{124}$$

$$\begin{aligned}\left\langle \Psi, \hat{\nu}H_{\rm hop}^{\uparrow}\hat{\nu}\Psi \right\rangle &= \left(\sum_{j=1}^{N_{\rm e}-1} \varepsilon_j \right) \langle \Psi, \hat{\nu}\Psi\rangle - \sum_{x\in\Lambda} \left(|\psi_x^{(1)}|^2 \sum_{p=1}^{N_{\rm e}-1} \varepsilon_p |\psi_x^{(p)}|^2 \right) \\ &\quad + \sum_{x\in\Lambda} \left(t_{x,x} |\psi_x^{(1)}|^2 \sum_{p=1}^{N_{\rm e}-1} |\psi_x^{(p)}|^2 \right), \end{aligned} \tag{125}$$

and

$$\begin{aligned}\left\langle \Psi, \hat{\nu} H_{\rm hop}^{\downarrow} \hat{\nu} \Psi \right\rangle &= \sum_{x,y\in\Lambda} \left\{ t_{x,y} \left(\sum_{p=1}^{N_{\rm e}-1} |\psi_x^{(p)}|^2 \right) \left(\sum_{q=1}^{N_{\rm e}-1} |\psi_y^{(q)}|^2 \right) (\psi_x^{(1)})^* \psi_y^{(1)} \right\} \\ &\quad + \sum_{x,y\in\Lambda} \left\{ t_{x,y} \left(\sum_{p=1}^{N_{\rm e}-1} (\psi_x^{(p)})^* \psi_y^{(p)} \right) \left(\sum_{q=N_{\rm e}}^{N_{\rm s}} \psi_x^{(q)} (\psi_y^{(q)})^* \right) (\psi_x^{(1)})^* \psi_y^{(1)} \right\} \\ &= \sum_{x\in\Lambda} \left\{ t_{x,x} \left(\sum_{p=1}^{N_{\rm e}-1} |\psi_x^{(p)}|^2 \right) |\psi_x^{(1)}|^2 \right\} + R, \end{aligned} \tag{126}$$

with

$$\begin{aligned} R &= \sum_{\substack{x,y\in\Lambda \\ \text{s.t. } x\neq y}} \left\{ t_{x,y} \left(\sum_{p=1}^{N_{\rm e}-1} |\psi_x^{(p)}|^2 \right) \left(\sum_{q=1}^{N_{\rm e}-1} |\psi_y^{(q)}|^2 \right) (\psi_x^{(1)})^* \psi_y^{(1)} \right\} \\ &\quad - \sum_{\substack{x,y\in\Lambda \\ \text{s.t. } x\neq y}} \left\{ t_{x,y} \left(\sum_{p=1}^{N_{\rm e}-1} (\psi_x^{(p)})^* \psi_y^{(p)} \right) \left(\sum_{q=1}^{N_{\rm e}-1} \psi_x^{(q)} (\psi_y^{(q)})^* \right) (\psi_x^{(1)})^* \psi_y^{(1)} \right\}, \end{aligned} \tag{127}$$

where we have used the fact that $\sum_{p=1}^{N_{\rm s}} \varepsilon_p (\psi_x^{(p)})^* \psi_y^{(p)} = t_{x,y}$ and $\sum_{p=1}^{N_{\rm s}} (\psi_x^{(p)})^* \psi_y^{(p)} = \delta_{x,y}$. We also made the obvious decomposition $H_{\rm hop} = H_{\rm hop}^{\uparrow} + H_{\rm hop}^{\downarrow}$.

By collecting all the estimates, we get

$$\begin{aligned} \Delta E &= \frac{\left\langle \widetilde{\Psi}, H_{\rm hop} \widetilde{\Psi} \right\rangle}{\left\langle \widetilde{\Psi}, \widetilde{\Psi} \right\rangle} - E_{\rm ferro} \\ &= \varepsilon_1 - \varepsilon_{N_{\rm e}} + \frac{2(\bar{t} - \varepsilon_1) \left\langle \Psi, \hat{\nu}\Psi \right\rangle}{1 - \left\langle \Psi, \hat{\nu}\Psi \right\rangle} + \frac{R}{1 - \left\langle \Psi, \hat{\nu}\Psi \right\rangle}, \end{aligned} \tag{128}$$

with

$$\bar{t} = \frac{\displaystyle\sum_{x\in\Lambda} \left(t_{x,x} |\psi_x^{(1)}|^2 \sum_{p=1}^{N_{\rm e}-1} |\psi_x^{(p)}|^2 \right)}{\displaystyle\sum_{x\in\Lambda} \left(|\psi_x^{(1)}|^2 \sum_{p=1}^{N_{\rm e}-1} |\psi_x^{(p)}|^2 \right)} \leq \max_{x\in\Lambda} t_{x,x}. \tag{129}$$

To proceed further, we assume that there is a constant $\alpha > 0$ independent of x, $N_{\rm s}$ and $p = 1, \ldots, N_{\rm s}$, and we have[70]

$$|\psi_x^{(p)}| \leq \frac{\alpha}{\sqrt{N_{\rm s}}}. \tag{130}$$

Note that if the model possesses translational invariance as in Appendix 15.9.5, then the bound is true with $\alpha = \sqrt{|\mathcal{U}|}$. We also assume that $\max_{x\in\Lambda} t_{x,x}$, $\max_{x\in\Lambda} \sum_{y\in\Lambda} |t_{x,y}|$, and ε_1 converge to finite quantities as $N_{\rm s} \to \infty$, which is again trivially valid in translation invariant models. With the requirement (130), we obtain $|R| \leq \text{const.}\rho^2$, which means that R gives only negligible contributions for small ρ. Since we also find $\langle \Psi, \hat{\nu}\Psi \rangle \propto \rho$, the bound (128) implies the desired estimate (51).

[70] This assumption rules out the possibility of localized eigenstates.

16 Solvable Model of a Spin-Glass by David Sherrington and Scott Kirkpatrick

We consider an Ising model in which the spins are coupled by infinite-ranged random interactions independently distributed with a Gaussian probability density. Both "spinglass" and ferromagnetic phases occur. The competition between the phases and the type of order present in each are studied.

Compelling experimental [1,2] and theoretical [3–5] evidence has accumulated in recent years suggesting that a new magnetic phase may occur in spatially random systems with competing exchange interactions. In this "spin-glass" phase, moments are frozen into equilibrium orientations, but there is no long-range order. Edwards and Anderson (EA) have demonstrated [3] that such a phase occurs within a novel form of molecularfield theory, and they propose that spin correlations between Gibbs-like replicas of the random system play the role of a spin-glass order parameter.

A closely related replica formalism has been employed in several recent papers [6,7] applying renormalization-group methods to random magnetic systems. The possibility of an EA-type [4] order parameter was not considered in that work, although some of the models studied [7] appear likely to exhibit spin-glass phases.

It is well known that molecular-field theory for a pure ferromagnet becomes exact in the thermodynamic limit for a constant infinite-ranged exchange interaction provided that the interaction is appropriately scaled with the number of spins in the system. [8] In this Letter we define and solve the analogous infinite-ranged problem for a disordered system. We obtain a spin-glass solution characterized by the EA order parameter in the appropriate regime of temperature and the strength of the exchange fluctuations. A simple interpretation of this order parameter is given. The various thermodynamic quantities and the competition with ferromagnetic long-range order are explored in some detail.

We consider N Ising spins interacting through infinite-ranged exchange interactions which are independently distributed with a Gaussian proba- bility density. The Hamiltonian is

$$\mathcal{H} = -\frac{1}{2}\sum_{i\neq i} J_{ij} S_i S_j, \quad S_i = \pm 1 \tag{1}$$

with the J_{ij} distributed according to

$$p\left(J_{ij}\right) = \left[(2\pi)^{1/2} J\right]^{-1} \exp\left[-\left(J_{ij} - J_0\right)^2 / 2J^2\right] \tag{2}$$

and J_0 and J scaled by

$$J_0 = \tilde{J}_0 / N, \quad J = \tilde{J} / N^{1/2} \tag{3}$$

so that $\tilde{J}_0$ and $\tilde{J}$ are both intensive. Following the usual procedure, we calculate the averaged free energy. (Averaging the free energy and not the partition function corresponds to treating a "quenched" rather than an "annealed" system.) With use of the identity

$$\ln x = \lim_{n\to 0} \left(x^n - 1\right) / n \tag{4}$$

the averaged free energy F may be expressed as

$$F = -kT \lim_{n\to 0} n^{-1} \left\{ \int \prod_{(ij)} \left[p\left(J_{ij}\right) dJ_{ij} \right] \mathrm{Tr}_n \exp\left(\sum_{\alpha=1,\ldots,n} \sum_{i\neq j} J_{ij} S_i{}^{\alpha} S_j^{\alpha} / 2kT \right) - 1 \right\}$$

$$= -kT \lim_{n\to 0} n^{-1} \left\{ \mathrm{Tr}_n \exp\left(\sum_{i\neq j} \left[\sum_{\alpha} S_i^{\alpha} S_j^{\alpha} J_0 / 2kT + \sum_{\alpha,\beta} S_i^{\alpha} S_j^{\alpha} S_i^{\beta} S_j^{\beta} J^2 / 4(kT)^2 \right] \right) - 1 \right\} \quad (5)$$

where α and β label n dummy replicas. Reordering and dropping terms which vanish in the thermodynamic limit yields

$$F = -kT \lim_{n\to 0} n^{-1} \left\{ \exp\left[J^2 N^2 n / 4(kT)^2 \right] \mathrm{Tr}_n \exp\left[\sum_{(\alpha\beta)} \left(\sum_i S_i{}^{\alpha} S_i^{\beta} \right)^2 J^2 / 2(kT)^2 + \sum_{\alpha} \left(\sum_i S_i{}^{\alpha} \right)^2 J_0 / 2kT \right. \right. \quad (6)$$

where $(\alpha\beta)$ refers to combinations of α and β with $\alpha \neq \beta$. Using the identity

$$\exp\left(\lambda a^2\right) = (2\pi)^{-1/2} \int dx \exp\left[-x^2/2 + (2\lambda)^{1/2} a x \right] \quad (7)$$

we rewrite (6) as

$$F = -kT \lim_{n\to 0} n^{-1} \left\{ \exp\left[\tilde{J}^2 N n / 4(kT)^2 \right] \int \left[\prod_{(\alpha\beta)} (N/2\pi)^{1/2} dy^{(\alpha\beta)} \right] \left[\prod_{\alpha} (N/2\pi)^{1/2} dx^{\alpha} \right] \right.$$

$$\times \exp\left[-N \sum_{(\alpha\beta)} \left(y^{(\alpha\beta)} \right)^2 / 2 - N \sum_{\alpha} \left(x^{\alpha} \right)^2 / 2 \right.$$

$$\left. \left. + N \ln \mathrm{Tr} \exp\left((\tilde{J}/kT) \sum_{(\alpha\beta)} y^{(\alpha\beta)} S^{\alpha} S^{\beta} + \left(\tilde{J}_0 / kT \right)^{1/2} \sum_{\alpha} x^{\alpha} S^{\alpha} \right) \right] - 1 \right\} \quad (8)$$

where the trace is now over n replicas at a single spin site.

It is assumed that the limit $n \to 0$ and the thermodynamic limit $N \to \infty$ can be interchanged. For integral $n \geqslant 2$, the integrals may be done by steepest descents. Since the replicas are indistinguishable, we consider only the extremum of the exponential for which all the $y^{(\alpha\beta)}$ are equal, as are all the x^{α}. We denote their values by y and x. This permits the replacement $\sum y^{(\alpha\beta)} S^{\alpha} S^{\beta} \to y \left[\left(\sum_{\alpha} S^{\alpha} \right)^2 - n \right]$, and $\left(\sum_{\alpha} S^{\circ} \right)^2$ may be absorbed by the introduction of a random field. [3] Continuation to arbitrary n, extraction of the terms linear in n as $n \to 0$, and the substitutions $y \to q(\tilde{J}/kT)$ and $x \to m \left(\tilde{J}_0 / kT \right)^{1/2}$ then yield

$$F = NkT \left\{ -\tilde{J}^2 (1-q)^2 / (2kT)^2 + \tilde{J}_0 m^2 / 2kT - (2\pi)^{-1/2} \int dz \exp\left(-z^2/2\right) \ln \left[2 \cosh\left(\tilde{J} q^{1/2} z / kT + \tilde{J}_0 m / k \right. \right. \right. \quad (9)$$

where q and m satisfy the simultaneous equations

$$q = 1 - (2\pi)^{-1/2} \int dz \exp\left(-z^2/2\right) \mathrm{sech}^2 \left[\left(\tilde{J} q^{1/2} / kT \right) z + \tilde{J}_0 m / kT \right] \quad (10a)$$

$$m = (2\pi)^{-1/2} \int dz \exp\left(-z^2/2\right) \tanh \left[\left(\tilde{J} q^{1/2} / kT \right) z + \tilde{J}_0 m / kT \right] \quad (10b)$$

To show the physical significance of m and q we note that the thermal average of the spin at site i, $\langle S_i \rangle$, and its square may be written

$$\langle S_i \rangle = (\partial/\partial h) \ln \mathrm{Tr} \exp \left(\sum_{i \neq j} J_{ij} S_i^\alpha S_j^\alpha / 2kT + h S_i{}^\alpha \right)_{h=0} \tag{11a}$$

$$\langle S_i \rangle^2 = (\partial/\partial h') \ln \mathrm{Tr} \exp \left[\sum_{i \neq j} J_{ij} \left(S_i^\alpha S_j^\alpha + S_i^\beta S_j^\beta \right) / 2kT + h' S_i^\alpha S_i^\beta \right]_{h'=0} \tag{11 b}$$

where $\alpha \neq \beta$ are dummy labels. Averaging over the J_{ij} distribution, which we denote by $\langle \rangle_J$, we see that $\langle \langle S_i \rangle \rangle_J$ and $\left\langle \langle S_i \rangle^2 \right\rangle_J$ are given by taking the $n \to 0$ limits respectively of $\langle S_i{}^\alpha \rangle$ and $\langle S_i{}^\alpha \times S_i{}^\beta \rangle_{\alpha \neq \beta}$ evaluated for a system characterized by the J-averaged n-ensemble partition function. This result is valid for finite-ranged interactions as well as infinite-ranged ones. Thus [9]

$$m \equiv \langle \langle S_i \rangle \rangle_J \tag{12a}$$

$$q \equiv \left\langle \langle S_i \rangle^2 \right\rangle_J \tag{12b}$$

independent of i. A nonzero q indicates magnetic order, while nonzero m (in addition to q) indicates that that order is ferromagnetic. When $m = 0$ but q is nonzero, we shall call the state a "spin-glass."

Equations (10) indicate that magnetic order sets in as kT is reduced below the greater of $\tilde{J}_0$ or $\tilde{J}$. If $\tilde{J}_0 > \tilde{J}$, the phase that is reached is ferromagnetic, but when the converse is true, spin-glass order ensues, and m remains zero for $kT < \tilde{J}$. The full phase diagram is plotted in Fig. 1, in terms of the dimensionless combinations $\tilde{J}_0/\tilde{J}$ and $kT/\tilde{J}$, and may easily be rescaled to describe models in which $\tilde{J}_0$ and $\tilde{J}$ are known functions of external parameters (see, e.g., Ref. 4).

For $\tilde{J}_0/\tilde{J} \gg 1$ the effects of fluctuations are weak, and one can show from (10a) and (10b) that $q \sim m^2$, in accord with the physical interpretation (12b) of q as the square of the modulus of the frozen mo-

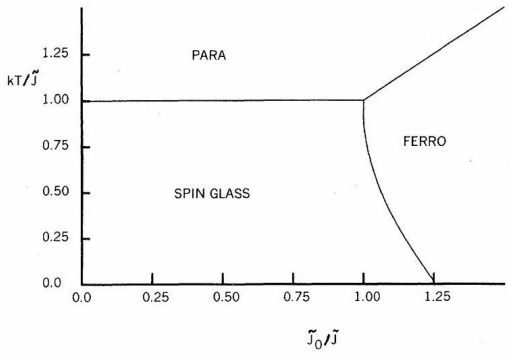


FIG. 1. Phase diagram of spin-glass ferromagnet. ment per site. The zero-temperature magnetization is diminished by weak fluctuations as

$$m \sim 1 - (2/\pi)^{1/2} \left(\tilde{J}/\tilde{J}_0 \right) \exp \left(-\tilde{J}_0^2 / 2\tilde{J}^2 \right) \tag{13a}$$

and vanishes continuously at the spin-glass phase boundary as

$$m \sim (18\pi)^{1/4} \left(\tilde{J}_0/\tilde{J} \right)^2 \left[(2/\pi)^{1/2} - \tilde{J}/\tilde{J}_0 \right]^{1/2} \tag{13b}$$

Values of $m(T)$ and $q^{1/2}$ obtained by numerical solution of (10a) and (10b) are plotted in Fig. 2 for various values of $\tilde{J}_0/\tilde{J}$. We note that the effect of fluctuations is strongest at low temperatures, causing the decrease in the magnetization as $T \to 0$ in Fig. 2, and a line of second-order transitions from ferromagnet to spin-glass in the phase diagram of Fig. 1.

The frozen moment, $q^{1/2}(T)$, as is shown in Fig. 2 and as can be derived from (10a) and (10b), is proportional to $(T_c - T)^{1/2}$ just below T_c, tends to unity as $T \to 0$, and is always greater than $m(T)$ at the same temperature. The linear low-temperature dependence of q and $q^{1/2}$,

$$1 - q(T) \sim (2/\pi)^{1/2}(kT/\tilde{J}) \exp\left(-\tilde{J}_0^2 m^2/2\tilde{J}^2\right) \tag{14}$$

as $T \to 0$, contrasts with that of $m(T)$ in a uniform Ising magnet, for which all temperature derivatives vanish at $T = 0$, since excitations from the ferromagnetic ground state require a

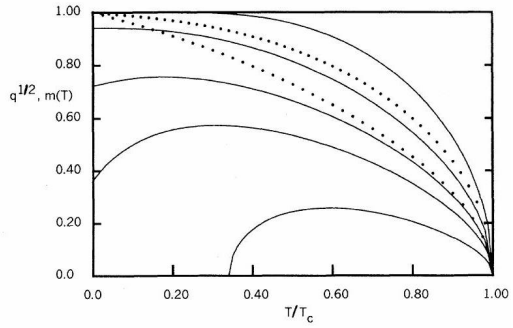


FIG. 2. Solid lines denote $m(T)$ for ratios $\tilde{J}_0/\tilde{J}$ of (top to bottom) $^{\infty}, 2.0, 1.51.3$, and 1.1. Dotted lines show $q^{1/2}(T)$ for $\tilde{J}_0/\tilde{J} = 2.0$ (upper line) and 0.0 (lower line).
finite energy.

The differential susceptibility, χ, may be obtained by repeating the steps leading to (9) and (10) with an external field term, $\sum_i HS_i$, in the Hamiltonian (1). This simply adds an extra contribution, H/kT, to the arguments of sech and tanh in (10a) and (10b). Differentiating (10b) with respect to H, and taking the limit $h \to 0$, we then obtain

$$\begin{aligned} \chi(T) &= [1 - q(T)] / \left\{kT - \tilde{J}_0[1 - q(T)]\right\} \\ &= \chi^{(0)} / \left(1 - \tilde{J}_0 \chi^{(0)}\right) \end{aligned} \tag{15}$$

where $\chi^{(0)}$ is the result for $\tilde{J}_0 = 0$. Above the ordering temperature, where $q = 0$, this is just a Curie-Weiss law. In the spin-glass phase, the fluctuations decrease χ, while $\tilde{J}_0$ enhances it. Two examples are plotted in Fig. 3. The dotted lines in Fig. 3 show the effect of a finite field, $H = 0.1\tilde{J}$, on the differential susceptibility in each case.

From (9), we obtain the internal energy, U:

$$U = -N\left[m^2\tilde{J}/2 + \tilde{J}^2\left(1 - q^2\right)/2kT\right] \tag{16}$$

In the spin-glass phase the leading term in the specific heat at low temperatures is

$$C \sim Nk(kT/\tilde{J})(2/\pi)^{1/2}\left[\left(\pi^2/12\right) - 1/2\pi\right] \tag{17}$$

At the spin-glass ordering temperature C has a cusp. For all $\tilde{J}_0, C$ equals $Nk\tilde{J}^2/2(kT)^2$ above the ordering temperature, [10] in contrast to the corresponding pure systems for which it vanishes. The linear temperature dependences seen in C, m, and $q^{1/2}$ suggest that the system

possesses excitations from the ground state whose density remains finite down to zero energy. The entropy S equals $Nk\left[\ln 2 - \tilde{J}^2/(2kT)^2\right]$ above the spin-glass ordering temperature, but goes to a negative limit, $-Nk/2\pi$, at $T = 0$. We speculate that this unphysical behavior has its origin in the interchange of limits $N \to \infty$ and $n \to 0$, but that the consequences are confined to low temperatures.

When an Ising system described by (1) and (2) with nearest-neighbor interactions is treated with mean-field theory (of the EA type), equations identical to (10) are obtained, with zJ_0 and $z^{1/2}J$ replacing J_0 and $\mathscr{J}$, where z is the average number of neighbors. [11] For interactions which are on the average antiferromagnetic or which include second-neighbor terms, analogous equations result in which m is replaced by the appropriate sublattice magnetization.

For the finite-ranged interactions occurring in real systems, the results presented here have at

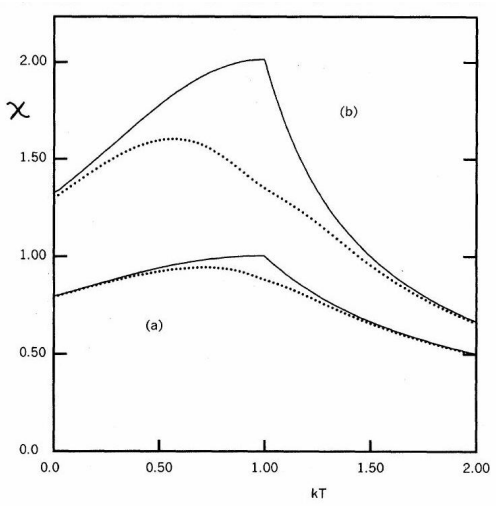


FIG. 3. Differential susceptibility without external field (solid lines) and with a field $\boldsymbol{H} = 0.1\tilde{J}$ (dotted lines) for $\tilde{J}_0/\tilde{J} = 0$, curves a, and $\tilde{J}_0/\tilde{J} = 0.5$, curves b.

best mean-field significance, and thus cannot treat critical phenomena correctly. However, it is evident that any more sophisticated treatment of critical properties in the presence of random competing exchange interactions should allow for the possibility of order parameters $\langle\langle S_i\rangle^n\rangle_J$ with $n > 1$, such as q discussed above.

17 Magnetic Monopoles in Spin Ice

Abstract

[70]*On sabbatical leave from Department of Physics, Imperial College, London SW7 2BZ, England.

[1] J. A. Mydosh, in Magnetism and Magnetic Materials-1974, edited by C. D. Graham, Jr., G. H. Lander, and J. J. Rhyne, AIP Conference Proceedings No. 24 (American Institute of Physics, New York, 1975), p. 131.

[2] D. E. Murnick, A. T. Fiory, and W. J. Kossler, to be published; J. M. D. Coey and R. W. Readman, Nature (London) 246,476 (1973).

[3] S. F. Edwards and P. W. Anderson, J. Phys. F: Met. Phys. 5, 965 (1975).

[4] D. Sherrington and B. W. Southern, J. Phys. F: Met. Phys. 5, L49 (1975).

[5] K. H. Fischer, Phys. Rev. Lett. 34,1438 (1975).

[6] V. J. Emery, Phys. Rev. B 11, 239 (1975) ; G. Grinstein, Ph.D. thesis, Harvard University, 1974 (unpublished); G. Grinstein and A. H. Luther, to be published. [7] A. Aharony, Phys. Rev. Lett. 34, 590 (1975).

[8] See, e.g., H. E. Stanley, Phase Transitions and Critical Phenomena (Oxford Univ. Press, New York, 1971) p. 91.

[9] The identification (12b) of q has also been pointed out by Fischer in Ref. 5.

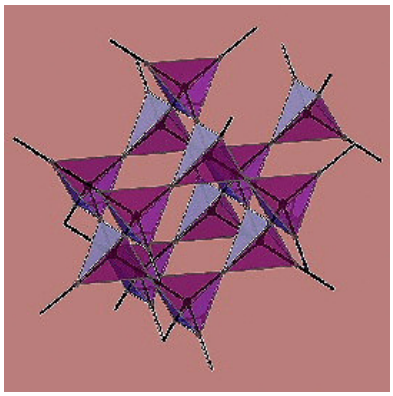


Figure 344: The pyrochlore and diamond lattices. The magnetic moments in spin ice reside on the sites of the pyrochlore lattice, which consists of corner-sharing tetrahedra. These are at the same time the midpoints of the bonds of the diamond lattice (black) formed by the centres of the tetrahedra. The ratio of the lattice constant of the diamond and pyrochlore lattices is $a_d/a = \sqrt{3/2}$. The Ising axes are the local [111] directions, which point along the respective diamond lattice bonds.

> Electrically charged particles, such as the electron, are ubiquitous. By contrast, no elementary particles with a net magnetic charge have ever been observed, despite intensive and prolonged searches[1]. We pursue an alternative strategy, namely that of realising them not as elementary but rather as *emergent* particles, i.e., as manifestations of the correlations present in a strongly interacting many-body system. The most prominent examples of emergent quasiparticles are the ones with fractional electric charge $e/3$ in quantum Hall physics[2]. Here we show that magnetic monopoles *do* emerge in a class of exotic magnets known collectively as spin ice[3, 5, 4]: the dipole moment of the underlying electronic degrees of freedom fractionalises into monopoles. This enables us to account for a mysterious phase transition observed experimentally in spin ice in a magnetic field[6, 7], which is a liquid-gas transition of the magnetic monopoles. These monopoles can also be detected by other means, e.g., in an experiment modelled after the celebrated Stanford magnetic monopole search[8].

Spin-ice materials are characterised by the presence of magnetic moments $\vec{\mu}_i$ residing on the sites of a pyrochlore lattice (depicted in Fig. 344). These moments are constrained to point along their respective local Ising axes $\hat{e}_i$ (the diamond lattice bonds in Fig. 344), and they can be modelled as Ising spins $\vec{\mu}_i = \mu S_i$, where $S_i = \pm 1$ and $\mu = |\vec{\mu}_i|$. For the spin ice compounds discussed here, $Dy_2Ti_2O_7$ and $Ho_2Ti_2O_7$, the magnitude μ of the magnetic moments equals approximately 10 Bohr magnetons ($\mu = 10\mu_B$). The thermodynamic properties of these compounds are known to be described with good accuracy by an energy term that accounts for the nearest neighbour exchange and the long ranged dipolar interactions[9, 10, 4]:

$$\begin{aligned} H &= \frac{J}{3}\sum_{\langle ij\rangle} S_i S_j \\ &+ Da^3 \sum_{(ij)} \left[\frac{\hat{e}_i \cdot \hat{e}_j}{|\mathbf{r}_{ij}|^3} - \frac{3\left(\hat{e}_i \cdot \mathbf{r}_{ij}\right)\left(\hat{e}_j \cdot \mathbf{r}_{ij}\right)}{|\mathbf{r}_{ij}|^5} \right] S_i S_j . \end{aligned} \tag{131}$$

The distance between spins is r_{ij}, and $a \simeq 3.54$Å is the pyrochlore nearest-neighbour distance. $D = \mu_0\mu^2/(4\pi a^3) = 1.41$K is the coupling constant of the dipolar interaction.

Spin ice was identified as a very unusual magnet when it was noted that it does not order to the lowest temperatures, T, even though it appeared to have *ferromagnetic* interactions[3].

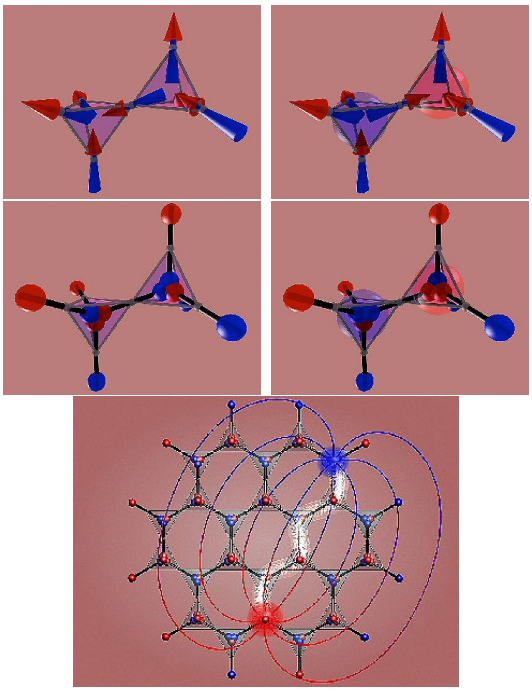


Figure 345: Mapping from dipoles to dumbbells. The dumbbell picture (c,d) is obtained by replacing each spin in (a,b) by a pair of opposite magnetic charges placed on the adjacent sites of the diamond lattice. Left: two neighboring tetrahedra obeying the ice rule, with two spins pointing in and two out, giving zero net charge on each site. Right: Inverting the shared spin generates a pair of magnetic monopoles (diamond sites with net magnetic charge). This configuration has a higher net magnetic moment and it is favoured by an applied magnetic field oriented upward in the figure (corresponding to a [111] direction). Large panel: A pair or separated monopoles (large red and blue spheres). A chain of inverted dipoles ('Dirac string') between them is highlighted, and the magnetic field lines are sketched.

Indeed, spin ice was found to have a residual entropy at low T [5], which is well-approximated by the famous Pauling entropy for (water) ice, $S \approx S_P = (1/2)\log(3/2)$ per spin. Pauling's entropy measures the huge ground state degeneracy arising from the so-called ice rules. In the context of spin ice, its observation implies a macroscopically degenerate ground state manifold obeying the "ice rule" that two spins point into each vertex of the diamond lattice, and two out.

We contend that excitations above this ground state manifold, i.e., defects that locally violate the ice rule, are magnetic monopoles with the necessary long distance properties. From the perspective of the seemingly local physics of the ice rule, the emergence of monopoles would at first sight seem rather surprising. To demystify it we will probe deeper into how the long range magnetic interactions contained in Eq. 131 give rise to the ice rule in the first place. We then incorporate insights from recent progress in understanding the entropic physics of spin ice, and the physics of fractionalisation in high dimensions[11, 12, 13, 14, 15], of which our monopoles will prove to be a classical instance.

Consider a modest deformation of Eq. 131, loosely inspired by Nagle's old work[16] on the 'unit model' description of water ice: we replace the interaction energy of the magnetic dipoles living on pyrochlore sites by that of dumbbells consisting of equal and opposite magnetic charges that live at the ends of the diamond bonds (see Fig. 345). The two ways of assigning charges on each diamond bond reproduce the two orientations of the original dipole. Demanding that the dipole moment of the spin be reproduced quantitatively fixes the value of the charge at $\pm\mu/a_d$, where the diamond lattice bond length $a_d = \sqrt{3/2}\,a$.

The energy of a configuration of dipoles is computed as the pairwise interaction energy of

magnetic charges, given by the magnetic Coulomb law

$$V(r_{\alpha\beta}) \;=\; \begin{cases} \frac{\mu_0}{4\pi}\frac{Q_\alpha Q_\beta}{r_{\alpha\beta}} & \alpha \neq \beta \\ \frac{1}{2}v_o Q_\alpha^2 & \alpha = \beta, \end{cases} \tag{132}$$

where Q_α denotes the total magnetic charge at site α in the diamond lattice, and $r_{\alpha\beta}$ is the distance between two sites. The finite "self-energy" $v_0/2$ is required to reproduce the net nearest neighbour interaction correctly. This expression – derived in detail in the annexed supplementary notes – is equivalent to the dipolar energy Eq. 131 up to corrections which are small everywhere, and vanish with distance at least as fast as $1/r^5$.

Consider first the ground states of the system. The total energy is minimised if each diamond lattice site is net neutral, i.e., we must orient the dumbbells so that $Q_\alpha = 0$ on each site. But this is just the above-mentioned ice rule, as illustrated in Fig. 345. Thus, one of the most remarkable features of spin ice pops right out of the dumbbell model: the measured low-T entropy agrees with the Pauling entropy (which follows from the short-distance ice rules), even though the dipolar interactions are long-ranged.

We now turn to the excited states. Naively, the most elementary excitation involves inverting a single dipole / dumbbell to generate a local net dipole moment 2μ. However, this is misleading in a crucial sense. The inverted dumbbell in fact corresponds to two adjacent sites with net magnetic charge

$$Q_\alpha = \pm q_m = \pm 2\mu/a_d, \tag{133}$$

a nearest neighbour monopole-antimonopole pair. As shown in Fig. 345, the monopoles can be separated from one another without further violations of local neutrality by flipping a chain of adjacent dumbbells. A pair of monopoles separated by a distance r experiences a Coulombic interaction, $-\mu_0 q_m^2/(4\pi r)$, mediated by monopolar magnetic fields, see Fig. 346.

This interaction is genuinely magnetic, hence the presence of the vacuum permeability μ_0, and not $1/\epsilon_0$. It takes only a finite energy to separate the monopoles to infinity (i.e., *they are deconfined*), and thus they are the true elementary excitations of the system: the local dipolar excitation fractionalizes!

The reader may now feel that we have slipped the baby in with the bathwater – that in taking the pictures from the dumbbell representation seriously we are somehow introducing monopoles where there were none to begin with. In general, it is of course well known that a string of dipoles arranged head to tail realises a monopole-antimonopole pair at its ends[18]. However, in order to obtain *deconfined* monopoles, it is essential that the cost of creating such a string of dipoles remain bounded as its length grows, i.e., that the relevant string tension vanish. This is evidently not true in a vacuum (such as that of the universe) where the growth of the string can only come at the cost of creating additional dipoles. Magnetic materials, which come equipped with vacua (ground states) filled with magnetic dipoles, are more promising. However, even here a dipole string is not always a natural excitation, and when it is – e.g., in an ordered ferromagnet – a string of inverted dipoles is accompanied by costly domain walls along its length (except, as usual, for one-dimensional systems[19]), causing the incipient monopoles to remain confined.

The magic of spin ice arises from its exotic ground states. The ice rule can be viewed as requiring that two dipole strings enter and exit each site of the diamond lattice. In a typical spin ice ground state, there is a "soup" of such strings: many dipole strings – of arbitrary size and shape – can be identified which connect a given pair of sites. Inverting the dipoles along any one such string creates a monopole-antimonopole pair on the sites at its ends. The associated energy cost does not diverge with the length of the string, unlike the case of an ordered ferromagnet, as no domain walls are created along the string, and the monopoles are thus deconfined.

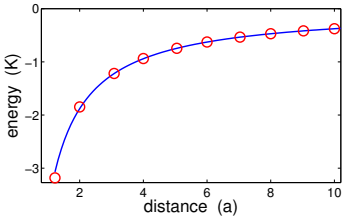


Figure 346: Monopole interaction. Comparison of the magnetic Coulomb energy $-\mu_0 q_m^2/(4\pi r)$ (Eq. 132, solid line) with a direct numerical evaluation of the monopole interaction energy in dipolar spin ice (Eq. 131, open circles), for a given spin-ice configuration (Fig. 345), as a function of monopole separation.

Notice that we did not make reference to the famous Dirac condition[17] that the fundamental electric charge e and any magnetic charge q must exhibit the relationship $eq = nh/\mu_0$ whence any monopoles in our universe must be quantised in units of $q_D = h/\mu_0 e$. This follows from the monopole being attached to a Dirac string which has to be un*observable*. [18] By contrast, the string soup characteristic of spin ice at low temperature makes the strings energetically un*important*, although they are observable and thus not quantised.

Indeed, the monopoles in spin ice have a magnitude

$$\begin{aligned} q_m &= \frac{2\mu}{a_d} \\ &= \frac{2\mu}{\mu_b}\frac{\alpha\lambda_C}{2\pi a_d} q_D \\ &\approx \frac{q_D}{8000}, \end{aligned} \tag{134}$$

where λ_C is the Compton wavelength for an electron, and α is the fine-structure constant. Amusingly, the charge of a monopole in spin ice can even be tuned *continuously* by applying pressure, as this changes the value of μ/a_d.

The monopoles are sources and sinks of the magnetic field $\mathbf{H}$, as is appropriate to the condensed matter setting. More precisely, like in other instances of fractionalisation[20], one can define a 'smeared' magnetic charge where $\nabla \cdot \mathbf{H}$ is the divergence of the magnetic field. For a monopole at the origin, separated by $L \gg \xi \gg a$ from any other monopoles, this gives $\rho_m(0) = \pm q_m$. The form of the magnetic induction $\mathbf{B}$ is also monopolar, but with the important difference that a compensating flux travels along the (unquantised) "Dirac string" of flipped dipoles created along with the monopole (see Fig. 345).

Our magnetic monopoles would in principle show up in one of the best-known monopole searches, the Stanford experiment to detect fundamental magnetic monopoles from cosmic radiation. This experiment is based on the fact that a long-lived current is induced in a superconducting ring when a monopole passes through it[8]. One can easily check that the presence of the Dirac string of flipped dipoles is immaterial to the establishment of a current.

The above observations are the central qualitative results of our work: ice-rule-violating defects are deconfined monopoles of $\mathbf{H}$, they exhibit a genuine magnetic Coulomb interaction, Eq. 132, and they produce Faraday electromotive forces in the same way as elementary monopoles would.

We reemphasize that the ice rule alone does not permit a consistent treatment of the excited states of the physical problem: crucially, the energetic interaction between our defects is absent altogether. Also, in previous discussions of the purely ice-rule problem and related short ranged problems[11, 12, 13] it has been noted that the defects do acquire a purely *entropic* Coulombic (i.e., $1/r$) interaction that has a strength that vanishes proportionally to T at low temperatures. This interaction will be present in addition to the magnetic Coulomb interaction discussed in

this paper, and is clearly much smaller as $T \to 0$. Also, it will *not* be accompanied by a magnetic field, it will not renormalise the monopole charge, and it will not be felt by a stationary magnetic test particle embedded in the lattice but not attached to a lattice site.

The most satisfactory way to demonstrate the presence of a monopole would be to measure the force on magnetic test particles, say by a Rutherford scattering experiment or by clever nanotechnological means. Unfortunately, for lack of elementary magnetic monopoles, one would have to choose dipoles as test particles, which significantly weakens such signatures.

An alternative strategy is to look for consequences of the presence of magnetic monopoles in the *collective* behaviour of spin ice. This is most elegantly achieved by applying a magnetic field in the [111] crystallographic direction. Such field acts as a (staggered) chemical potential, see Fig. 345, favouring the creation of monopoles of a given sign on either sublattice of the diamond lattice.

We thus have a tuneable lattice gas of magnetic monopoles on the diamond lattice. The basic structure of the phase diagram as a function of magnetic field and temperature can be inferred from a series of papers by M. E. Fisher and collaborators[21] in the context of ionic lattice gases and Coulombic criticality. At high T, there is no phase transition but a continuous crossover between the high- and low-density phases as the chemical potential is varied. At low T, a first-order phase transition separates the two regimes. This transition terminates in a critical point at (h_c, T_c), not unlike the liquid-gas transition of water. This serves as a useful diagnostic as the liquid-gas transition is absent for a nearest-neighbour spin-ice model, in which defects interact only entropically. In that case, it is known that there cannot be a first-order transition in the limit of low T [22].

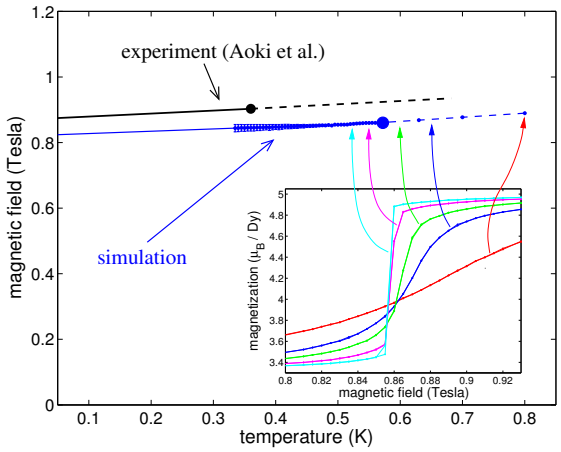


Figure 347: Phase diagram of spin ice in a [111] field. The location of the monopole liquid-gas transition from numerics (blue line) compared to experiment (black line). The solid line is a first-order transition terminated by a critical endpoint (fat circle). The dashed lines are crossovers. Inset: magnetisation curves showing the onset of first-order behaviour as the temperature is lowered. Our simulations cover the range $0.335\,\mathrm{K} < T < 0.8\,\mathrm{K}$ for 1024 spins. At the lowest temperatures, the parallel tempering code we use in our simulations of the Ewald-summed dipolar interaction no longer completely suppresses the hysteresis, and we have extended the first-order transition line using Clausius-Clapeyron.

To confirm this scenario, we have demonstrated by Monte Carlo simulations that the actual phase diagram of dipolar spin ice model has precisely this structure. In order to rule out the appearance of the liquid-gas transition due to effects introduced by the approximations leading to Eq. 132, we simulated directly the original dipolar spin ice model, Eq. 131. The resulting phase diagram is depicted in Fig. 347. The critical endpoint is located around $(T_c, h_c) =$

$(0.57 \pm 0.06\text{K}, 0.86 \pm 0.03\text{T})$. The error bars are mainly due to finite-size effects, as the intensive nature of the simulations of long-range dipolar interactions prohibits simulating very large systems.

This scenario is indeed observed experimentally in spin-ice materials[6, 7], and our results provide a natural explanation. Spin ice in a [111] magnetic field is a very interesting problem which has already attracted considerable attention. The low-density phase of monopoles is known as kagome ice, a quasi two-dimensional phase with algebraic correlations and a reduced residual entropy[6, 7, 23]. The high-density phase is an ordered state with maximal polarisation along the field direction. Experimental results on the liquid-gas transition and its endpoint are also displayed in Fig. 347 for comparison. Our numerical results are in good qualitative agreement with both experiment and the analytic calculations of Ref. [21]. Our value of the critical field agrees with Ref. [6] within a few percent, less than the uncertainty due to demagnetisation effects[6, 7]. However, the experimental value of T_c is about a third lower than the numerical one, most likely due to farther-neighbour (exchange) interaction terms, which despite being small can shift the location of a transition temperature considerably[10].

The presence of a liquid-gas transition was noted to be very remarkable as there are few, if any, other experimentally known instances in localised spin systems[6]. No mechanism was known to account for this phenomenon, and our theory of magnetic monopoles fills this gap.

The existence of magnetic monopoles in a condensed matter system is, given their intellectual pedigree, exciting in itself. (The reader familiar with the beautiful work on the anomalous Hall effect will notice that the monopoles appearing there are not excitations and do not involve the physical magnetic field[24].) This is further heightened by their being a rare instance of high-dimensional fractionalisation, a phenomenon of great interest in fields as diverse as correlated electrons and topological quantum computing[25]. We hope our analysis will encourage our experimentalist colleagues to try their hand at directly detecting these monopole. There are surely many avenues to explore in search of useful signatures: scattering, transport and noise measurements, flux detection, to mention just a few.

The dumbell picture

This material presents a detailed derivation of the dumbell Hamiltonian used extensively in our paper. We start from the generally accepted Hamiltonian which contains a sum of nearest-neighbour exchange and long range dipolar interactions,

$$\begin{aligned} H &= \frac{J}{3} \sum_{\langle ij \rangle} S_i S_j \\ &+ Da^3 \sum_{(ij)} \left[\frac{\hat{e}_i \cdot \hat{e}_j}{|\mathbf{r}_{ij}|^3} - \frac{3\left(\hat{e}_i \cdot \mathbf{r}_{ij}\right)\left(\hat{e}_j \cdot \mathbf{r}_{ij}\right)}{|\mathbf{r}_{ij}|^5} \right] S_i S_j . \end{aligned} \tag{135}$$

The magnetic moment of a spin is denoted by μ, which equals approximately 10 Bohr magnetons ($\mu = 10\mu_B$) for the spin ice compounds discussed here (namely, $Dy_2Ti_2O_7$ and $Ho_2Ti_2O_7$). The distance between spins is r_{ij}, and $a \simeq 3.54\text{Å}$ is the pyrochlore nearest-neighbour distance. $D = \mu_0 \mu^2/(4\pi a^3) = 1.41\text{K}$ is the coupling constant of the dipolar interaction.

A dipole can be thought of as a pair of equal and opposite charges of strength $\pm q$, separated by a distance $\tilde{a}$, such that $\mu = q\tilde{a}$. The dipolar part of the Hamiltonian is reproduced exactly in the limit $\tilde{a} \to 0$. Here we choose $\tilde{a}$ to equal the diamond lattice constant $a_d = \sqrt{3/2}\, a$, therefore fixing $q = \mu/a_d$. The two ways of assigning charges reproduce the two orientations of the original dipole, as illustrated in Fig. 348.

Given a configuration of N dipoles, let us label $\{q_i,\ i = 1, \ldots, 2N\}$ the $2N$ charges in the corresponding dumbell configuration. The magnetic Coulomb interaction between the charges is given by

$$\mathcal{V}(r_{ij}) \;=\; \begin{cases} \frac{\mu_0}{4\pi}\frac{q_i q_j}{r_{ij}} & r_{ij} \neq 0 \\ v_0 q_i q_j & r_{ij} = 0, \end{cases} \tag{136}$$

which reproduces the interaction between two dipoles exactly in the limit of large separation between them (modulo an innocuous "'self-energy" term discussed below). We can also tune the onsite contribution v_0 so as to match the interaction energy $\pm J_{\rm eff} = \pm(J + 5D)/3$ between two neighbouring dipoles on the lattice (see for example Fig. 349). For two neighbouring spins both pointing into the shared tetrahedron (right panel of Fig. 349) one has

$$\begin{aligned} J_{\rm eff} &= v(0) - 2v(r_{12}) - 2v(r_{23}) + v(r_{13}) \\ &= v(0) - v(r_{12}) - v(r_{23}) + 2E_{\rm self} + v(r_{13}), \end{aligned} \tag{137}$$

where $v(r) = |\mathcal{V}(r)|$, and $E_{\rm self} = -v(r_{12}) = -v(r_{23})$ is the self-energy of a dipole in the dumbell picture. For two neighbouring spins pointing into and out of the shared tetrahedron (as in the left panel of Fig. 349), we obtain

$$\begin{aligned} -J_{\rm eff} &= -v(0) + 2v(r_{12}) + 2v(r_{23}) - v(r_{13}) \\ &= -v(0) + v(r_{12}) + v(r_{23}) + 2E_{\rm self} - v(r_{13}). \end{aligned} \tag{138}$$

Taking the difference between the two equations above, and using the fact that $r_{12} = r_{23}$ and therefore $v(r_{12}) = v(r_{23})$, one gets

$$v(0) = J_{\rm eff} + 2v(r_{12}) - v(r_{13}) \tag{139}$$

where

$$\begin{aligned} r_{12} &= a_d \\ r_{13} &= 2a \\ v(r) &= \begin{cases} \frac{\mu_0}{4\pi}\frac{(\mu/a_d)^2}{r} & r \neq 0 \\ v_0 \left(\frac{\mu}{a_d}\right)^2 & r = 0 \end{cases} \end{aligned}$$

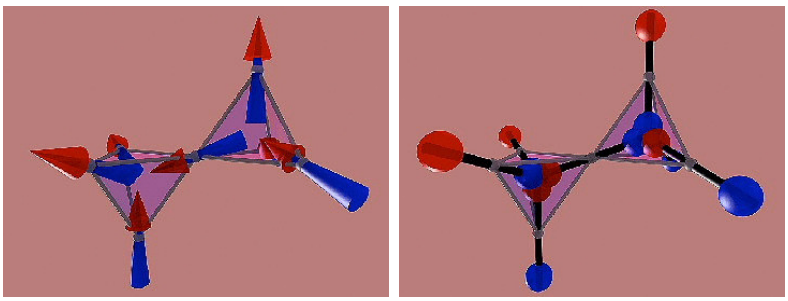


Figure 348: Mapping from dipoles to dumbells. Left: two neighboring tetrahedra obeying the ice rule, with two spins pointing in and two out, giving zero net charge on each site. Right: The corresponding dumbell picture obtained by replacing each spin by a pair of opposite magnetic charges placed on the adjacent sites of the diamond lattice.

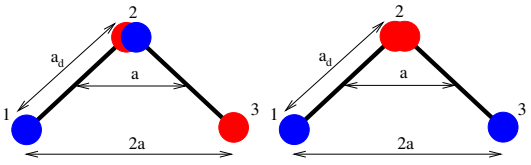


Figure 349: Schematic examples of two neighbouring dipoles on the pyrochlore lattice.

and in the end

$$v(0) = \frac{J+5D}{3} + \frac{\mu_0}{4\pi}\left(\frac{\mu}{a_d}\right)^2\left[\frac{2}{a_d} - \frac{1}{2a}\right] \tag{140}$$

Using the fact that $D = \mu_0\mu^2/4\pi a^3$ and that $a = a_d\sqrt{2/3}$, the expression above reduces to

$$v_0\left(\frac{\mu}{a_d}\right)^2 = \frac{J}{3} + \frac{4}{3}\left[1+\sqrt{\frac{2}{3}}\right]D. \tag{141}$$

Now, the energy of a charge configuration can be straightforwardly resummed and written in terms of the total charges Q_α on each site α of the diamond lattice, i.e., $Q_\alpha = q_{i_1}+q_{i_2}+q_{i_3}+q_{i_4}$ for the quartet of charges with $r_{i_1} = r_{i_2} = r_{i_3} = r_{i_4} \equiv r_\alpha$. The energy can thus be computed from a pairwise Coulomb interaction of the magnetic charges Q_α:

$$V(r_{\alpha\beta}) \;=\; \begin{cases} \frac{\mu_0}{4\pi}\frac{Q_\alpha Q_\beta}{r_{\alpha\beta}} & \alpha \neq \beta \\ \frac{1}{2}v_0 Q_\alpha^2 & \alpha = \beta, \end{cases} \tag{142}$$

This gives the same total energy as Eq. 136 up to an unimportant overall constant $(1/2)\sum q_i^2 = N(\mu/a_d)^2$. Thus, it is also equivalent to the dipolar energy Eq. 135 up to corrections which are small everywhere, and vanish with distance at least as fast as $1/r^5$ for each dipole pair. In particular, one can show that v_0 is large enough, for the compounds we consider, that it enforces $Q_\alpha = 0$ everywhere in the ground state. These are just the ice rules, and the degeneracy of all the states follows from the fact that all terms with $r_{\alpha\beta} > 0$ vanish for $Q_\alpha \equiv 0$.

Part VI
Experiments

18 High magnetic fields for fundamental physics

Rémy Battesti, Jerome Beard et al.

Abstract

Various fundamental-physics experiments such as measurement of the birefringence of the vacuum, searches for ultralight dark matter (e.g., axions), and precision spectroscopy of complex systems (including exotic atoms containing antimatter constituents) are enabled by high-field magnets. We give an overview of current and future experiments and discuss the state-of-the-art DC- and pulsed-magnet technologies and prospects for future developments.

18.1 Introduction

While production of high magnetic fields is a mature field, recent years have seen significant developments in both pulsed and continuous-operation (often referred to as "DC") magnets. These developments are driven by applications in materials sciences, nuclear magnetic resonance (NMR), thermonuclear-energy research, particle accelerators, and particle detectors (see, for example, a High-Magnetic-Field study by the US National Academy of Sciences [1]).

In this review we discuss the state-of-the-art of magnet technologies with an eye towards their application to "small-scale" fundamental physics such as searches for magnetic birefringence of the vacuum and for axions/axion-like particles (ALPs) with haloscope and helioscope devices and light-through-wall experiments, etc.

There are several laboratories around the world (Fig. 350 and Table 350) pursuing research into high magnetic fields, including their production and scientific applications. In Europe, there is the French Laboratoire National des Champs Magnétiques Intenses (LNCMI) which has two sites, one in Grenoble with a specialization in DC magnets and one in Toulouse with a specialization in pulsed magnets, the German Dresden High Magnetic Field Laboratory (Hochfeld-Magnetlabor Dresden, HLD), and the High Field Magnet Laboratory (HMFL) in Nijmegen, the Netherlands. The European high-field laboratories operate as coordinated infrastructure within the framework of the European Magnetic Field Laboratory (EMFL). In America, there is the US National High Magnetic Field Laboratory (NHMFL) with one location at Los Alamos, New Mexico, and two locations in Florida (Tallahassee and Gainesville). In Asia, there are two laboratories in China, the High Magnetic Field Laboratory of the Chinese Academy of Sciences (CHMFL) in Hefei and the Wuhan National High Magnetic Field Center (WHMFC), and four laboratories in Japan: the Tsukuba Magnet Laboratory (TML), the High Field Laboratory for Superconducting Materials in Sendai, the International Megagauss Science Laboratory (IMGSL) in Kashiwa, and the Center for Advanced High Magnetic Field Science in Osaka. The laboratories and the parameters of some of their magnets are listed in Table 350.

Generation of magnetic fields involves flowing an electric current through a conductor (typically, a coil), with the field intensity B being proportional to the current I. For a coil with a fixed resistance R, the heat dissipation I^2R is proportional to the square of the magnetic field. At the same time, mechanical pressure also scales as B^2, with a proportionality coefficient of approximately $4\,\mathrm{atm/T^2}$. Heating and mechanical forces are the two essential challenges for the design of high-field electromagnets. The stored energy in magnetic field is proportional

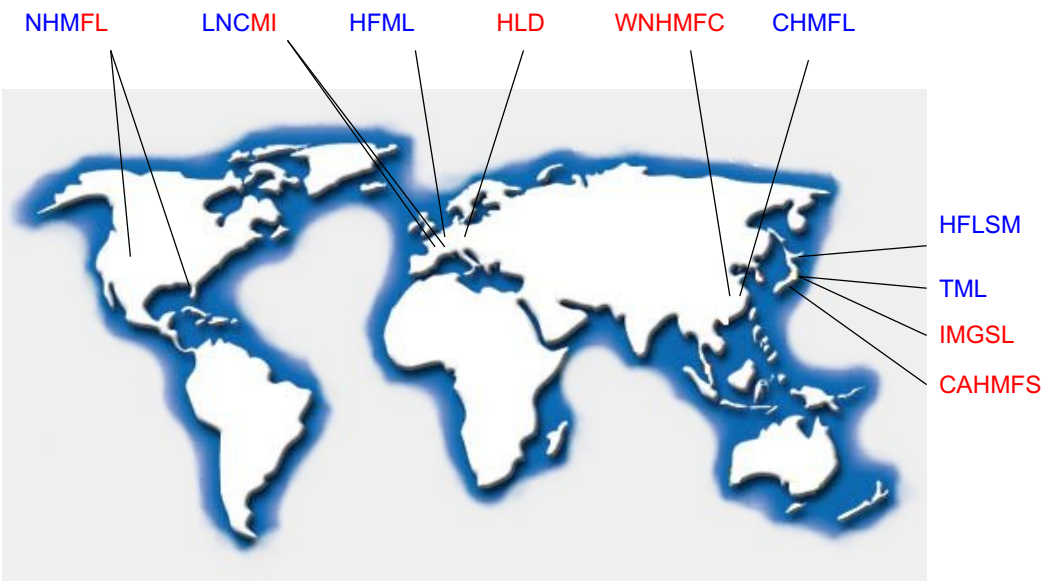


Figure 350: Laboratories around the world pursuing high pulsed (red) and DC (blue) magnetic fields.

to the integral of B^2 over the volume of the magnet, so the size of the field volume is an important characteristic of the magnet along with the maximum field. Another key parameter is, of course, whether a magnet is DC or pulsed and, in the latter case, additional important parameters are the duration and temporal profile of the pulse, as well as the pulse repetition rate.

The approaches to meeting the heat-dissipation and mechanical-stability challenges are different for different types of magnets as summarized in Fig. 351 and discussed in Sec. 18.4. However, in all cases, materials are crucially important for the development of advanced magnets; such development is thus a task at the intersection of physics, engineering, and materials sciences.

Some of the world's highest-field nondestructive magnets are listed in Table 27. However, it is frequently the case that the highest-field magnet is not necessarily the best choice for a particular experiment. Indeed, as summarized in Table 28, the figure-of-merit (FOM), depending on the experiment, could be the stored energy B^2V (where B is the magnetic field and V is the field volume), or the effective B^2L (where L is the length of the field), or something completely different such as the ability to tune the magnet to a desired field value, or the ability to tune the field over a broad range, while maintaining spatial homogeneity and low ripple.

Maxwell's equations possess symmetry with respect to electric and magnetic fields. While this symmetry is explicitly seen in Gaussian units where electric field is measured in the same units as magnetic induction (gauss), the symmetry is somewhat obscured in the International System. To compare electric and magnetic fields, one may use the fact that 1 T corresponds to $3 \cdot 10^8$ V/m. This allows us to "compare" strong laboratory electric and magnetic fields. For example, the highest static electric fields that can be produced in macroscopic volumes are significantly below 10^8 V/m, which is a mere 0.3 T! Clearly, magnetic fields win at the "strong-field game," particularly when one is studying the properties of the vacuum such as field-induced birefringence!

The present review is structured as follows. We begin with a discussion of some of the current fundamental-physics experiments (Sec. 18.2) using strong fields, leaving aside many important topics such as magnets for accelerators, high-energy detectors, NMR, thermonuclear fusion, and neutron scattering, only mentioning some of the corresponding developments in

Table 26: High-magnetic-field laboratories around the globe. SD stands for "semi-destructive, SC stands for superconducting, " B/T is magnetic field/temperature; the dimension stands for the diameter of the magnet's bore. "Best performance" does not include many practical factors such as whether it represents routine operation or a rare record and to what extent the facility is available to users.

Name	Location	Type	Best performan
Laboratoire National des Champs Magnetiques Intenses			
- DC-field facility	Grenoble, France	DC	37 T in 34 mm
- pulsed-field facility	Toulouse, France	pulsed	98.8 T in 8 mm
			209 T in 8 mm
High Field Magnet Laboratory	the Netherlands	DC	38 T in 32 mm
Hochfeld Labor Dresden	Dresden, Germany	pulsed	95.6 T in 16 m
National High Magnetic Field Laboratory			
- DC-field facility	Tallahassee, USA	DC	45 T in 32 mm
- pulsed-field facility	Los Alamos, USA	pulsed	100.7 T in 10 m
			240 T in 8 mm
- high-B/T facility	Gainesville, USA	DC	16.5 T at 0.4 m
Wuhan National High Magnetic Field Center	Wuhan, China	pulsed	90.6 T in 12 m
Chinese High Magnetic Field Laboratory	Heifei, China	DC	40 T in 32 mm
International Megagauss Science Laboratory	Kashiwa, Japan	pulsed	75 T in 17 mm
			250 T in 8 mm
High Field Laboratory for Superconducting Materials	Sendai, Japan	DC	30 T in 32 mm
			24.6 T in 52 m
Center for Advanced High Magnetic Field Science	Osaka, Japan	pulsed	60 T
Tsukuba Magnet Laboratory	Tsukuba, Japan	DC	32 T in 52 mm
			35 T in 30 mm

passing. We then present several ideas for future experiments in Sec. 18.3. State-of-the-art magnet technologies are discussed in Sec. 18.4, while magnetic-field measurement techniques are discussed in Sec.18.5. Conclusions and outlook on the future of the field are presented in Sec. 18.6.

Table 27: Highest-field magnets currently in operation. Highest field does not necessarily imply high figure-of-merit (see text).

Country	Field (T)	Laboratory	feature
DC magnets			
USA	45	NHMFL	hybrid
China	40	CHMFL	hybrid
Europe	38	HFML	
Japan	35	TML	hybrid
Pulsed nondestructive magnets			
USA	100.7	NHMFL	
Europe	98.8	LNCMI	
China	90.6	WHMFC	
Japan	86	IMGSL	

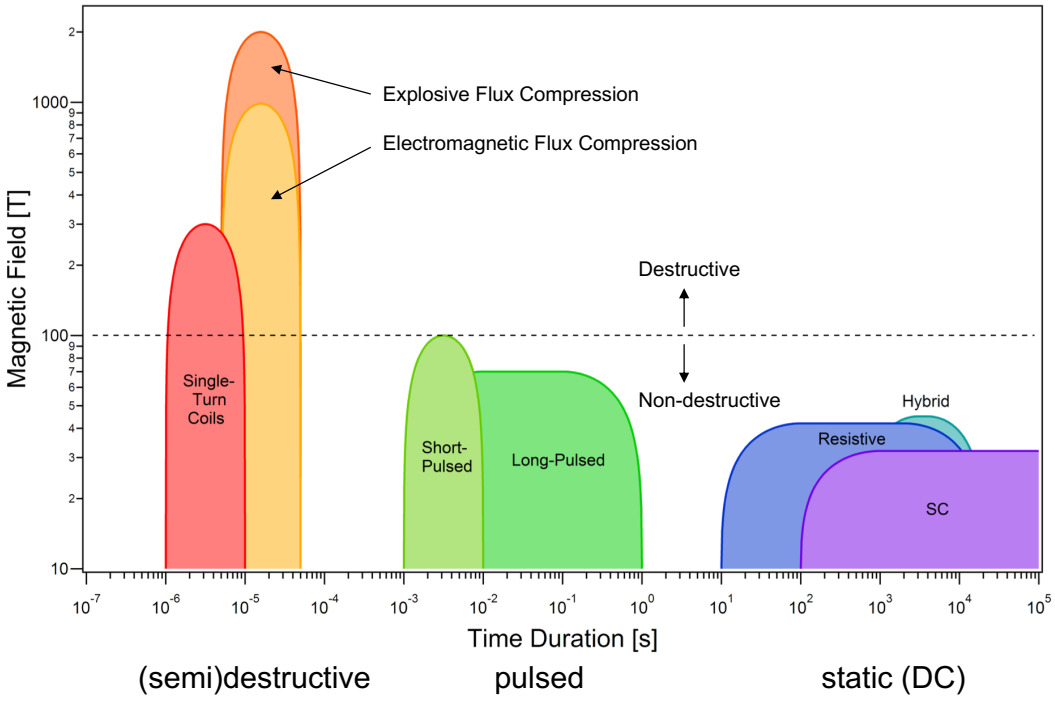


Figure 351: Overview of strong-magnet methods; see also Table 26.

Table 28: Figure-of-merit (FOM) for various fundamental-physics experiments. L, A and V are the characteristic length, transverse area and volume of the magnetic-field region. The last column lists the sections of this paper where the corresponding experiments are discussed.

Experiment	FOM	Examples	Section
Vacuum birefringence	B^2L	BMV, PVLAS, OVAL	18.2.1
Light shining through wall	B^4L^4	ALPS, OSQAR, ...	18.2.2
Helioscope	$B^2L^2A = B^2VL$	CAST, IAXO	18.2.2
Haloscope (Primakoff)	B^2V	ADMX, HAYSTAC, ORGAN, CULTASK ...	18.2.2
Haloscope (other)	None of the above	CASPEr, QUAX, ...	18.2.2

18.2 Experiments

In this section, we discuss several classes of experiments that use strong magnetic field to probe the quantum vacuum close to the conditions where nonlinear-electrodynamics effects such as vacuum optical birefringence (Sec. 18.2.1) are expected to be observable and that are sensitive to various beyond-the-standard-model scenarios. One of the major goals of high-magnetic-field experiments is a search, either direct or indirect, for possible constituents of dark matter (Sec. 18.2.2).

18.2.1 Vacuum birefringence

Vacuum-birefringence tests with quasi-stationary magnetic fields

In the framework of classical electrodynamics, governed by Maxwell's linear equations, vacuum isotropy and the speed of light are not affected by the application of electromagnetic fields. This property distinguishes vacuum from material medium. Going beyond classical electrodynamics, for example, by accounting for virtual particle-antiparticle fluctuations in the framework of quantum electrodynamics (QED), mediating effective interactions among electromagnetic fields [2, 3, 4], one finds that nonlinear phenomena such as photon-photon scattering [5, 6, 7] and magnetic-field-induced birefringence of vacuum [8, 9, 10] exist, albeit in regimes that are not easy to achieve experimentally. At the same time, QED can be seen as a special case of nonlinear electrodynamics (NLED) [11], therefore, a more general framework exists that allows a uniform description and comparison of different experiments [12] that test QED predictions and constrain alternative models that predict bigger nonlinear effects. A prominent example is Born-Infeld electrodynamics [13]. Historically, nonlinear self-interaction of electromagnetic field as a consequence of the Dirac theory was first considered by Heisenberg and Euler [2]; see the pertinent historical review [14]. For a modern theoretical perspective, see [15, 16, 17] and references therein. While there are indications for the relevance of vacuum birefringence for explaining the observed linear polarization of the light from a neutron star [18, 19], unambiguous detection in the laboratory is being attempted by several groups.

The general idea for vacuum-birefringence experiments [20] is illustrated in Fig. 352. The

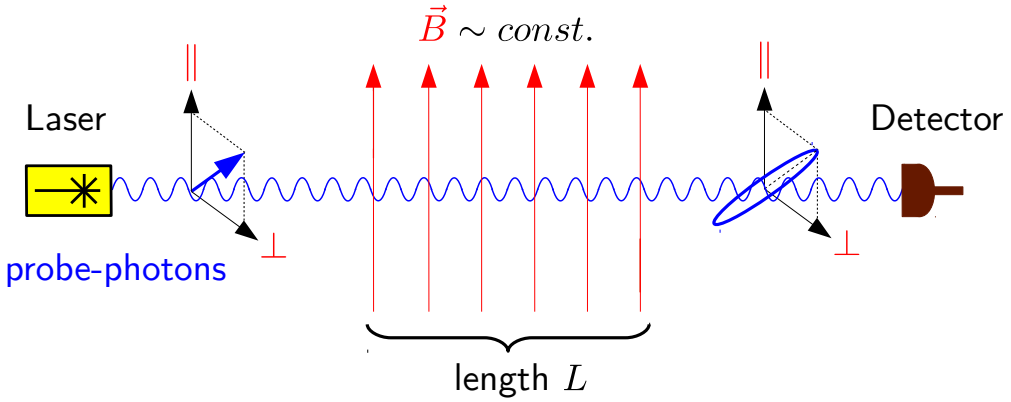


Figure 352: The concept of the vacuum-birefringence experiments.

initial polarization of the light is at 45° to the magnetic field; the difference in the light speed for the light polarized along and perpendicular to the magnetic field induces light ellipticity

given by

$$\Phi = 2\pi \frac{L}{\lambda} \Delta n \,, \tag{143}$$

where L is the length of the magnetic field, λ is the wavelength of the light, and Δn is the difference of the refractive indices for light polarized parallel and perpendicular to the magnetic field (the Cotton-Mouton effect[71] of the vacuum). More specifically, we have (in SI units)

$$\Delta n = n_{\parallel} - n_{\perp} = k_{\mathrm{CMV}} |\vec{B}|^2 \,, \tag{144}$$

where k_{CMV} is the Cotton-Mouton coefficient of the vacuum. According to standard QED,

$$k_{\mathrm{CMV}} = \frac{1}{\mu_0} \frac{\alpha}{\pi} \frac{1}{30} \left(\frac{e\hbar}{m^2 c^2} \right)^2 \approx 4.0 \cdot 10^{-24} \mathrm{T}^{-2}, \tag{145}$$

where μ_0 is vacuum permeability, $\alpha \approx 1/137$ is the fine structure constant and m is the electron mass. Up to a numerical constant, the difference of the refractive indices due to the vacuum Cotton-Mouton effect of Eq. (144) is just the square of the ratio of magnetic energy of a magnetic dipole equal to a Bohr magneton immersed in field B, $e\hbar B/(2m)$, and the electron rest energy, mc^2. The quantity $\left(\frac{e\hbar}{m^2c^2}\right)^{-1} \sim 10^{10}\,\mathrm{T}$ is the critical magnetic field. Its electric-field analogue governs the onset of spontaneous electron-positron pair production. Upon propagation of initially linearly polarized light traveling perpendicular to a quasi-static magnetic field, the magnetic field induced change in light polarization is analyzed (Fig. 352). With practically achievable values of B^2L, the expected optical phase shifts are small and challenging to detect with current polarimetry techniques, though the experiments are getting tantalizingly close to the QED predictions (Fig. 353).

As noted above, as long as the transverse dimensions of the magnet are sufficient to accommodate the laser beam, the figure-of-merit for a vacuum-birefringence experiment is:

$$\mathrm{FOM}_{\text{Vacuum birefr.}} = B^2 L. \tag{146}$$

Currently ongoing vacuum-birefringence experiments include the BMV (Biréfringence Magnétique du Vide) experiment in Toulouse, the PVLAS (Polarizzazione del Vuoto con Laser) experiment in Ferrara, and the OVAL (Observing Vacuum with Laser) experiment in Tokyo. Some key parameters of these experiments are summarized in Table 29. Note that it is customary to present the maximum value of the field B_{max}, the effective value of B^2L which is an integral of B^2 along the path of the light beam, as well as the effective value of L, which is the effective value of B^2L divided by maximum B^2_{max}. All current experiments employ optical cavities in order to enhance the effect of the birefringence, by a factor of the order of the cavity finesse. For pulsed experiments, "filtered" values for B^2L are listed in Table 29. What this means is the following: high-finesse optical cavities used in these experiments "trap" light for times that exceed the duration of the magnetic-field pulse. In this situation, the cavity acts as a low-pass filter for the signal. For instance, in BMV, a characteristic signal bandwidth is $\approx 100\,\mathrm{Hz}$, while the cavity cuts at $\approx 70\,\mathrm{Hz}$. The final signal is almost half of what one would calculate without taking into account the cavity filtering. One can also see the filtering effect in the following way. Photons while trapped between the cavity mirrors "see" the pulsed magnetic field changing so the effect is given by an averaged value of the magnetic field, smaller than the nominal one. The PVLAS modulating at 5 Hz has almost no cavity filtering. Because of cavity filtering, there is nothing to gain by increasing the cavity finesse much further, or by shortening the magnetic-field pulse.

[71] So named after A. Cotton and H. Mouton who studied magnetic birefringence effects in liquids in the beginning of the 20th century.

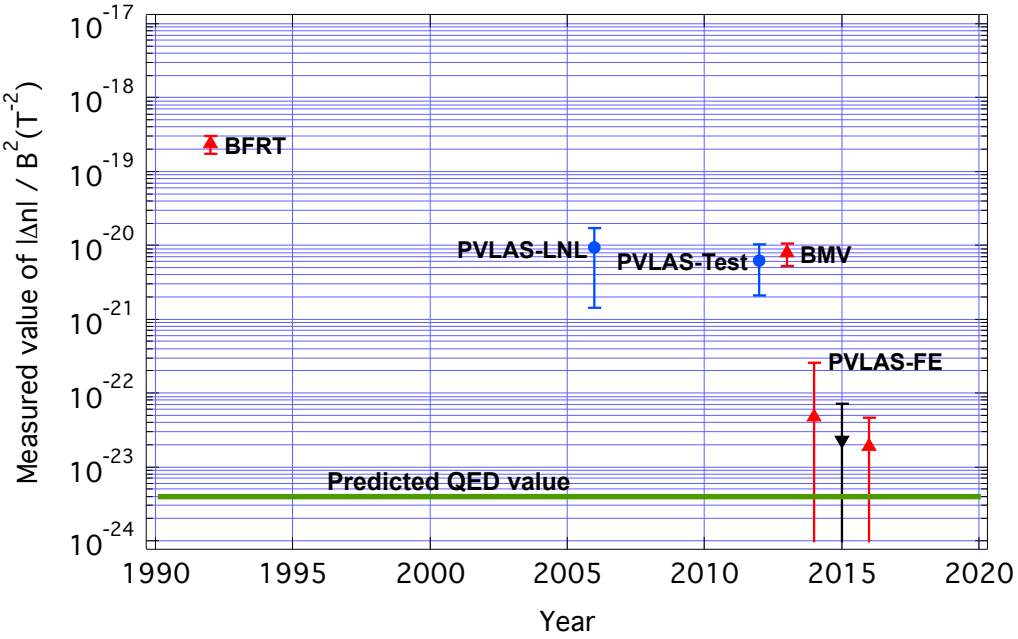


Figure 353: Time evolution of the absolute values of the measured vacuum magnetic birefringence normalized to B^2. Red up triangles refer to positive central values for $\Delta n/B^2$ whereas the blue circles and black down triangles refer to undetermined sign and negative central values, respectively. The error bars correspond to 1σ. Several measurements correspond to central values larger than the declared errors indicating the presence of undetermined systematics. The values derive from the following references: BFRT (the Brookhaven-Fermilab-Rutherford-Trieste experiment) [21], PVLAS-LNL (LNL: Italian National Laboratories in Legnaro) [22], PVLAS-Test [23], BMV [24], PVLAS-FE (FE: the University of Ferrara) [25].

Since the science run with the parameters shown in Table 29, the BMV experiment is being significantly upgraded [30]. The so-called BMV 2 is based on a new pulsed magnet developed at the LNCMI-Toulouse [31], the XXL-Coil (Fig. 354). A prototype was tested up to 30 T corresponding to $B^2L \approx 278\,\mathrm{T^2m}$ and one magnet is currently in use at LNCMI providing up to 18 T, $B^2L \approx 100\,\mathrm{T^2m}$. The rise time of the magnetic pulse is 6 ms, corresponding to a typical frequency for B^2 of about 80 Hz, which is not too far above the optical-cavity frequency cut-off, limiting the filtering effect due to the high finesse. The total pulse duration is 14 ms with cooling duration between two pulses fixed to (conservative) one minute. The repetition rate is not limited by the Joule heating in the magnet and the cooling duration could be reduced about one order of magnitude for the final data-accumulation process. Commissioning of the apparatus is nearing completion, first data runs are expected by the end of 2018.

A pulsed magnet for vacuum birefringence measurements and axion-like particle searches (Sec. 18.2.2) was developed at the University of Tokyo [32], demonstrating 9 T over a length of over 0.8 m at a 0.1 Hz repetition rate. The figure-of-merit for the vacuum birefringence experiments is $B^2L = 54\,\mathrm{T^2m}$, achieved at a repetition rate that is much higher than for other pulsed magnets used for such experiments. The system utilizes "racetrack" coils (Fig 355, left) and takes full advantage of the fact that high B^2L can be be achieved in a small field volume, resulting in relatively small heating and allowing for efficient cooling. The high repetition rate of the system is enabled by an energy-recycling scheme (Fig 355, right), where the field-pulse temporal profile is bipolar and a significant fraction of the energy is thus returned to the capacitor bank at the end of the pulse.

Operation of pulsed magnets is associated with considerable mechanical shock and acoustic

	BMV [24]	PVLAS [26, 25]	OVAL [27]
Magnet type	pulsed	rotating permanent	pulsed
Maximum field $B_{\max}$ [T]	6.5	2.5	9.0
Field length L [m]	0.14	$0.82 + 0.82$	0.17
B^2L [T^{2}m]	5.8	$5.06 + 5.06$	13.8
"Filtered" B^2L [T^{2}m]	2.7	$5.06 + 5.06$	4.1
Wavelength λ [nm]	1064	1064	1064
Cavity finesse $\mathcal{F}$	445 000	700 000	320 000
Integration time [s]	0.3	$2 \cdot 10^6$	0.6
Repetition rate [Hz]	0.0017	continuous	0.17
Uncertainty in Δn [10^{-23}/T^2]	270	2.7	110 000

Table 29: Parameters of the three operating vacuum-birefringence experiments. Effective values are given for L and B^2L, see text. Note that the OSQAR experiment at CERN originally aimed at measuring vacuum magnetic birefringence using an LHC dipole magnet providing $B^2L = 1150\,\mathrm{T}^2\mathrm{m}$ [28]; however, it has operated as a light-shining-through-wall experiment [29], see Sec 18.2.2.

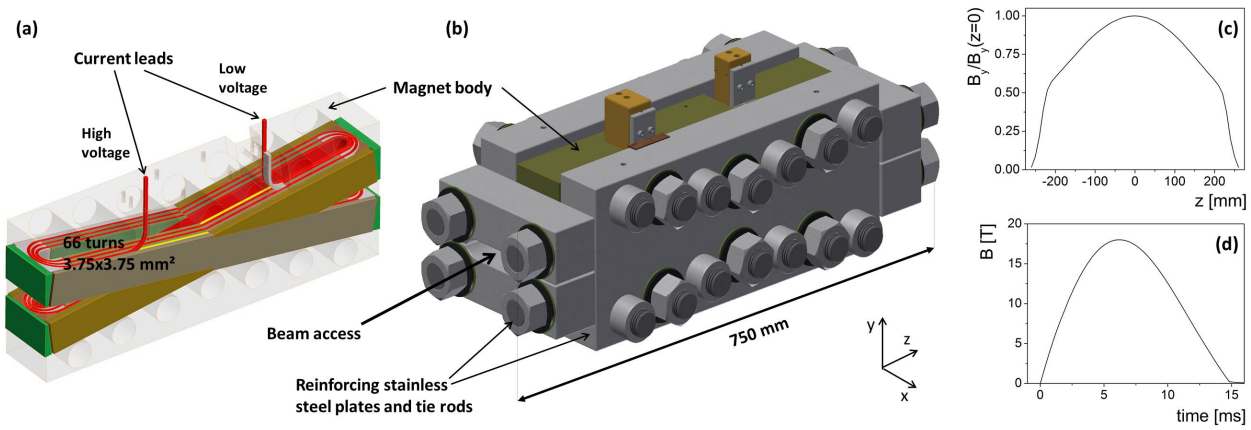


Figure 354: The LNCMI XXL-Coil (the magnet currently used in the BMV 2 experiment). (a) Some winding details of this X-shape magnet. It consists of two interlaced coils that leave openings on two sides for the beam access. The conductor (in red) is in Hogänäs Glidcop® AL-15 insulated with Kapton®. (b) Final assembly with stainless-steel reinforcement and brass connectors. The magnet is installed in a Dewar (not shown) and cooled with liquid nitrogen before a pulse. Optical-access room-temperature port is 13 mm diameter. The symmetry in the alternating layers from one branch of the X to the other limits the axial components of the magnetic field to a few percent of the total field. Copper screens (not shown) about one centimeter thick, are added on both sides to decrease Faraday component of the magnetic field by about one order of magnitude at the positions of the cavity mirrors of the BMV experiments. (c) Spatial profile of the transverse magnetic field along the beam path. (d) Temporal profile of the magnetic field for an 18 T pulse. Figure courtesy LNCMI-Toulouse.

disturbance presenting a significant challenge to the operation of precision experiments. What nevertheless makes such measurements possible is the finite propagation time of acoustic waves. For example, acoustic shock arrives at the optical elements in the OVAL experiment some 4 ms after the magnetic pulse.

The PVLAS experiment takes advantage of several techniques to maximize the signal and isolate it from possible systematics. A 3.3 m long Fabry-Perot cavity increases the single-pass

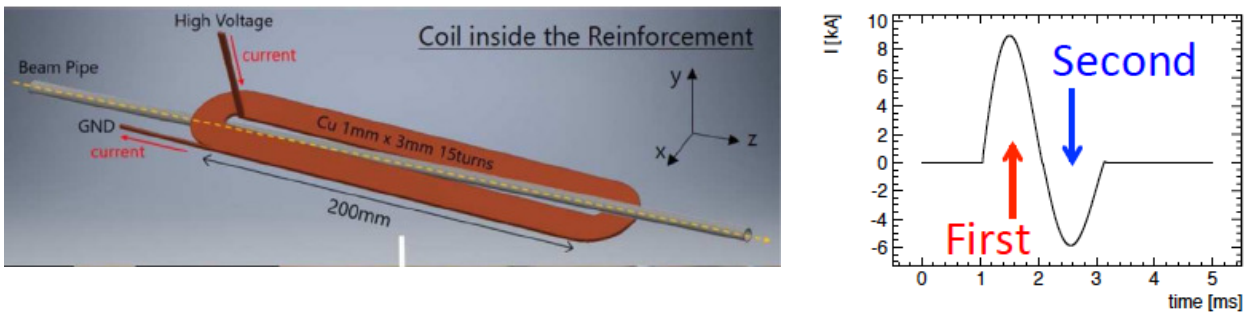


Figure 355: Left: one of the four racetrack coils used in the University of Tokyo high-repetition-rate pulsed magnet [32]. Right: temporal profile of the current during a single pulse. Part of the energy stored in the capacitor bank is recycled.

ellipticity by a factor of $N_{\rm pass} = 2\mathcal{F}/\pi$, where $\mathcal{F}$ is the finesse of the cavity (in PVLAS, $\mathcal{F} = 7{\cdot}10^5$ and the circulating light power is 40 kW; the light wavelength is 1.064 μm); a heterodyne technique is used to ensure that the observed signal is linear in the induced ellipticity (similar "tricks" are also used in other polarimetry experiments). PVLAS uses two permanent magnets, each producing 2.5 T field with field length of about 0.8 m, from AMT&C (Moscow), which are Halbach-array dipoles with 20 mm bore. The Halbach-array configuration provides "self-shielding:" the field outside the array is small. Additional ferromagnetic shielding is installed around the sides and ends of the magnets. As a result, the fields at the outer surfaces of the cylindrical magnets is less that 0.1 mT. The magnets rotate around their axis in order to modulate the signal. The two magnets can be independently rotated with frequencies up to 23 Hz, which important for identifying and suppressing systematic effects. The advantages of the permanent magnets include their relatively low cost ($\approx$ 100 k€for both magnets; no power supplies needed), the absence of running costs, as well as good mechanical stability. It was possible to balance the magnets with extra wights ($\approx$ 10 g) to minimize vibrations.

The standard QED vacuum birefringence (144) for $B = 2.5$ T is $\Delta n = 2.5 \cdot 10^{-23}$, which, for the current parameters of the PVLAS experiment corresponds to induced ellipticity of $5 \cdot 10^{-11}$ rad.

The present ellipticity sensitivity of PVLAS (which is significantly above the shot noise) is $\approx 2.5 \cdot 10^{-7}$ rad/$\sqrt{\rm Hz}$ corresponding to birefringence sensitivity $\Delta n \approx 1.2 \cdot 10^{-19}/\sqrt{\rm Hz}$, so it would take about three years of data taking to reach statistical sensitivity corresponding to the QED expectations.

What is the way forward for the vacuum-birefringence experiments? In the case of DC magnets, from the point of view of the B^2L figure-of-merit, switching to accelerator superconducting dipoles appears particularly attractive and realistic. However, rotation or sufficiently fast modulation of the field does not appear practical in this case, so the experiment would have to do with some other ways to modulate the signal. Polarization-modulation is a possible approach [33, 34], but systematics-free polarimetry with a necessary sensitivity is yet to be demonstrated. An interesting idea [35, 36, 37] involves adding magnets to the arms of a gravitational-wave detector such as VIRGO, taking advantage of the existing ultrahigh-precision optical interferometers at these facilities.

Probing vacuum birefringence with high-intensity lasers

While most of the current review is devoted to production of and experiments with static or quasi-static magnetic fields in macroscopic volumes, there are other ways to address some of the related fundamental physics using strong electromagnetic fields associated with focused ultrafast laser pulses. While we do not attempt to survey this vast research area, as one

particular example, we discuss vacuum-birefringence experiments along these lines and compare their impact to the more traditional vacuum-birefringence experiments discussed in Sec. 18.2.1.

The field strength associated with laser pulses can exceed those of quasi-static laboratory fields by many orders of magnitude [38] (albeit over much smaller lengths and time scales). For pulsed laser beams focused into a micrometer-sized spot, these could be as high as $10^5 - 10^6$ T for magnetic fields, and, correspondingly, as high as over 10^{12} V/cm for electric fields, i.e., some three orders of magnitude over the characteristic atomic field.

The principle scheme of a vacuum birefringence experiment employing a high-intensity laser field to polarize the quantum vacuum was put forward in [39], envisioning the combination of an optical high-intensity laser pulse as pump and a bright linearly polarized x-ray pulse as probe. As the birefringence-induced signal is inversely proportional to the wavelength of the probe and directly proportional to the number of photons available for probing, x-ray probes seem most promising, particularly given the tremendous progress in x-ray polarization purity measurements achieved in recent years [40, 41]. An experiment of this type will become possible, for example, at the Helmholtz International Beamline for Extreme Fields (HIBEF) at the European x-ray free-electron laser (XFEL). Its potential feasibility with state-of-the-art technology was recently demonstrated [42, 43, 44]; for related studies, see [45, 46, 47]. Gamma-ray probes were also proposed to detect the effect [48, 49, 50, 51, 52].

For further signatures of quantum vacuum nonlinearity in strong electromagnetic fields of focused high-intensity laser pulses, we refer the reader to the pertinent reviews [16, 53, 54, 55, 56, 57, 58, 59] and references therein.

18.2.2 Searches for ultralight bosonic dark matter

If we assume that galactic dark matter (DM) is dominated by a single particle species, we can find the number density of the DM particles by dividing the observed DM mass density in our galaxy ($\approx 0.4\,\mathrm{GeV/cm^3}$) by the mass of the particle. If the DM particles are fermions, according to the Fermi-Dirac distribution, a fraction of the particles move with respect to the galaxy with the Fermi velocity, even at zero temperature of the gas. This has an interesting and important consequence: if the DM particle's mass is less than $\sim 10\,\mathrm{eV}/c^2$, it has to be a boson. Indeed, the Fermi velocity for a gas of such light particles exceeds $\sim 10^{-3}c$, the escape velocity of the galaxy, which is incompatible with the observation that DM is bound to the galaxy.

A gas of bosonic particle can be equivalently thought of as a field (similar to the photon/electromagnetic-field duality). According the field theory, for instance, the Klein-Gordon equation for (pseudo)scalars, the field oscillates at the Compton frequency of the bosons mc^2/h. Assuming that DM moves in the galaxy with "virialized" velocities with magnitudes $v \approx 10^{-3}c$, the bosons' total energies $\approx mc^2(1 + v^2/2c^2)$ acquire a distribution corresponding to the effective Q-factor of the oscillation of $(v/c)^{-2} \approx 10^6$. This shows that the expected dark-matter field corresponds to a highly monochromatic radiation with a coherence time corresponding to roughly a million oscillation periods and coherence length of roughly a thousand Compton lengths of the boson.

A variety of spin-zero (axions, axion-like particles, dilatons, etc.), spin-one (dark and hidden photons, etc.) and higher-integer-spin particles have emerged as possible DM candidates [60, 61, 62, 63].

The axion is a pseudoscalar (odd intrinsic parity, spin-zero) particle that was hypothesized in the 1970s to explain the experimental fact that strong nuclear interactions do not violate CP invariance, the symmetry with respect to the combined operation of charge conjugation and spatial inversion. At least the violation is at a small enough level that it has evaded detection in the past six decades of dedicated searches. In recent years, axions and a more general class of axion-like particles (ALPs, which are like axions, except they do not solve the "strong-CP

problem") have drawn renewed attention because of the realization that they could be excellent candidates for dark matter, as well as offering a solution to a number of other outstanding problems.

Axion and ALP searches can be classified into categories depending on where these particles are produced. In "light shining through wall" (LSW) experiments (see [64] and references therein; Fig. 356), axions/ALPs are created in the interaction between intense laser light and a static field of a strong magnet (this is known as "Primakoff production"). The light photon and the static field mix to produce the scalar particles that then travel freely through a wall and are detected by converting them back to photons after they cross the wall. While the wall is transparent for axions/ALPs, it blocks the photons completely. In "helioscope" experiments (Fig. 358), the production of axions or ALPs is entrusted to the Sun (Helios), but detection is accomplished in the laboratory, as in LSW experiments. "Haloscope" experiments directly detect the DM from the galactic halo. Note that, in contrast to haloscope experiments, LSW and helioscope experiments are sensitive to "new" particles that may or may not be abundant in the galaxy and be part of DM.

Light shining through walls

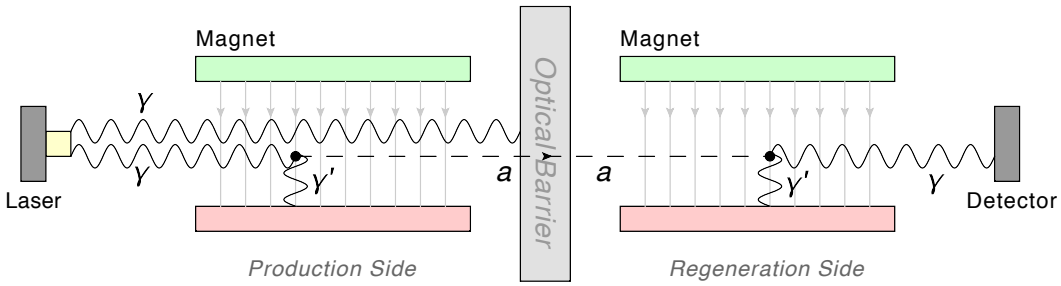


Figure 356: The concept of the light-shining-through-walls experiments. The γ' photons are associated with the quasistatic magnetic field.

The general idea of light-shining-through-a-wall (LSW) experiments is illustrated in Fig. 356. On the production side of the setup, laser photons (γ) transform into axions/ALPs in the presence of the quasistatic field providing photons γ'. The probability of such conversions is given by

$$P_{\gamma\to a} = \frac{1}{4}\left(g_{a\gamma\gamma}BL\right)^2 F(qL), \tag{147}$$

where $g_{a\gamma\gamma}$ is the coupling constant characterizing the strength of the axion/ALP-photon coupling vertex, $q = p_\gamma - p_a$ (p_γ and p_a are, respectively, the photon and axion/ALP momenta), L is the length of the magnetic-field region. The form factor accounting for the difference between the speed of the axion/ALP and the speed of light is

$$F(qL) = \left[\frac{\sin\left(\frac{1}{2qL}\right)}{\frac{1}{2qL}}\right]^2. \tag{148}$$

An identical calculation for the probability of the conversion of the axion/ALP back to a photon in the detection region (and assuming the identical length and magnitude of the magnetic field) leads to the overall probability

$$P_{\gamma\to a\to\gamma} = \frac{1}{16}\left(g_{a\gamma\gamma}BL\right)^4 = 6\cdot 10^{-38}\left(\frac{g_{a\gamma\gamma}}{10^{-10}\ \mathrm{GeV}^{-1}}\frac{B}{1\ \mathrm{T}}\frac{L}{10\ \mathrm{m}}\right)^4, \tag{149}$$

where we assumed sufficiently light axions/ALPs, so that $qL \ll 1$ and $F(qL) \approx 1$.

Various techniques can be used to significantly boost the sensitivity of LSW experiments. Using an optical cavity, one can "trap" the photons used to produce ALPs in the magnetic-field region thus increasing the conversion probability. Less obviously, a resonant optical cavity on the detection side gives a similar enhancement ([65] and references therein). A survey of the LSW experiments and their key parameters is given in Table 30, while the current sensitivity and future prospects are presented in the parameter-exclusion plot in Fig. 357.

Experiment	Reference	Photon energy [eV]	Laser power	Power buildup	Magnetic field strength B[T]
ALPS	[66]	2.33	4 W	$P_p = 300$	5
BRFT	[21]	2.47	3 W	$P_p = 100$	3.7
BMV	[67]	1.17 (14 pulses)	$8 \cdot 10^{21}\gamma$/pulse	-	12.3
GammeV	[68]	2.33 (3600 pulses)	$4 \cdot 10^{17}\gamma$/pulse	-	5
OSQAR	[69]	2.33	18.5W	-	9
ALPS-II	[70]	1.16	30W	$P_p = 5000$ $P_r = 40000$	5
LSW with X-Rays	[71]	50200 90700	10 mW 0.1 mW	-	3
LSW with Pulsed Magnets and Synchrotron X Rays	[72]	9500	46 mW	-	8.3 T and 5.7 T pulsed (duration 1ms)

Table 30: Overview of experimental parameters of previous and future LSW experiments: the photon energy, the initial laser power, the power-buildup in the production and regeneration side (P_p and P_r), as well as the magnetic field strength and length in production and regeneration sides (B_P, B_R, L_P, L_R). For all the cases, $B = B_p = B_r$ and $L = L_p = L_r$.

The choice of magnets for LSW experiments is driven by the requirement that the magnetic field should be transverse in both the production and regeneration regions.[72] As in the case of the vacuum-birefringence experiments (Sec. 18.2.1), it is the length of the field that enters the figure-of-merit (Table 28). Long transverse-field magnets are used in accelerators, and so LSW experiments have benefited from the use of surplus magnets from accelerators such as LHC.

Note also that the "physics case" for the LSW experiments is not limited to axion/ALP searches; there is, indeed, a broad range of "new physics" that can be explored [65].

[72] This follows from the fact that the amplitude of the photon-axion/ALP conversion and the inverse process are both proportional to a rotational invariant $\vec{\epsilon} \cdot \vec{B}$, where $\vec{\epsilon}$ describes the electric field of the light.

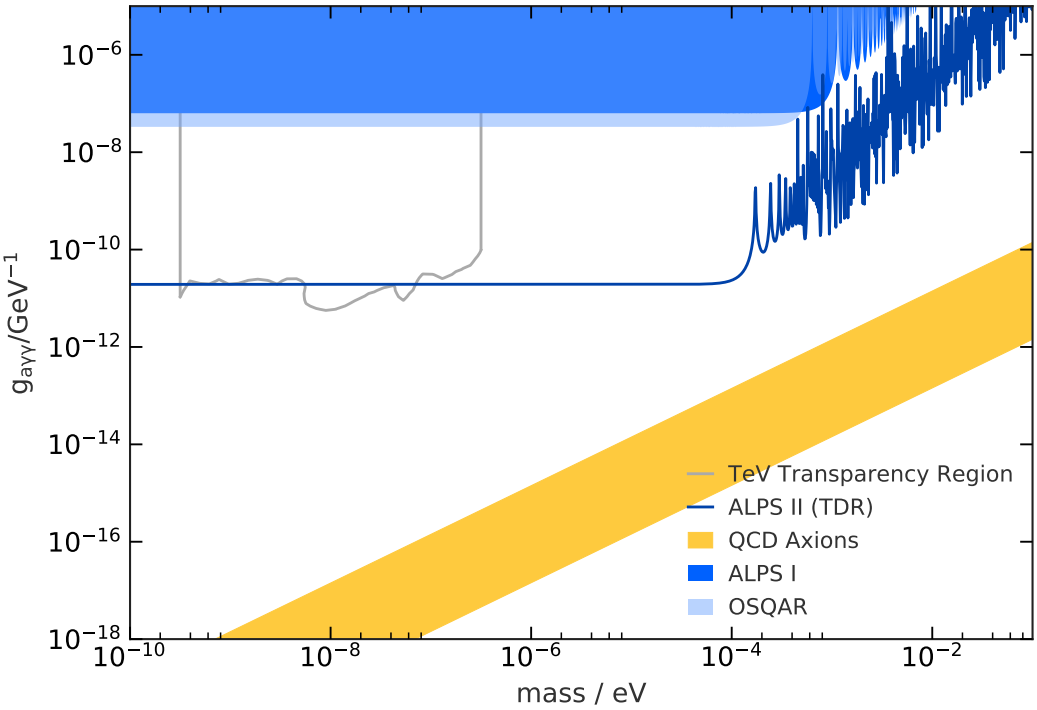


Figure 357: Overview of the current leading limits in LSW experiment (ALPS-I, OSQAR) and the projected sensitivity of the ALPS-II Experiment. The QCD-axion band as well as the region for ALPs explaining TeV transparency is indicated. "TeV transparency" refers to a hint for a possible range of ALP parameters from astrophysical observations that suggest excessive transparency of the universe for TeV gamma rays. The idea is that the gamma rays may "hide" from absorption by converting to ALPs and back in the cosmic magnetic field. TDR stands for technical design report. This overview plot was created with ALPlot (https://alplot.physik.uni-mainz.de).

Helioscopes

The general layout of axion helioscopes is illustrated in Fig. 358. The figure-of-merit for a helioscope reads:

$$\mathrm{FOM}_{\mathrm{Helio}} \;=\; B^2L^2A = B^2VL. \tag{150}$$

This is a simplified equation that assumes uniform transverse magnetic field of magnitude B over the volume V of the detector of length L and area A. If the field varies over the volume, the FOM needs to be evaluated by appropriate volume integration taking into account that the amplitude of the axion-to-photon conversion scales as BdL, hence the L^2 dependence in Eq. (150). It is important to note that this FOM is different from the FOM encountered in other experiments (Table 28).

An example of a helioscope is the CERN Axion Solar Telescope (CAST) [73]. The experiment is constructed around a 9 T LHC test magnet built in the 1990s. It has been very successful over the years, providing new significant limits on the axion-photon coupling.

The currently planned International Axion Observatory (IAXO) [74, 75] is promising to reach an unprecedented level of sensitivity and discovery potential. In terms of the magnet, IAXO moves away from the geometry of accelerator dipoles and towards the detector-inspired barrel geometry, allowing for a better helioscope FOM (Eq. 150). IAXO design draws upon the experience of CAST, and aims at building a new large-scale magnet optimized for axion searches and extensively implementing focusing of the light produced in the solar-axion/ALP conversion and low-background photon detection techniques. The IAXO superconducting magnet will have a toroidal multibore configuration [76]. The current design considers a 22 m long, 5.2 m diameter toroid assembled from eight coils, effectively generating an average (peak) 2.5 (5.1) tesla in the eight bores of 600 mm diameter. This represents a 300 times better B^2L^2A FOM compared to the CAST magnet. The toroid's stored energy will be 500 MJ. The design is inspired by the ATLAS barrel and end-cap toroids [77, 78], the largest superconducting toroids built and presently in operation at CERN. Each of the eight magnet bores will be equipped with a detection line composed of an x-ray-optics telescope and a low-background detector. Beyond the magnet, several improvements are foreseen also in the optics and detector parameters. Figure 359 shows the conceptual design of the overall IAXO infrastructure [74].

As a first step towards IAXO, construction of a scaled-down version of the setup, BabyIAXO (Fig. 358), is contemplated. The BabyIAXO magnet will be only 10 m long and will feature one bore of 60 cm in diameter. It will have a peak field of 4.1 T, an average field of 2.5 T, and a total stored energy of 27 MJ. BabyIAXO will be equipped with only one set of optics and a detector, but of similar dimensions as for the final IAXO systems. BabyIAXO will therefore constitute a representative prototype for the final infrastructure. With the FOM of this demonstrator already exceeding that of CAST by an order of magnitude, it will also provide relevant physics outcomes at an intermediate level between the current best CAST limits and the full IAXO prospects. The design and operational experience with BabyIAXO, in particular with the magnet, is expected to provide relevant feedback for the technical design of the full infrastructure and enable improvements in the ultimate FOM.

A list of helioscopes and their representative parameters are presented in Table 31.

Haloscopes

If axions or ALPs constitute the dark matter of our galactic halo and if their mass corresponds to a Compton frequency in the microwave range, they can be detected in the laboratory via their conversion to photons in a microwave cavity permeated by a magnetic field [89, 90]. The signal power to be detected is

$$W \propto g_{a\gamma\gamma}^2 \frac{\rho_{halo}}{m_a} B^2 V \cdot C \cdot Q, \tag{151}$$

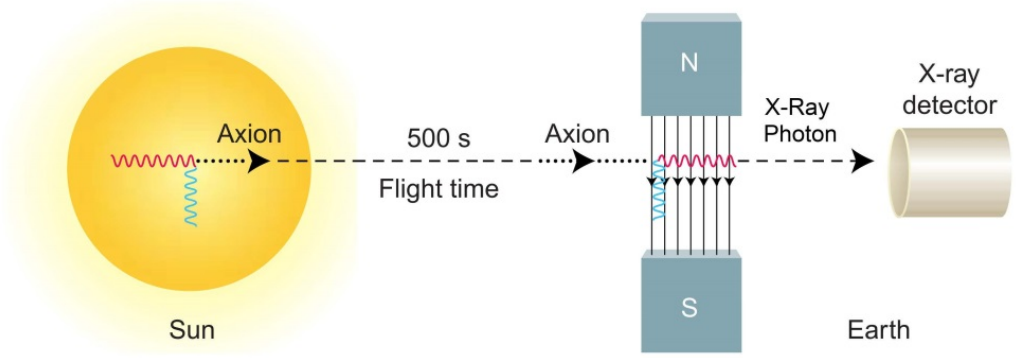


Figure 358: The axion-helioscope concept. Axions are produced in the sun and travel towards the earth. In the presence of a transverse magnetic field in the haloscope (corresponding to the blue photon in the figure), the axions are converted into x-ray photons and detected. The energy spectrum of the x-ray photons corresponds to that of the axions produced in the sun.

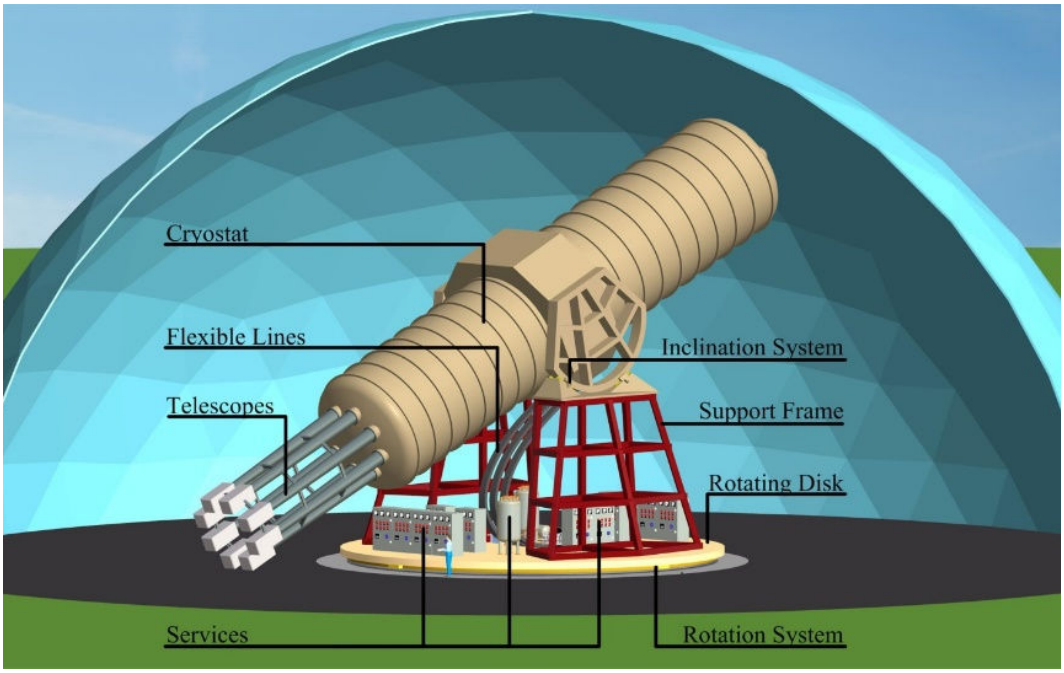


Figure 359: General view of the IAXO design whose key part is a 22 m long eight-coil toroidal magnet enclosed in a 25 m long cryostat. Figure from [74].

where $\rho_{halo} \approx 0.4$ GeV/c^2cm^{-3} is the density of the galactic dark matter, m_a is the axion/ALP mass, V is the volume of the cavity, $C \approx 0.5$ is the cavity-mode form factor and $Q \approx 10^5 - 10^6$ is the cavity quality factor. Resonant conversion occurs when the frequency of the cavity is close to the mass of the ALP, i.e.,

$$h\nu = m_a c^2[1 + O(\beta^2)/2], \tag{152}$$

where $\beta \approx 10^{-3}$ is the galactic virial velocity (which is an estimate of our relative velocity with respect to galactic DM) and h is the Planck constant. The signal is thus highly monochromatic. However, the expected finite bandwidth of the signal limits the required quality factor for the cavity to $Q \approx 10^6$. The search for axions is performed by tuning the cavity frequency in small overlapping steps and the time integration at each scanned frequency is one of the key limiting

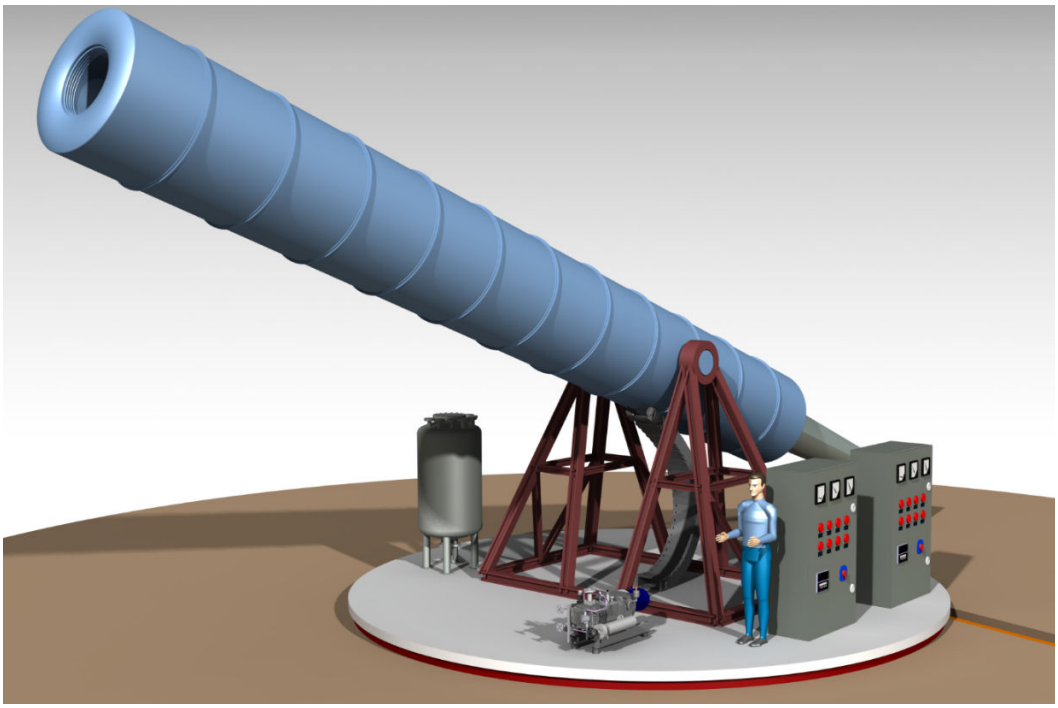


Figure 360: BabyIAXO: the proposed scaled down (there is one only one 10 m long magnet), fully functional initial-demonstration version of IAXO. Figure courtesy CERN/ATLAS Magnet Team.

Experiment	References	Status	B (T)	L (m)	A (cm^2)	Focussing	$g_{a\gamma\gamma}$ (10^{-10} GeV^{-1})
Brookhaven	[79]	past	2.2	1.8	130	no	36
SUMICO	[80, 81]	past	4	2.5	18	no	6
CAST	[82, 83, 84, 85, 86]	ongoing	9	9.3	30	yes	0.66
BabyIAXO	[87]	in design	~2.5	10	2.8×10^3	yes	0.2
IAXO	[88, 74]	in design	~2.5	22	2.3×10^4	yes	0.04

Table 31: Past, present and future helioscopes with the key magnet parameters. The last column represents the sensitivity achieved or expected in terms of the magnitude of the axion/ALP-photon coupling constant. The numbers for the BabyIAXO and IAXO helioscopes correspond to the design parameters considered by the collaboration.

factors. The signal-to-noise ratio can be approximated by the Dicke radiometer equation [91]

$$\mathrm{SNR} = \frac{W}{k_B T_{syst}} \left(\frac{t}{\Delta\nu} \right)^{1/2}, \tag{153}$$

with W being the detection power (in the range of 10^{-23} W), k_B—the Boltzmann constant, $T_{syst} = T + T_N$—the sum of the physical temperature T and the intrinsic amplifier noise temperature T_N, t—integration time, and $\Delta\nu$—the bandwidth. SNR can be maximized by lowering T_{syst}.

Note that the FOM for the magnets used in haloscopes based on axion/ALP-photon conversion is B^2V [see Eq. (151)], which is proportional to the magnetic energy stored in the field.

The "flagship" haloscope experiment is the Axion Dark Matter eXperiment (ADMX) at the University of Washington in Seattle. The experiment is currently running at the level of sensitivity high enough to detect QCD axions (i.e., axions capable to solve the strong-CP problem) with masses around 2.7 μeV. The ADMX magnet was built by Wang NMR (Dublin,

California), delivered in 1993. This superconducting NbTi magnet was designed to operate at 8.5 T in a persistent mode, however, the current switch necessary to put the magnet into the persistent mode upon charging failed on delivery, and ADMX has used this magnet, running with current leads, at 7.6 T, for all searches so far. An upgrade to low-noise superconducting quantum interferometer-device (SQUID) amplifiers necessitated an additional 8 T bucking magnet to create a field-free region about 1 m above the main solenoid (Fig. 361, left). Inside the bucking coil, nested cylindrical layers of mu-metal and niobium define a field-free region that houses a low-noise RF-SQUID amplifier and circulators.

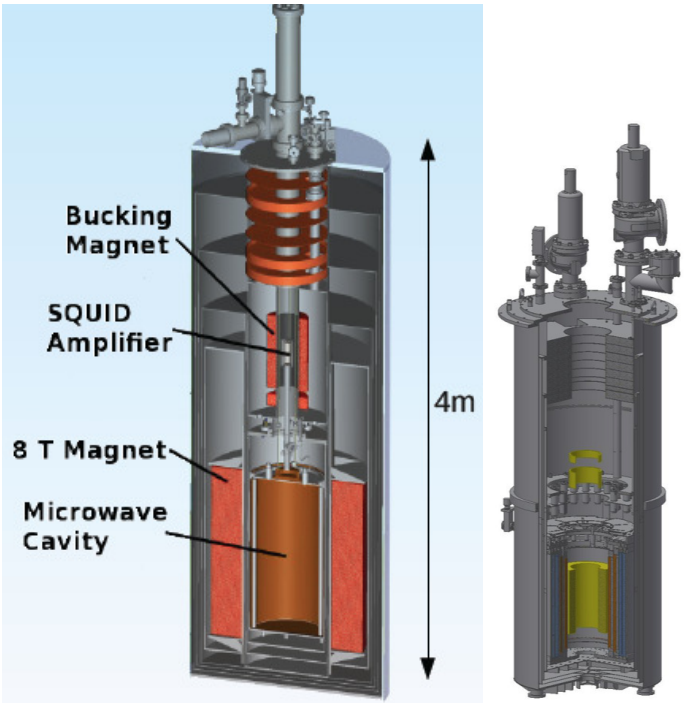


Figure 361: Left: The magnet used by the ADMX experiment. The barrel is 50 cm in diameter and has capacity to accommodate a 220-liter cylindrical microwave cavity. The purpose of the small bucking magnet is to create a zero-field region where the field-sensitive SQUID amplifiers are placed. Right: Design of the new 24 T magnet for ADMX. The outer diameter of the magnet is about 0.9 m. The coils are color-coded according to: yellow-HTS; red-Nb_3Sn; blue NbTi. Figure courtesy NHMFL.

A design for the next-generation ADMX magnet is currently in the works. It envisions a 24 T static field and a bore diameter of 16 cm (smaller than the 50 cm in the current magnet). Indeed, this will accommodate a smaller microwave cavity providing access to higher frequencies and, correspondingly, higher axion/ALP masses. The power in the photons produced from conversion of dark-matter axions in the presence of magnetic field is proportional to B^2V [Eq. (151)]. Conventional cavity resonators have geometry dictated by the boundary conditions necessary to allow the resonances of the transverse-magnetic (TM) modes of the cavity in the magnetic field to correspond to the axion mass to be probed via $h\nu = m_a c^2$. Here ν is the frequency of the cavity TM mode used for the search, usually the TM_{010} mode. Therefore the only degree of freedom that can be altered at a given axion mass is the strength of the magnetic field. The current ADMX magnet has a 7.6 T central field, so a higher-field magnet with the proposed design field of 24 T would raise the signal-to-noise ratio by a factor of 12.6. The NHMFL design (Fig. 361, right) utilizes three nested sets of windings, the outer windings being niobium-titanium, the intermediate windings being niobium-tin, and the innermost windings being YBCO tape. The choice of winding materials is guided by the quench fields in the

different superconductors; closer to the bore of the magnet the windings must be quench–free at higher fields. As of February of 2018, there is no proposal submitted to any funding agency for a magnet upgrade; however such a proposal has a high priority once the first results from phase two of ADMX have been published. Publication of these results is expected in early spring of 2018.

A closely related experiment to ADMX is the Haloscope At Yale Sensitive To Axion Cold dark matter (HAYSTAC) [92, 93] that was previously known as ADMX-HF and X3. HAYSTAC uses a 9.4 T DC magnet made by Cryomagnetics that is cooled with a pulse-tube refrigerator. Both the magnet and the main bore (17.5 cm in diameter, 50 cm long) are "dry" in the sense that the only helium in the system is the ^{3}He/^{4}He mixture in the dilution refrigerator. A small cavity gives sensitivity to high-mass axions (in a range complementary to that of ADMX), currently already reaching the sensitivity necessary to detect a QCD axion. A disadvantage of a dry magnet system is that such a magnet quenches within minutes in the event of a power loss. The HAYSTAC collaboration lived through such an unfortunate experience; however, despite of the initial pessimistic assessment of the "destroyed" magnet, miraculously, it was possible to fully repair the system within two months.

Experiment	References	Status	B (T)	V (m^3)	B^2V (T^{2}m^3)	Multiple Cavities
						m
ADMX	[94]	present	7.4			no
		future	24	0.008	4.8	no
HAYSTAC	[93]	present	9.4	0.0015	0.12	no
ORGAN	[95]	present	7	0.003	0.16	
		future	14	0.0014	0.27	
CAPP	[96]	commission	12	0.033	4	no
		present	18	0.001	0.3	no
		comission	25	0.0025	1.4	no
						yes
Grenoble	[97]	proposal	9/17.5/27/ 40/43	0.49/0.05/0.0029/ $2.9{\cdot}10^{-5}$/$0.45{\cdot}10^{-5}$	13-0.01	yes/no

Table 32: Present and future haloscopes based on the Primakoff production of microwave photons by dark-matter axions/ALPs. In this case, V refers to the volume within microwave cavities placed into magnetic field. After CAPP runs the 12 T and the 25 T magnets, they plan to combine them and reach 37 T within the volume of the 25 T magnet, with a significant boost in the sensitivity of the axion/ALP search.

A new haloscope has been proposed by a French/Korean collaboration that utilizes the Grenoble hybrid magnet described in Sec. 18.4.1 that boosts the sensitivity by both optimizing the magnet FOM and operating at low (mK-range) temperatures [97]. Compared to other haloscopes such as ADMX, the superconducting outsert magnet alone offers an unprecedented value of $B^2V \approx 40\,\mathrm{T^2m^3}$ to search for axion/ALPs in the mass range of 1-3 μeV. A further increase of B^2V up to $75\,\mathrm{T^2m^3}$ is possible but would imply major changes in the structure of the Grenoble hybrid magnet as well as significant investments, which might be considered at a later stage. On the other hand, adding resistive magnets to operate at higher fields and employing small cavities will allow further unique possibilities for this new haloscope to probe axions/ALPs of larger masses, so overall the range of axion/ALP masses from μeV to hundreds of μeV can be covered with FOM of B^2V = 13-0.01 T^{2}m^3 (Table 32).

It should be noted that magnets of a variety of geometries can be considered for use for axion/ALP searches [98] and a pilot haloscope search with a toroidal magnet was recently completed [99]. The significance of this is that large toroidal magnets that are built for plasma-

physics research (Tokamaks) can, in principle, be used for axion searches [100]. Indeed, Tokamaks like the International Thermonuclear Experimental Reactor (ITER) can have very large $B^2V \gtrsim 3 \cdot 10^4\,\mathrm{T}^2\mathrm{m}^3$ (c.f. Table 32).

A summary of the Primakoff-haloscope results and projections is given in Table 32 and Fig. 362. Note that assuming that dark matter is dominated by a single species of axion/ALP, haloscopes have higher sensitivity to the coupling constant $g_{a\gamma\gamma}$ compared to other experiments searching for axions/ALPs in the same mass ranges.

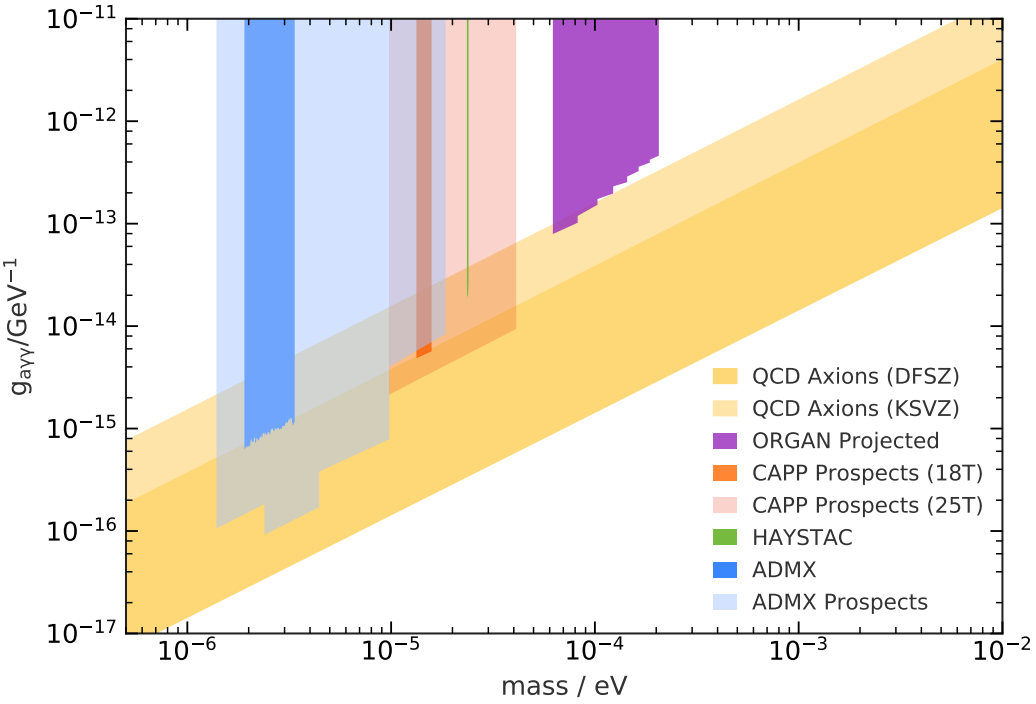


Figure 362: Current and projected limits on the axion-photon coupling constants from Primakoff-haloscope experiments under the assumption that dark matter is dominated by a single-species of axions or ALPs. This overview plot was created with ALPlot (https://alplot.physik.uni-mainz.de).

Several new experiments and proposals address the possibility that the DM particles could be lighter than μeV. In principle, the entire range down to $\approx 10^{-22}\,$eV is "fair game," limited from below by the requirement that the de Broglie wavelength of the particle not exceed the size of the galaxy.

The DM Radio experiment [101] will search for axion/ALP and hidden photon dark matter over a wide range of sub-μeV masses. The concept of DM Radio is simple: like ADMX, it is a sensitive electromagnetic resonator (a superconducting lumped-element circuit), shielded from external noise. The dark matter passes through the shield and excites the resonator when the resonant frequency is tuned to the dark-matter mass. DM Radio will cover many orders of magnitude in mass and coupling for the hidden photon in a relatively small and fast experiment. This broad frequency coverage is possible since DM Radio relies on a lumped element (LC) circuit. To search for axions, DM Radio needs a magnet [Fig. 363]. In the presence of a magnetic field, DM radio will search for axions in a frequency range complementary to ADMX and CASPEr (see below). The DM Radio group would welcome collaboration with a facility that can provide a large magnet, with minimal fringing fields to prevent driving superconductors normal, and compatible with a high-quality-factor resonator, so that the resonant axion signal can be optimally detected.

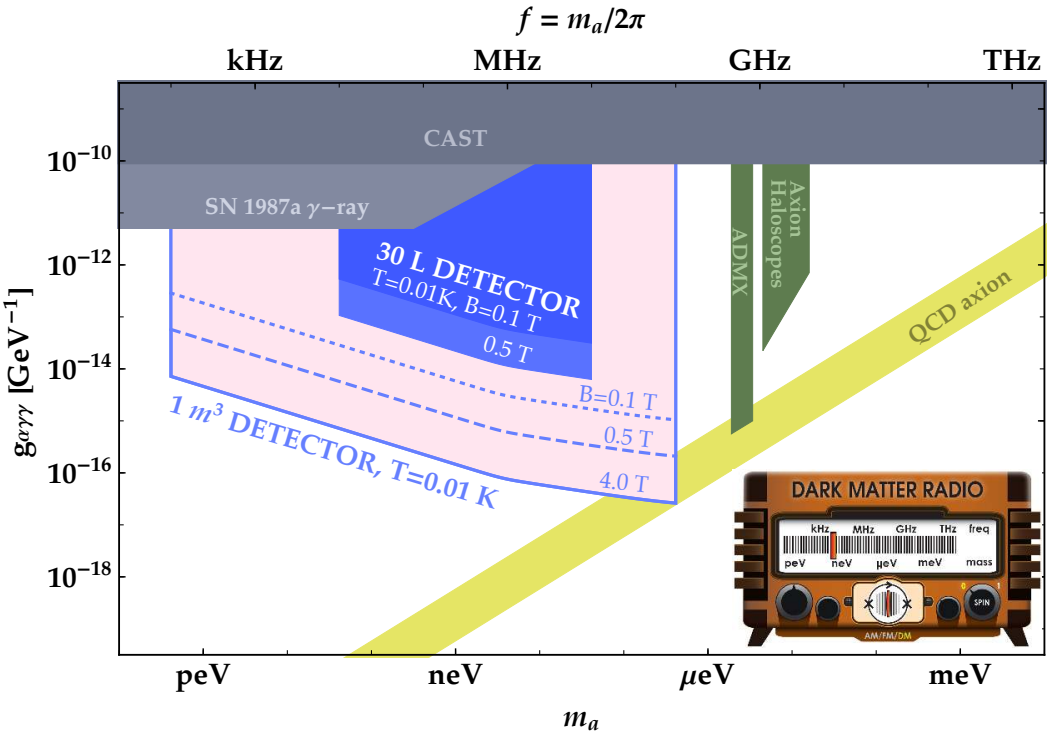


Figure 363: Projected sensitivity of various stages of the Dark Matter radio (DM radio) experiment to axion-photon coupling. The inset at the lower right shows an emblem of the experiment accurately capturing some of the essential features. Figure courtesy DM radio collaboration.

A similar "lumped-circuit" approach to dark-matter axion/ALP search is pursued by the MIT-based ABRACADABRA (A Broadband/Resonant Approach to Cosmic Axion Detection with an Amplifying B-field Ring Apparatus) experiment [102]. The ABRACADABRA collaboration envisions a three-stage program where the stages are defined by the magnet. The first stage ABRACADABRA-10 cm is a demonstrator experiment which will begin to cut into as yet unexplored ALP parameter space. It is currently preparing for its first data run with a 1 T toroidal magnet with inner radius of 3 cm, maximum outer radius of 6 cm, and maximum height of 12 cm. ABRACADABRA-75 cm scales the magnet to an inner radius of 25 cm, a height of 75 cm, and magnetic field of 5 T. This experiment would have some sensitivity to the QCD axion depending on the readout approach. The ultimate experiment would be a large magnet with specifications on the order of a 10 T field with an inner radius of 4 m, an outer radius of 8 m, and a height of 12 m. This may be prohibitively expensive to build, so the collaboration is exploring alternative readout strategies and geometries that may allow similar sensitivities with a more affordable magnet.

If DM consists of axions/ALPs, there exist additional possibilities to search for them based on the fact that, in addition to axion/ALP coupling to electromagnetic field (with coupling constant $g_{a\gamma\gamma}$), there are two more nongravitational coupling of axions/ALPs that are predicted by theory (see, for example, [103]). One of these, corresponds to axion/ALPs interaction with two gluons instead of two photons. As a consequence, a nucleon and nonzero-spin nuclei acquire time-dependent electric-dipole moments (EDMs) whose time evolution follows that of the axion/ALP dark-matter field, oscillating at the axion/ALP Compton frequency and having a typical "Q-factor" of 10^6. It is important to note that EDMs violate both parity (P) and time-reversal invariance (T) on the account of the corresponding intrinsic properties of the axion/ALPs. The other possible axion/ALP interaction with normal matter is the so-called derivative interaction, where the gradient of the axion/ALP field (colloquially known as "axion

wind") acts as a pseudo-magnetic field causing nuclear spin precession. In contrast to the usual magnetic field, the pseudo-magnetic field can penetrate magnetic shielding. A search for both these interactions is conducted by the Cosmic Axion Spin-Precession experiment (CASPEr) collaboration [104, 105], with separate setups. The search for the axion/ALP-induced oscillating EDM (CASPEr-Electric) is conducted at the Boston University, while the setup sensitive to the derivative coupling (CASPEr-Wind) is at the Helmholtz Institute in Mainz. In both CASPEr-Electric and CASPEr-Wind a sample of polarized nuclei is immersed in a magnetic field whose magnitude is scanned. When the Larmor precession frequency becomes resonant with the axion/ALP mass, the DM field induces resonant spin-flips, and the resulting time-dependent transverse magnetization of the sample is detected with a SQUID magnetometer, so the experiments are, in effect, highly sensitive continuous-wave NMR spectrometers. Because CASPEr-Electric requires an electric field present in the sample and CASPEr-Wind does not, the two setups operate with different samples: a solid ferroelectric material containing ^{207}Pb in the former case and liquid ^{129}Xe in the latter.

The requirements for magnets to be used in CASPEr are strict: they should be broadly scannable and have temporal instabilities and field inhomogeneity at better than ppm levels over the entire field range. Phase-one CASPEr-Electric is using a commercial 9 T Cryomagnetics superconducting magnet, specified to have 0.1% homogeneity over a 1 cm^3 volume. The rated current at 9 T is 64 A, and the charging rate is 0.1 A/s, therefore the magnet can be charged or discharged in ten minutes. The charging rate is an important parameter since the CASPEr-Electric experimental protocol includes sample pre-polarization in a large magnetic field, before a sweep to the target magnetic field range for the axion search. If magnetic shielding is necessary to prevent external magnetic noise, an open question is how the magnetic field profile inside the magnet will be affected by the presence of a shield. The planned magnet for CASPEr-Wind has a spatial inhomogeneity over a 10 mm diameter spherical volume corresponding to < 1 ppm NMR linewidth during a 0.2 T sweep. The main field will be sweepable from -14.1 to 14.1 T and the room-temperature bore will be 89 mm in diameter. The magnet will be equipped with a set of 4 to 8 cryoshims. Particular attention needs to be given to minimizing vibrations.

An experiment analogous to CASPEr-Wind but searching for axion/ALP coupling to electrons rather then nuclei called QUAX (QUaerere AXion) is being developed at the University of Padova [106]. Its principle is detection of electron spin-flips in a magnetized solid in a magnetic field induced by the pseudo-magnetic DM field.

18.2.3 Neutrino-mass measurements

Neutrinos were assumed massless in the framework of the Standard Model, however, neutrino-oscillation experiments have shown that at least some of the masses associated with the three neutrino flavors are nonzero; moreover, the neutrino mass eigenstates are superpositions of flavor eigenstates (and vice versa). Apart from neutrino oscillations, one can glean information on neutrino masses from the analysis of cosmological data, from the neutrinoless double-beta-decay experiments, and from accurate measurements of the electron spectra in beta-decays.

Most commonly used for mass measurements is beta-decay of tritium resulting in a ^{3}He nucleus, an electron, and an electron antineutrino with a release of 18.6 keV–a small energy on the nuclear scale. Because the decay results in three particles in the final state, the spectrum of the decay electrons is continuous. The relative effect of the (anti)neutrino mass is most pronounced at the "endpoint," where the electron carries maximal kinetic energy, close to the overall energy released in the decay. A major difficulty of the measurements close to the endpoint is that there are few decays resulting in highest-energy electrons. For example, only $2 \cdot 10^{-13}$ of the tritium beta-decays result in electrons within 1 eV from the endpoint.

The state-of-the-art beta-spectrometer Karlsruhe Trititum Neutrino Experiment (KATRIN)

[107] employs a building-size vacuum chamber and aims to achieve neutrino-mass sensitivity of better than 200 meV in the next years, limited by the ability to control systematic effects. Given the fundamental importance of neutrino-mass determinations, it is highly desirable to develop alternative, perhaps, smaller-scale, experiments that are not affected by the same systematics as KATRIN.

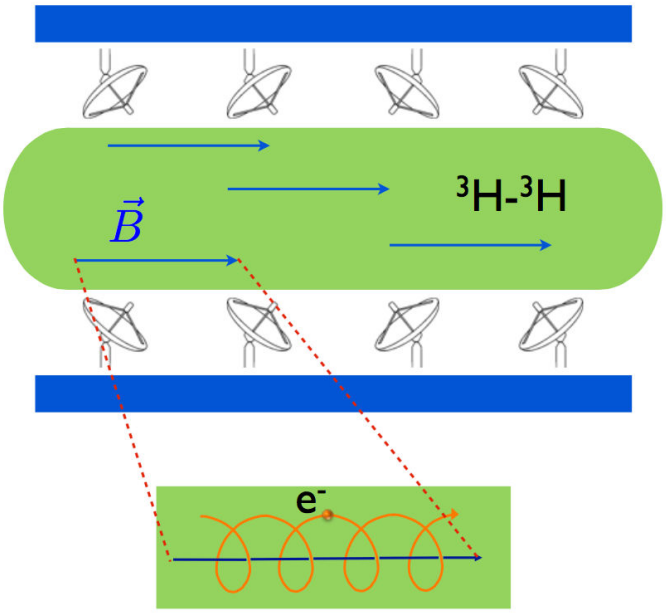


Figure 364: The concept of the neutrino-mass measurement via determination of the cyclotron frequency of electrons emitted in beta-decay of tritium [108].

One such alternative approach [108] involves applying a strong magnetic field around the source and measuring the energy-dependent cyclotron frequency

$$f = \frac{1}{2\pi} \cdot \frac{eBc^2}{m_e c^2 + E} \tag{154}$$

of the electrons produced in beta-decay (Fig. 364), taking advantage of the fact that frequencies tend to be the physical quantities most amenable to precision measurements. For example $\delta E = 1$ eV of kinetic energy near the endpoint translates into a fractional cyclotron-frequency offset of $\approx \delta E/(m_e c^2) \approx 2$ ppm, which can be readily detected if the electron is observed long enough.

In the "Project 8" experiment [109], a gas volume is filled with tritium, a magnetic field is applied, and antennas are used to detect the radiation emitted by electrons undergoing the cyclotron motion. With $B = 1$ T, the cyclotron frequency is close to 28 GHz (the K_a microwave band). In order to precisely measure the applied magnetic field and calibrate the apparatus, Project 8 uses the internal-conversion electrons from a ^{83m}Kr source, which emits electrons of several discrete energies, one of them close to the energy of the electrons at the tritium endpoint.

When it comes to detection of the radiation from single 18 keV electrons undergoing cyclotron motion in a 1 T field, the challenge lies in the low power of the emitted radiation, on the order of 1 fW in total. Nevertheless, with the use of advanced microwave techniques, this feat was accomplished by the Project 8 collaboration [110] in a prototype experimental setup at the University of Washington.

Sufficient observation time to establish the energy resolution is ensured by a shallow (few mT) magnetic trap in which the electrons are localized until they scatter off of gas atoms.

Energy resolutions at the eV level have been achieved this way.

With the proof-of-principle of this measurement technology accomplished, two conditions must be met in a future setup for successful neutrino-mass measurement: a) accumulation of sufficient event statistics, which directly translates into a stable fiducial population of 10^{18} tritium atoms in the apparatus and b) excellent control of systematic effects, in particular those that affect the energy resolution. One of the dominant factors is that the final-state spectrum of molecular tritium is broadened due to the rovibrational excitations of the daughter molecule, providing a sensitivity limit of 100 meV for molecular tritium. To achieve its design sensitivity of 40 meV, Project 8 will therefore have to employ atomic, rather than molecular, tritium.

However, the presence of atomic tritium imposes an additional requirement. Contact with any surface catalyzes recombination of atomic tritium. To prevent recombination, tritium atoms can as well be trapped magnetically through their non-zero nuclear magnetic moment.

The task for the Project 8 magnet system is therefore threefold:

- provide a highly uniform ($\Delta B/B$ at a ppm level) 1 T solenoidal field over a $\approx 10\,\mathrm{m}^3$ fiducial volume in which the electrons can emit cyclotron radiation,
- provide shallow mT traps in which the decay electrons are localized during observation,
- provide a conservative potential with a closed $|B| \approx 2\,\mathrm{T}$ contour inside all surfaces to prevent recombination of atomic tritium. The magnet system must superimpose this trapping field on the 1 T background field without disturbing its ppm uniformity in the fiducial region.

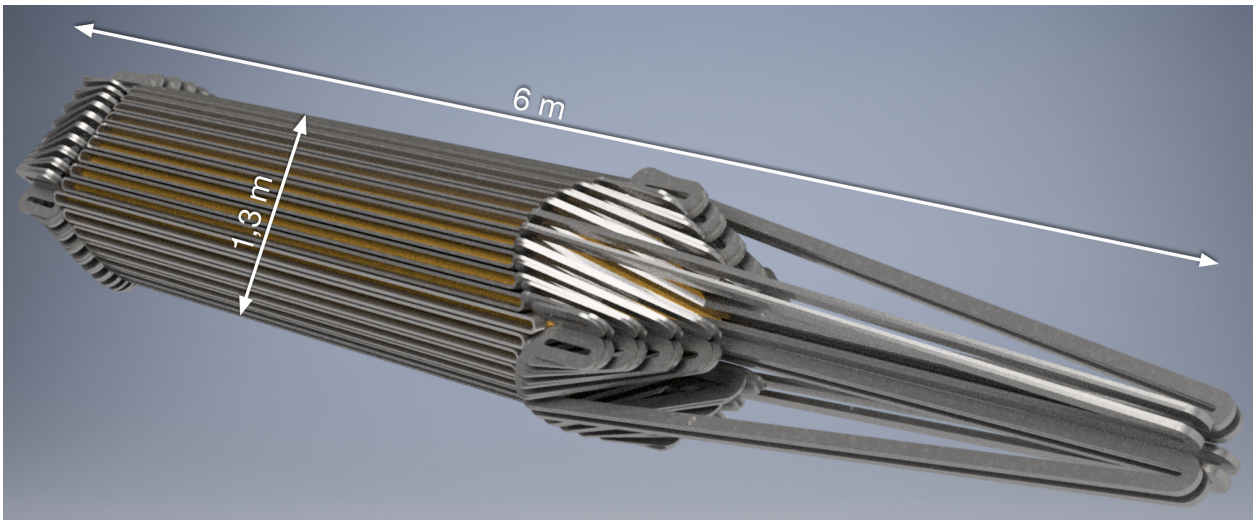


Figure 365: Design sketch of a possible Project 8 atomic-tritium trap consisting of 60 bars of 4×8 cm cross-section. The fiducial volume in which the field homogeneity is good enough for cyclotron-radiation-emission spectroscopy (CRES) is about 25% of the physical volume.

While the design of the uniform background field and electron traps is straightforward, design of the atom trap provides several challenges. A cylindrical multipole magnet with discrete axial conductors carrying opposing currents creates a magnetic field that increases radially from its central axis. Such a Ioffe-Pritchard trap has been used to confine antihydrogen by the ALPHA collaboration. The order of the multipole in a Ioffe trap determines the radial dependence of the field, where higher orders increase the steepness of the radial increase and, therefore, the fiducial volume in which the uniformity requirement is met. However, employing multipoles of too high order, the 2 T contour would not project inward far enough to hold

atoms away from the physical surface. Additional challenges arise from closing off the cylinder in the axial direction and because the trap must be filled with atomic tritium from the outside. Figure 365 shows a first design sketch of a possible atomic-trap configuration, in which the trap is loaded through a protruding quadrupole section, which opens up into the actual trap volume ensuring a continuous 2 T contour inside all surfaces for a current density of ≈ 200 A/mm^2 per conductor.

While the Project 8 requirements call for a complex, superconducting Ioffe-Pritchard trap to store 10^{18} tritium atoms in a large fiducial volume with a 1 T solenoidal field that is uniform at the ppm level, in the center of a 2 T deep magnetic minimum, a first engineering design for a ≈ 1 m^3 atom-trapping demonstrator magnet is being currently undertaken by the collaboration.

18.3 Ideas for future experiments

Atoms and molecules have proven to be very useful for fundamental-physics research [111]. In this section, we discuss several cases where strong magnetic fields can potentially enable new breakthroughs in this area.

18.3.1 Parallel spectroscopy of complex systems at high magnetic fields: Spectroscopy 2.0

Specific experiments require atoms or molecules with particular sets of properties. For example, atomic parity-violation effects are enhanced for heavy atoms and when opposite-nominal-parity atomic states are close in energy or, as a limiting case, are degenerate. The enhancement of the effect of the interaction under study due to the proximity of energy levels mixed by the interaction is not specific to parity violation; in perturbation theory, the mixing is inversely proportional to the energy denominator. While there are known "accidental" degeneracies, the levels do not always possess a complete set of desired properties. For example, dysprosium atoms have nearly degenerate metastable opposite-parity states of the same total electronic angular momentum that can be tuned to full degeneracy by applying magnetic fields on the order of tens of millitesla. While these states may appear ideal for parity-violation experiments [112], the relevant matrix elements were found to be strongly suppressed [113].

In designing fundamental-physics experiments, it is helpful to have a choice of systems with a range of properties in order to optimize the sensitivity to particular interactions of interest. In the case of atoms, one can turn to atomic databases [114], however, the spectra for atoms are not as complete as one would have hoped in the 21st century! This is particularly true for the rare-earth and actinide atoms and for highly-excited states. But "new" relatively low-lying levels are being found even in lighter systems such as barium [115].

The reasons for the considerable gaps in atomic tables is that dense spectra are difficult to analyze both from the experimental and theoretical points of view.

To address these issues, a novel approach to spectroscopy of dense atomic and molecular spectra is proposed [116]. It combines three developments in spectroscopy: 1. The ability to record the spectra in a massively parallel manner using femtosecond frequency-comb techniques; 2. Availability of strong (up to over 100 T) tunable magnetic fields enabling a new dimension in spectroscopy; and finally, 3. Developments in many-body atomic and molecular theory methods, which, combined with the power of modern computers, allows for unambiguous interpretation of the multidimensional spectra with massive information content.

Here, we outline a vision for a joint effort between experimentalists and theorists to develop a modern atomic spectroscopy program, with a possibility of extension to molecular systems in the future. This program will improve the knowledge base of atomic spectra, using state-of-the-art tools developed in recent years by experimentalists and atomic theorists. The program will cover studies of atomic spectra in strong magnetic fields, providing a new, largely unexplored dimension for atomic spectroscopy.

The advantages of extending our spectroscopy program to atomic systems subject to large magnetic fields are numerous. The information provided would allow for nearly simultaneous measurement of large numbers of transition probabilities, magnetic g-factors of states, and the magnetic and electric dipole interaction strengths between states brought close to or through level crossings/anti-crossings. These level crossings may appear frequently at relatively modest fields (1-10 T) in heavy elements with partially filled d- and f-electron shells. Such fields can be achieved in relatively simple laboratory experiments.

Comprehensive study of the spectra of rare-earth elements (REE) in magnetic fields is additionally relevant for modern astrophysics. For example, highly anomalous abundances of

REE are observed in stars with strong magnetic field ($\sim$ 0.1 − 1 T) [117, 118]. At present, neither the origin of the magnetic field nor the abundance anomalies are well understood. Since information about conditions in cosmic objects is obtained from astronomical spectra, the reliability of derived physical parameters depends directly on the accuracy of atomic data such as transition probabilities and oscillator strengths. Although some data on the REE are available [119], they are by far insufficient to meet all the scientific needs. This makes new laboratory measurements of current importance.

Sufficiently complex elements have spectra with chaotic properties [120, 121, 122]. The analysis techniques necessary for understanding these spectra bridge many disciplines, ranging from collisions in ultracold atoms [123] to dielectronic recombination (a process where an electron is captured by an atom and this capture is accompanied by excitation of a initially bound electron) [124] and the spectra of nuclei. Chaotic systems are particularly sensitive to small perturbations and can be used to study fundamental interactions and search for new physics beyond the standard model. For example, statistics of complex spectra may be used to constrain exotic T-odd, P-even interactions that indicate physics beyond the standard model [125]. A long-term goal is to continue the spectroscopy program at a high-magnetic-field facility, where magnetic fields up to 100 T or higher may be achieved. This would allow for the study of atomic spectra in environments found only in the vicinity of neutron stars, white dwarfs, or black holes [126, 127].

When the density of states reaches some critical level, the system becomes chaotic. Strong magnetic fields mix states with different total angular momenta and this effectively increases the level density. For some systems this may lead to a transition to chaotic behavior. Such "controlled chaos" is interesting both from the theoretical and practical points of view. It can improve our understanding of chaos in complex quantum systems, while providing new tools to study weak interactions in atoms.

How can magnetic field help with identification of electronic states? To start with, atomic g-factors have been traditionally used for level identification as they depend on the angular-momentum quantum numbers, for example, L, S and J for the LS-coupling scheme. In dense spectra, there are level (anti)crossings as a function of magnetic field as illustrated in Fig. 366 for the case of singly ionized Ni. The magnetic-dipole (M1) operator only mixes states with the same configuration, and so the presence of anticrossings provides valuable information for level identification. This information is complementary to that from spectroscopy without magnetic field that provides information on electric-dipole (E1) couplings which are nonzero only with different, opposite-parity configurations. Furthermore, magnetic field mixes levels with different total angular momenta J thus removing the usual E1-transition selection rule $\Delta J = 0, \pm 1$. Thus many transitions appear in the spectra that are forbidden in the absence of magnetic field.

In a complex system of energy eigenstates, it is often useful to discuss statistical properties of state mixing by small perturbations rather than mixing of individual pairs of states. In fact, application of magnetic field can have a profound effect on statistical properties of such mixing leading to statistical enhancement of all small perturbations. Studying such phenomena can improve our understanding of chaos in complex quantum systems, while providing new tools to study weak interactions in atoms. Figure 367 shows a histogram for the calculated mixing of neighboring states of opposite parity for singly ionized nickel (Ni II) by a P-odd interaction [128]. We see that mixing grows with the magnetic field, but remains small. We conclude that for Ni II in the range of energies considered in this example (within approximately 30 000 cm^{-1} from the ground state) there is no transition to chaos even in magnetic fields of ≈ 170 T. In Fig. 367 we see that the number of mixed states grows while the typical size of the mixing remains the same. It is predicted that at higher magnetic fields, there should be more of the larger mixing coefficients, however, it is not known yet if this will look like a phase transition

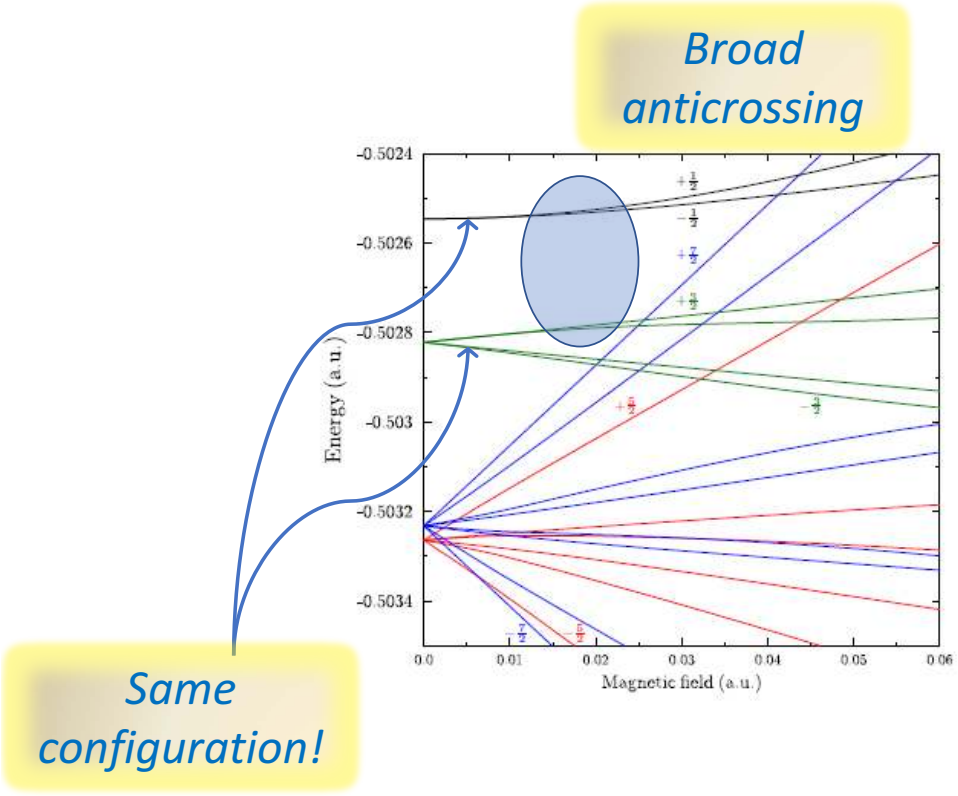


Figure 366: Selected energy levels of Ni^+ (Ni II) illustrating level-anticrossing behavior. Energies of electronic states and magnetic fields are given in atomic units (a.u.); an atomic unit of magnetic field is $c_1 e^2/a_0 \approx 1.7 \cdot 10^3$ T; e is the elementary charge, a_0 is the Bohr radius, and c_1 is a constant appropriate for the units in which e and a_0 are expressed.

or a gradual process. The average level spacing in Ni II is a few hundred cm^{-1}. In lanthanides and actinides, however, typical level density is much higher and transition to chaos may be possible. This can lead to much stronger enhancement of the mixing and high sensitivity to exotic fundamental interactions.

The success of the proposed spectroscopy program relies on one's ability to acquire spectra of complex system immersed in a strong (possibly, time-dependent) magnetic field in a massively parallel fashion. Advances in coherent light sources such as frequency combs provide an opportunity for basing the new broadband spectroscopy program on detection of absorption of light rather than fluorescence. In absorption, the initial electronic states of transitions can be known and controlled with high certainty. This simplifies the step of converting from an element's spectra to its electronic level structure and assigning configurations. The principle of broadband absorption spectroscopy with frequency combs was demonstrated by several groups, including the spectroscopy of molecular iodine [129], spectroscopy of HfF^+ [130], and trace-gas detection [131]. In these examples absorption spectra were parallel-acquired for a wavelength range of several nanometers, with precision on the order of several MHz. These publications demonstrate the feasibility of high resolution, broadband spectroscopy with frequency combs, albeit with limited wavelength coverage. This wavelength range was limited either by the frequency comb itself or the wavelength coverage of the dispersive imaging systems. Extending the spectral range to hundreds of nanometers will be highly desirable for the proposed program. For the light source, octave spanning frequency combs are a well-established technology, and can be acquired as complete commercial systems.

Finally, once the spectra have been obtained as a function of magnetic field, novel theoretical

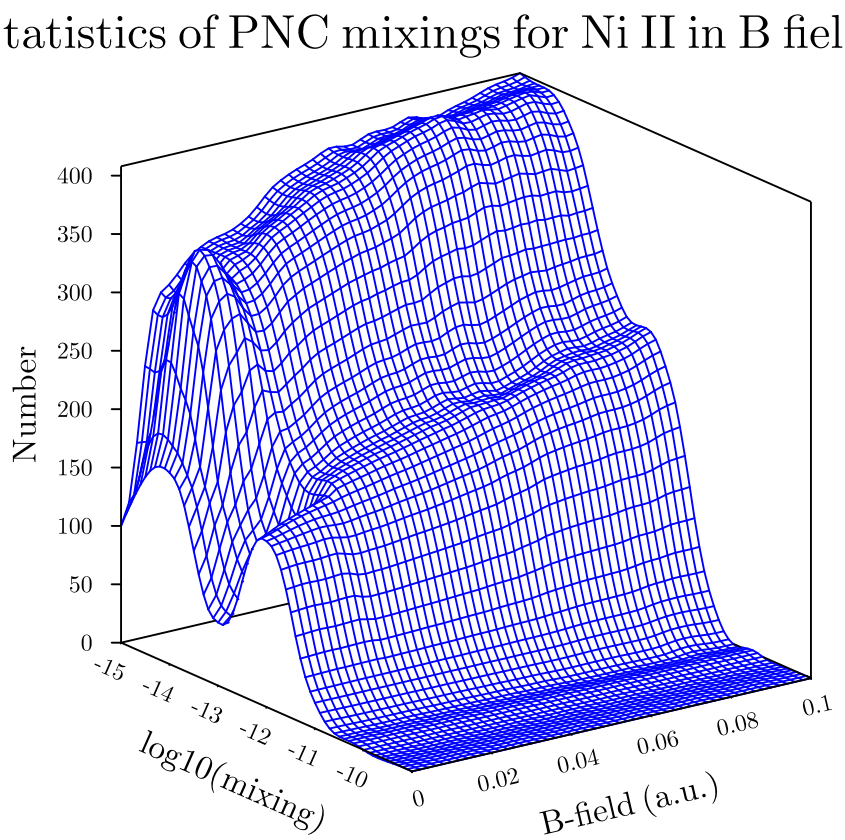


Figure 367: Calculations showing the statistics of opposite parity state mixing as a function of magnetic field. Atomic energy levels of both parities are calculated at a given magnetic field. Then, weak-interaction-induced mixings of neighboring states of opposite parity is evaluated. Taking logarithm of the absolute value of the mixing coefficients, one determines how many of the mixing values lie in a certain range, from which the plot is produced. We see that the number of mixing coefficients with magnitudes in the range between 10^{-15} and 10^{-12} grows with B, but there is no growth for larger mixings above 10^{-11}.

techniques will be applied to analyze these data and extract the target information. An example of such theoretical tool is the recently developed software package designed to model atomic spectra with a combination of configuration-interaction and many-body perturbation-theory analysis (CI+MBPT) [132].

18.3.2 Positron capture by atoms

According to calculations [133], about half of the atoms in the periodic table have bound states with the positron. However, such bound states have not been observed so far. Experimental observation of such bound systems, apart from being a test of atomic theory, would mark a creation of a fundamentally new class of atomic systems of particular interest for fundamental physics because they involve an antiparticle. For example, spectroscopy of atoms containing an antiparticle can be used in searches for exotic spin-dependent interactions among the atomic constituents [134].

The authors of Ref. [133] suggested a possibility to capture positrons to a shallow bound level using a pulse of a strong magnetic field. The idea here is similar to the use of Feshbach resonances for molecular formation out of cold atoms, where the system can be brought into a bound state by sweeping the magnetic field. The same techniques can be used to capture

electrons to an atom forming a negative ion.

18.3.3 Search for dark matter using atomic-physics methods

In Sec. 18.2.2, we discussed the CASPEr and QUAX experiments, where a spin system is brought into resonance with the Compton frequency of the DM field via tuning the energy splitting between the spin states by varying the leading magnetic field. The availability of high-field magnets may expand the range of the DM-particle masses that may be explored. Various other possibilities have also been discussed in the literature.

Atomic and molecular transition induced by axion/ALPs were considered by Zioutas and Semertzidis [135]. Axions/ALPs can induce 1^+ (i.e., $M1$) and 0^- transitions, where the number indicates the tensor rank of the transition operator and +/- indicate whether there is (-) or there is no (+) change of parity in the transition. Sikivie [136] considered atomic $M1$ transitions induced by the derivative coupling of axions/ALPs to nucleons and electrons. The transition probability is proportional to the square of the axion-fermion coupling. It was proposed to detect the axion/ALP-induced transitions by resonantly photoionizing excited atoms. Since the method was discussed in the context of the detection of the galactic-halo axion/ALPs, it is assumed that the atomic energy interval of the axion/ALP-driven transition is resonant with the frequency corresponding to the axion/ALP mass.

Following Sikivie's approach, Santamaria et al. [137] considered an experimental realization using molecular oxygen at sub-kelvin temperatures, where the transition frequency is scanned by applying a strong magnetic field and the detection is done via resonant multiphoton-ionization (REMPI) spectroscopy. Another system considered in this context is optically pumped Er^{3+}-doped yttrium-aluminum garnet (YAG) crystals [138].

We note that related to the discussion of the axion/ALP-induced atomic transitions is the recent work on "cosmic parity violation" [139], where $E1$ transitions between the atomic and molecular states of the same nominal parity induced by the background cosmic field that may be part of DM were considered.

18.4 Current and forthcoming magnet technologies

18.4.1 DC magnets

In this section, we attempt to orient the reader in the various types of modern DC magnets that include a variety of resistive magnets, purely superconducting magnets, and hybrid magnets that combine both resistive and superconducting magnets. The development of various magnet types is driven by specific applications and available materials.

CERN is one of the world's centers for development and use of superconducting magnets. Its 14 TeV center-of-mass energy proton-proton collider (LHC) is the largest superconducting system incorporating more than 9000 magnets. The 8 T dipole magnets in the LHC's 27 km long tunnel are based on a 28-strand NbTi/Cu cable with a critical current at 1.9 K of about 20 kA corresponding to a maximum field of 9 T. The feasibility of future higher-energy colliders critically depends on the magnet development. Recent years have seen significant progress towards dipoles with field strength of 16 T.

Detector magnets are typically quite distinct from accelerator magnets, for instance, in their geometry ("barrel" vs. "tube") and volume, with detector magnets typically having orders of magnitude larger volume and stored energy. The world's largest detector-magnet system is currently that of the ATLAS detector at LHC [140]. It consists of a barrel toroid, two end-cap toroids and a central solenoid, generating a 2 T field for the inner detector and a 1 T for the peripheral muon detectors. The detector is 21 m in diameter and 25 m long, with a 8300 m^3 volume containing magnetic field. The magnets operate at 4.7 K with a current of 20.4 kA and a peak field of 4.1 T.

As emphasized above, generating high magnetic fields can only be done by circulating high current densities, which are intrinsically accompanied by strong Lorentz forces. Superconductors allow circulating such currents without ohmic losses, up until a certain limiting value, imposed by the material's intrinsic properties. There is, therefore, no need for large power supplies, nor for large cooling installations, albeit at the expense of the cryogenics needed to maintain the magnet at low temperatures. For the current workhorse material of high-field superconductivity, Nb_3Sn, this current corresponds to a magnetic field on the conductor of around 25 T. The new family of high transition temperature (Tc) cuprate superconductors has critical magnetic field values above 100 T, and therefore seems to be the dream material for the magnet engineer. However, the problem of the Lorentz forces remains, and neither the mechanical strength of high-Tc superconductors, nor their processability, are optimal to build magnets that can handle high forces. The highest magnet field that has been generated with a high-Tc superconducting (HTS) insert is currently 46 T (Sec. 18.4.1). The engineering freedom using non-superconducting metals and alloys is much larger and fields above 100 T have been generated, albeit only in pulsed mode (Sec. 18.4.2) because of the ohmic heating problem. Thus, if sufficient electrical power, and the corresponding cooling capacity are available (plus the budget to operate them), the magnet that generates the highest magnetic fields will be a hybrid one (Sec. 18.4.1), a superconducting outsert, providing several tens of tesla, in which one operates a resistive insert. Only if HTS with the same mechanical properties as the best non-superconducting conductors become available, can one hope to generate the highest possible fields with only superconductors.

Conductors for DC and pulsed resistive magnets

The highest-field DC magnets in Tallahassee, Hefei, Nijmegen, Tsukuba, and Sendai use a copper-silver alloy that is ≈ 24 weight% Ag. The Cu-Ag sheet used for DC resistive magnets is a bulk micro-composite consisting of a eutectic[73] Cu-Ag phase and a proeutectic Cu phase.

[73] Eutectic is a mixture of substances in specific proportions that melts and freezes at a single temperature.

Inside the eutectic phase are layers of Cu and Ag with spacing $< 1\,\mu$m. Inside the proeutectic Cu phase are Ag filaments $< 1\,\mu$m in diameter. This is the material of choice for DC magnets operating in the 30-45 T range as it provides an optimal combination of mechanical strength and electrical conductivity. However, for the higher-field pulsed magnets, strength becomes more important, and Cu-Nb is preferred given its higher strength/conductivity ratio.

Resistive Bitter-type magnets

Conventional superconducting electromagnets – meaning those wound with standard low-temperature superconducting NbTi and Nb_3Sn wire [see Sec. 18.4.1 for recent developments in electromagnets wound with high-temperature (high-Tc) superconducting wire] typically operate at cryogenic temperatures (1.8-5 K) and can generate magnetic fields up to about 25 T in small volumes (typically only a few cm^3 for fields > 22 T), this upper limit being imposed by the field- and temperature-dependent critical current of the superconducting wire itself. To achieve higher magnetic fields, electromagnets must be constructed using resistive conductors, typically copper or higher-strength copper alloy (Sec. 18.4.1). However, this means that the substantial Joule heating in the coil which can amount to many megawatts of power must be continuously dissipated, otherwise the magnet will overheat and fail catastrophically. For steady-state DC electromagnets, the ability to dissipate this heat imposes a limit to the maximum achievable field. The geometry of monolithic solenoids wound with conventional wire makes it difficult to efficiently remove heat. A breakthrough was achieved in the 1930s by Francis Bitter at the Massachusetts Institute of Technology, who used a stack of water-cooled flat copper plates or discs as the conductor instead of conventional wire. These "Bitter plates" were drilled out with many round holes (Fig. 368) to allow cooling water to flow through and around the plates, giving increased cooling capacity [141]. Using Bitter plates, eventually 28 T could be achieved using 20 MW of electrical power. A further advance was the development of the "Florida-Bitter plate" in the 1990s at the NHMFL in Tallahassee; these designs used plates with staggered rows of specifically-engineered elongated holes to optimize the mechanical strength, current-carrying capacity, and cooling capacity of the magnet. Using these Florida-Bitter plates, 30 T resistive magnets were developed in 1995, and 35 T magnets were developed in 2005. These magnets typically had 32 mm diameter bores, were powered with a 20 MW power supply ($\approx$ 40000 A, $\approx$ 500 V), and were continuously cooled with chilled de-ionized water flowing at rates of order 15000 liters/minute [142]. Florida-Bitter-type electromagnets are now in place at many of the high-field magnet laboratories around the world, including the 38.5 T resistive magnet in Hefei (China) and the 37.5 T resistive magnet at Nijmegen (Netherlands).

Driven by advances in materials science and magnet engineering, progress with Bitter-type electromagnets continues to this day. In 2017, an all-resistive 41.4 T magnet was demonstrated at the NHMFL in Tallahassee [143], Fig. 369, using Florida-Bitter plates of high-strength copper-silver alloy and an upgraded 32 MW power supply.

While the primary scientific use of Bitter-type magnets is research in condensed-matter physics and materials science, these magnets may also prove useful for the various types of fundamental-physics experiments discussed in this review, particularly in the cases where large field volumes are not required.

Superconducting magnets

Designing a modern high-performance magnet operating with multi-kA currents that needs to be safe, reliable and show no degradation over many years is a multi-scale challenge (Fig. 370), starting from superconductor-material engineering at the atomic level, all the way to the complete magnet-system design. It is important to understand and control all the links in the

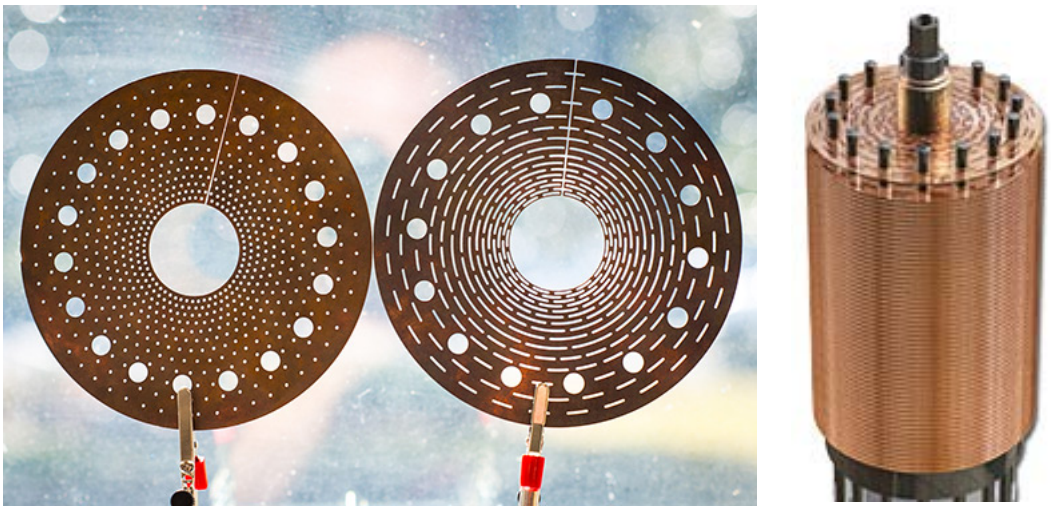


Figure 368: Left: Images of the original "Bitter-plate" design (round holes) and the more recent "Florida-Bitter-plate" design (staggered oval holes), used in the construction of high-field resistive electromagnets. The holes in the plates allow for the efficient flow of cooling water through the magnet, which is needed to dissipate the considerable (10s of MW) resistive heating. The central hole in the plates is about 40 mm in diameter. Right: these plates, with separating insulating layers, can be stacked in a helical configuration to construct a complete magnet.

Figure 369: Photograph of the 41.4 T resistive magnet at the NHMFL-Tallahassee (August 2017). This magnet has a 32 mm bore and is powered with a 32 MW power supply. Its design is based on Florida-Bitter plates (see text) using high-strength copper-silver-alloy conductor. Photo courtesy of the NHFML-Tallahassee.

chain.

From the discovery of superconductivity by Heike Kammerlingh Onnes in 1911, it took about 60 years to develop the first practical multitesla magnets. One of the first such magnets using NbTi wire was developed by Oxford Instruments in 1971 (Fig. 371). Superconducting-magnet records generally follow developments in superconductors. A representative parameter

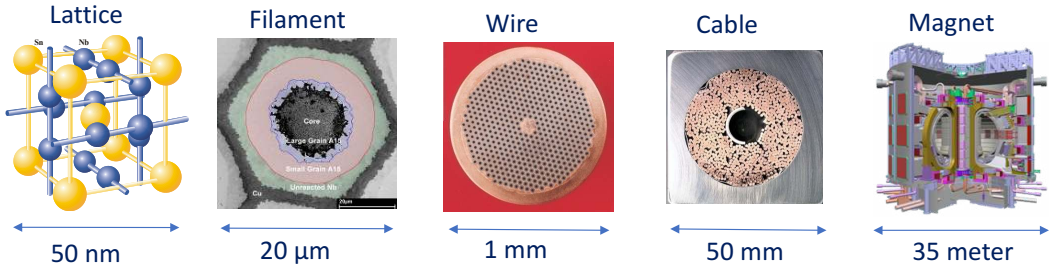


Figure 370: Designing a reliable high-performance magnet is a challenge that spans many scales, from material design at the atomic scale, all the way to engineering of a complete magnet system.

is current density. In the last 30 years, the current densities increased from 500 to 3500 $\mathrm{A/mm^2}$.

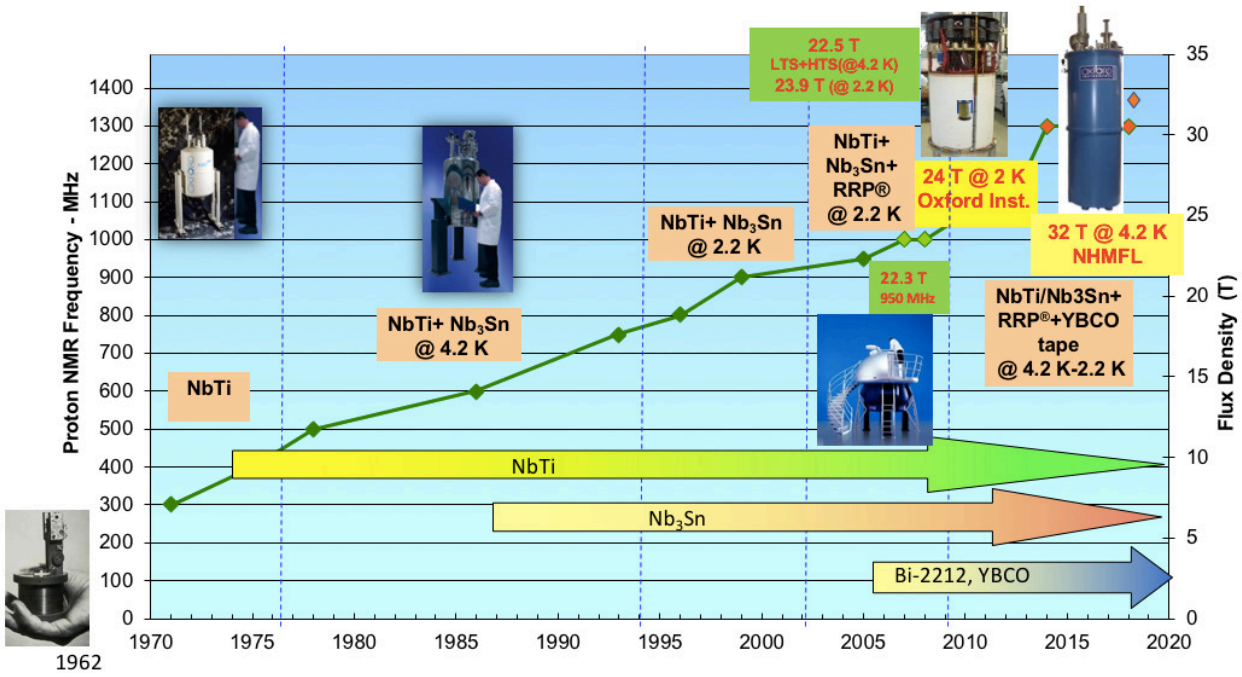


Figure 371: Solenoid magnets for NMR as a representative example of the progress in magnet technology. NHMFL: National High Magnetic Field Laboratory, Tallahassee, Florida.

High-critical-temperature (high-Tc) superconductors are key to much of the modern SC-magnet technology. The first-generation (1G) high-temperature superconducting (HTS) wires that have been commercially available since 1990s are based on the BSCCO (bismuth-strontium-calcium-copper-oxide) material. The more recent second-generation (2G) wire with better electrical and mechanical properties is based on ReBCO—rare-earth-bismuth-copper oxide. The rare-earth elements are typically gadolinium, yttrium, samarium, or neodymium.

Key properties of some superconducting materials are summarized in Table 33.

Superconducting magnets based on high-temperature superconducting wire

As noted above, the field- and temperature-dependent critical current of conventional (low-Tc) superconducting wire (Nb_3Sn wire) limits the maximum achievable magnetic field to of order

Material Name	Class	Critical Temperature (K)	Critical Field B_{c2}	Critical Field@2.2 K	Geometr
NbTi	LTS	9.8	9.5 T @ 4.2 K	11.5 T	Multi-filame round & rectang
Nb_3Sn	LTS	18.1	20 T @ 4.2 K	23 T	Multi-filame round wi
MgB_2	MTS	39	5−10 T @ 4.2 K 1−3 T @ 10 K	N/A	Multi-filame round wi
Bi−2212	HTS	90−110	40 T @ 4.2 K 10 T @ 12 K	N/A	Multi-filame round wi
Bi−2213	HTS	90−110	40 T @ 4.2 K 8 T @ 20 K 4 T @ 65 K	N/A	Tape
YBCO	HTS	92−135	45 T @ 4.2 K 12 T @ 20 K 8 T @ 65 K	N/A	Tape

Table 33: Superconducting materials. LTS, MTS, and HTS stand for low-, medium-, and hight-temperature superconductors. N/A means that these materials as are typically not used below 4.2 K. They can operate at lower temperatures but without particular advantage.

25 T. However, high-temperature (high-Tc) superconductors such as YBCO can exhibit significantly higher critical currents, raising the possibility of a new generation of very high-field superconducting magnets with fields significantly exceeding 25 T. As with many aspects of high-field magnet development, the primary challenge lies in the materials science and engineering of the superconducting wire (or tape) itself. Being ceramic oxides, high-Tc superconductors exhibit significantly different electric and mechanical properties from those of low-temperature superconductor materials. A breakthrough was achieved in late 2017 with the demonstration of an all-superconducting 32 T magnet at NHMFL-Tallahassee [143]; Fig. 372. Using a combination of a low-Tc superconducting outsert (developed in partnership with Oxford Instruments) and a high-Tc superconducting insert (using YBCO conductor developed in partnership with SuperPower Inc.), this 32 T magnet offers a stable and homogeneous field. Given the high critical current of YBCO and other high-Tc materials, it is anticipated that superconducting magnets with considerably higher field can be developed in the coming years [143].

The feasibility of extending the field range of DC magnets all the way to 100 T (the field range so far only accessible in the pulsed mode) was considered [144] with an optimistic conclusion that such a magnet should be possible with present-day technology. The "first-cut" design is based on GdBCO superconductor arranged in a nested-coil configuration (39 coils in total), with a band of high-strength steel over each coil. The design envisions a 20-mm cold bore, a nearly 5.6-m outermost winding o.d., and nearly 17 m winding height. The GdBCO conductor tape is assumed to be 12 mm wide and just under 0.1 mm thick, with a total length over 12 500 km. Its magnet's inductance is 43 kH and its magnetic energy is 122 GJ at 2400 A.

High-reliability HTS cable for operation at low temperature is currently available from several companies. A HTS cable consists of many different layers, as shown in Fig. 373. Magnets based on such cables are still in the development stage; due to the high cost of manufacturing these magnets it is imperative to have a reliable quench protection mechanism. Even though such quench-protection systems have been developed, a more passive system could be advan-

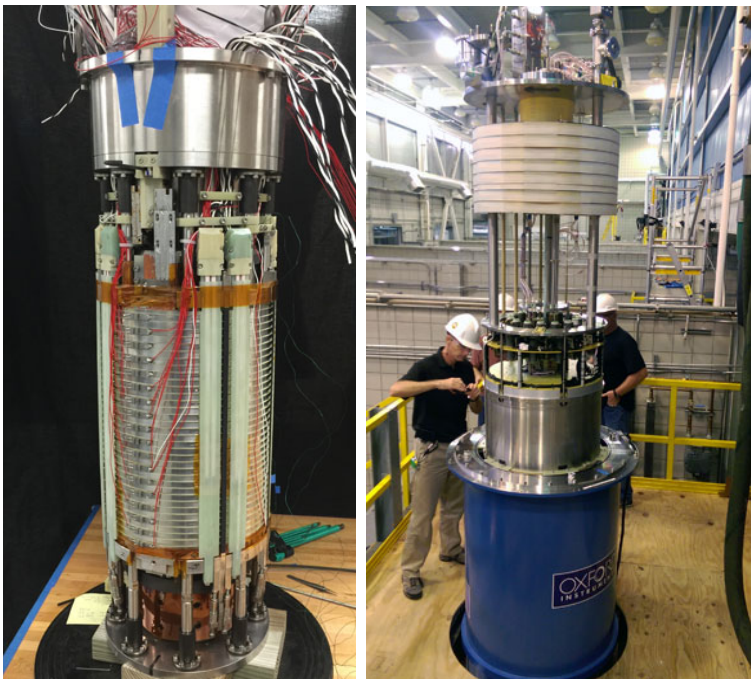


Figure 372: Left: photograph of the innermost two high-Tc-superconductor (YBCO) coils of the all-superconducting 32 T magnet, whose successful operation was demonstrated in December 2017 at the NHMFL-Tallahassee. Total height of the pictured assembly is about 1 m. These YBCO coils are integrated within an outer coil set made of conventional (low-Tc) superconducting wire. The LTS outsert and the cryostat were designed and developed by Oxford Instruments NanoScience. Right: magnet assembly. Photos courtesy of the NHFML-Tallahassee.

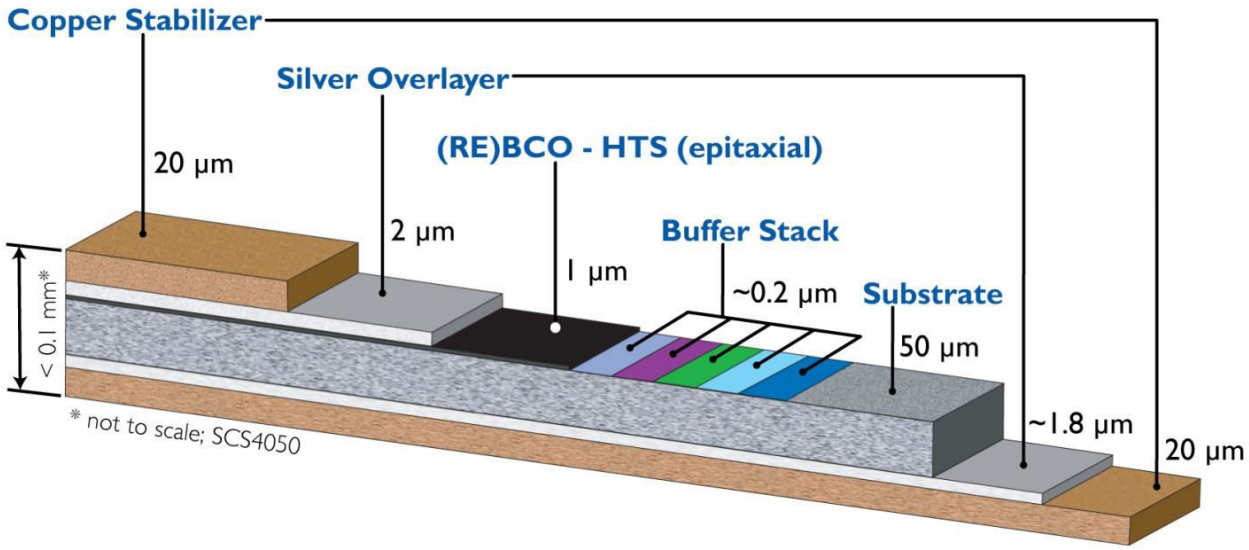


Figure 373: High-temperature superconducting cable design by SuperPower, offering high pinning capabilities to withstand the presence of a large transverse magnetic field. Figure from [145].

tageous to fully protect the magnets. When a quench is developed, it is imperative to safely extract the large stored magnetic energy and prevent generating the equivalent heat in a small concentrated region that can destroy the coil. An alternative to an active quench-protection system is to use no insulation between the turns in the coil; that way the current after a quench

will immediately be diverted into the different turns instead of heating up the quenched spot. No-insulation magnets were invented and applied at the Brookhaven National Laboratory [146] using low-temperature superconductors. But such coils cannot be charged fast which makes them unusable in most accelerator applications. However, some axion/ALP searches require steady-state magnet operation, where the field stability is not critical, the magnet cycle time can be long, and the magnet field uniformity is not critical either.

No-insulation, high temperature superconducting magnets, were first introduced in 2010 at the MIT lab of Francis Bitter [147, 148]. The Center for Axion and Precision Physics Research at the Institute for Basic Science of the Republic of Korea (IBS-CAPP) already has in its laboratory two HTS, no-insulation magnets [149] made by SuNAM (http://www.i-sunam.com). One magnet with an aperture of 3.5 cm is capable of reaching 26 T, while the second one is capable of providing 18 T in a 7 cm inner diameter, over 40 cm long magnet-bore aperture. The 18 T, no-insulation, HTS magnet is the first such magnet currently used in the field of axion dark matter. In addition, IBS-CAPP is commissioning a 25 T, 10 cm inner diameter, 40 cm long-solenoid, no-insulation, HTS magnet with the Brookhaven National Laboratory magnet Division, scheduled for delivery in 2019. The 25 T magnet will also be used to search for axion dark matter.

The no-insulation, HTS magnets exhibit superb stability in steady-state operation, while tests on quenches indicate a robust system even after multiple quenches [147, 148]. Charging the magnet takes much longer than an equivalent for a magnet with insulation, with magnetic field lagging behind the current as it is distributed in all turns at the beginning. However, the current eventually prefers to follow the superconductor instead of the copper stabilizer of Fig. 373. The B-field drifts to its final value over longer times, but the field difference does not constitute a significant parameter in the axion dark-matter search, making these magnets ideally suited for this application.

Hybrid magnets

DC magnetic fields up to about 45 T can be achieved using a hybrid approach, wherein a Bitter-type resistive electromagnet (now the 'insert') is placed within the bore of a large superconducting 'outsert' magnet [150, 151]. Using an 11.5 T superconducting outsert and a 33.5 T resistive insert, a 45 T hybrid magnet (32 mm bore) was first demonstrated in 1999 at the NHMFL-Tallahassee.

A recent advance in hybrid magnet design are the so-called "series-connected" hybrid magnets, in which a superconducting outsert magnet and a resistive (Florida-Bitter type) insert magnet are driven in series by the same power supply, rather than by independent power supplies as is typically the case. Connecting the superconducting outsert and the resistive insert in series has the benefit of significantly reducing, due to the large inductance of the superconducting magnet, the high-frequency magnetic field fluctuations that are otherwise typically present in purely resistive Bitter-type magnets. The result is an extremely "quiet" magnetic field that is well suited to precision measurements requiring high stability (e.g., spin resonance [152]). A 36 T series-connected hybrid magnet with a 40 mm warm bore and $\approx$ 1 ppm uniformity over 1 cm^3 was recently developed and tested at the NHMFL in Tallahassee [153, 154]. With its high stability and excellent field homogeneity, this system is suitable for electron-spin-resonance studies at frequencies up to 1 THz and NMR at $\sim$ 1.5 GHz proton frequency.

As an example of the hybrid-magnet technology we briefly describe a system under construction at LNCMI-Grenoble (Fig. 374) that, in a first step, will produce a continuous magnetic field of 43 T in a 34 mm warm bore [155]. There are only a few hybrid magnets over 40 T currently operating worldwide [156] as there remains a technological challenge due to the necessity for the superconducting coil to withstand quenches and large electromechanical forces from the warm insert, especially under fault conditions. A feature of the LNCMI system (also

available for some other hybrid magnets) is that it is reconfigurable, offering the users a range of configurations from 43 T in 34 mm to 9 T in 800 mm diameter bore. This is accomplished by using different combinations of superconducting and resistive coils. The superconducting coils are based on the proven Nb-Ti/Cu material and are cooled with pressurized superfluid He to 1.8 K, producing in excess of 8.5 T. The magnet requires a dedicated He liquefier with a production capacity of 140 l/h. The 13 km long superconductor is made from a flat Rutherford cable (a design named after the laboratory where it was introduced) containing 19 Nb-Ti strands of 1.6 mm diameter soft-soldered onto a copper-silver hollow "stabilizer" with both ends connected to the pressurized superfluid He bath. Operation at low temperatures enhances the superconducting properties of the conductor while the use of superfluid helium allows to profit from its uniquely high thermal conductivity. As an illustration, the heat flux transported by conduction between 1.9 and 1.8 K in a 1 m long static channel of superfluid He can be up to about three orders of magnitude higher than that of a bar of oxygen-free high thermal conductivity (OFHC) copper of the same geometry. The resistive part of the 43 T, 34 cm-diameter-bore hybrid magnet will be a combination of resistive polyhelix and Bitter coils using 24 MW of electrical power [157]. By combining the Bitter insert alone with the superconducting coil, another hybrid-magnet configuration will allow to produce 17.5 T in a 375 mm diameter aperture with 12 MW. A crucial part of the magnet design is the control and protection system. Indeed, in case of a fault, it would be necessary to safely guide and dissipate the 110 MJ of energy stored in the hybrid magnet. A upgrade of the Grenoble hybrid magnet to exceed 45 T in the 34 mm diameter bore will be possible thanks to the ongoing two-phase upgrade of the electrical power installation, from 24 up to 30 MW and then up to 36 MW.

Looking towards the future, preliminary designs for a 60 T hybrid magnet have been recently initiated at NHMFL [150].

18.4.2 Pulsed magnets

To attain magnetic fields in excess of $\approx$ 45 T (the approximate limit of hybrid DC resistive magnets today), pulsed magnets are used. In effect, this avoids the significant "cooling problem" described above; in a pulsed magnet the field is ramped up to a high value and back down again before Joule heating in the copper-alloy windings destroys the magnet. High magnetic fields $>$ 50 T and even exceeding 100 T can be achieved nondestructively for timescales typically in the range of 1-100 ms. For pulsed magnets, the design challenge centers primarily on high-strength materials (both reinforcement materials and the conductor itself) and clever engineering design to cope with the huge magnetic pressures that exist due to the Lorentz forces associated with the high fields (50-100 T) and currents ($\sim$ 30 000 A) in the magnet. Most often, a large capacitor bank capable of storing many megajoules of energy is used to power the magnet; however there are also other schemes employing large motor-generators (essentially, flywheels) as power sources. For example, at the NHMFL-Los Alamos a 1.4 GW motor-generator is used to store up to 600 MJ of (rotational kinetic) energy in a 300-ton steel shaft, which is used to drive one of the two large pulsed magnets, described briefly in the following section.

Magnetic fields significantly in excess of 100 T can be achieved for microsecond time durations in small volumes of order 1 cm^3 using copper coils, but these methods are either semi-destructive or completely destructive. So-called "single-turn coils" are made from sheets of copper plate (typically 2-3 mm thick) that are formed into a single loop with diameter of order 1 cm (Fig. 376). A current pulse of $\approx$ 3 MA from a fast (low-inductance) capacitor bank can generate magnetic fields in the range of 150-300 T for a few microseconds before the coil itself is vaporized [159, 160]. A significant advantage of the single-turn-coil technique is that the experiment in the coil (usually) remains intact; in this sense, the technique is semi-destructive. Single-turn magnet systems are currently operating at the Tokyo, Toulouse, and Los Alamos

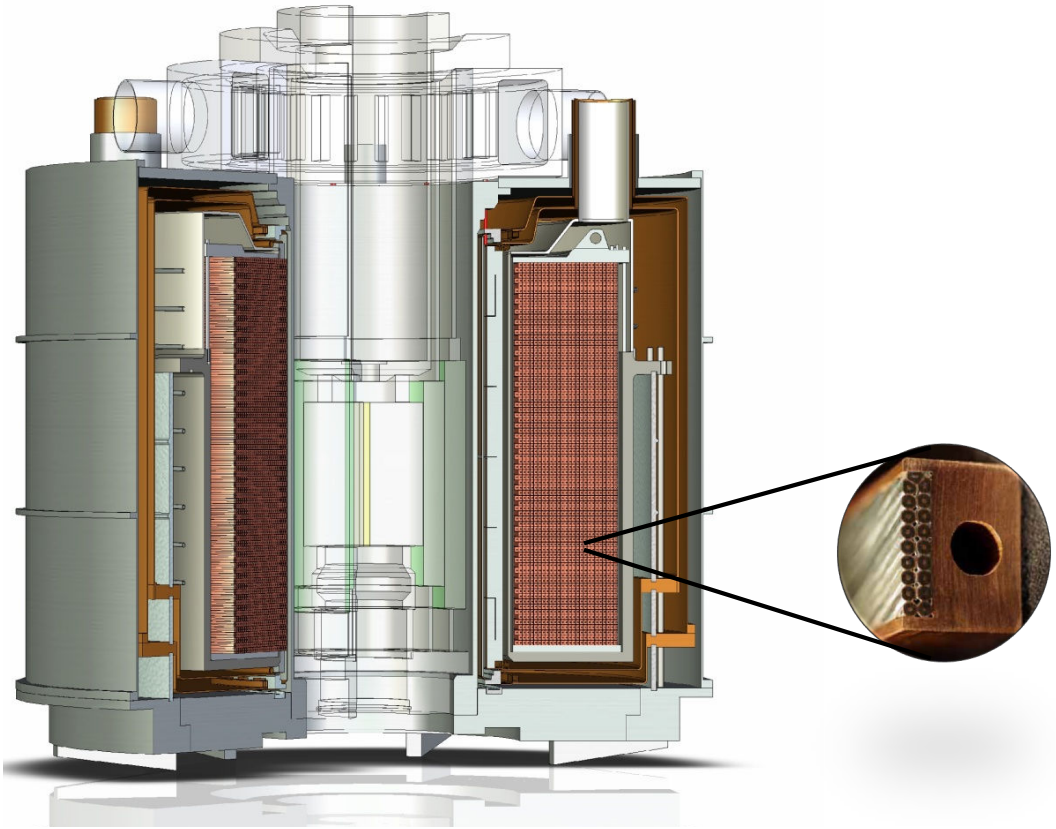


Figure 374: Three-dimensional cut-away view of the cryostat and superconducting coil of the Grenoble hybrid magnet under construction in collaboration with Saclay Nuclear Research Center (CEA-Saclay). The superconducting coil is made of the assembly of 37 interconnected double pancakes with inner-bore diameter of 1111 mm and a total height of 1400 mm. The total weight of the hybrid magnet is about 53 tons, including the 24 ton mass cooled down to 1.8 K with superfluid He. The inset shows a superconductor of 17.92 mm in height $\times$ 12.96 mm in width.

magnet laboratories.

Significantly larger magnetic fields exceeding 1000 T can be achieved on microsecond timescales using explosive or electromagnetic flux-compression techniques that are entirely destructive. Here, a background seed field of several tesla is first introduced into a large metal "liner" (effectively a short copper cylinder with diameter of order 10 cm). The liner and the magnetic flux lines contained within are then rapidly compressed down to a small diameter, resulting in a corresponding significant increase in magnetic field. Flux-compression techniques using high explosives were explored in the 1990s at both Los Alamos (USA) and Sarov (formerly Arzamas-16, Russia), where magnetic fields in the range of 1000-2800 T were reported [161]. Alternatively, electromagnetic flux-compression (EMFC) methods utilize a large single-turn primary coil placed around the liner. When the primary coil is energized (typically with $\approx$ 5 MA of pulsed current from a capacitor bank), the nearly-equal-but-opposite secondary current induced in the liner results in forces that drive the liner inward, boosting the magnetic field. EMFC methods are currently practiced at the IMGSL (Japan), where magnetic fields exceeding 600 T are routinely achieved (see Ch. 7 of the book [162] and [163]). Generation of fields approaching 1000 T via electromagnetic flux compression were reported recently [164].

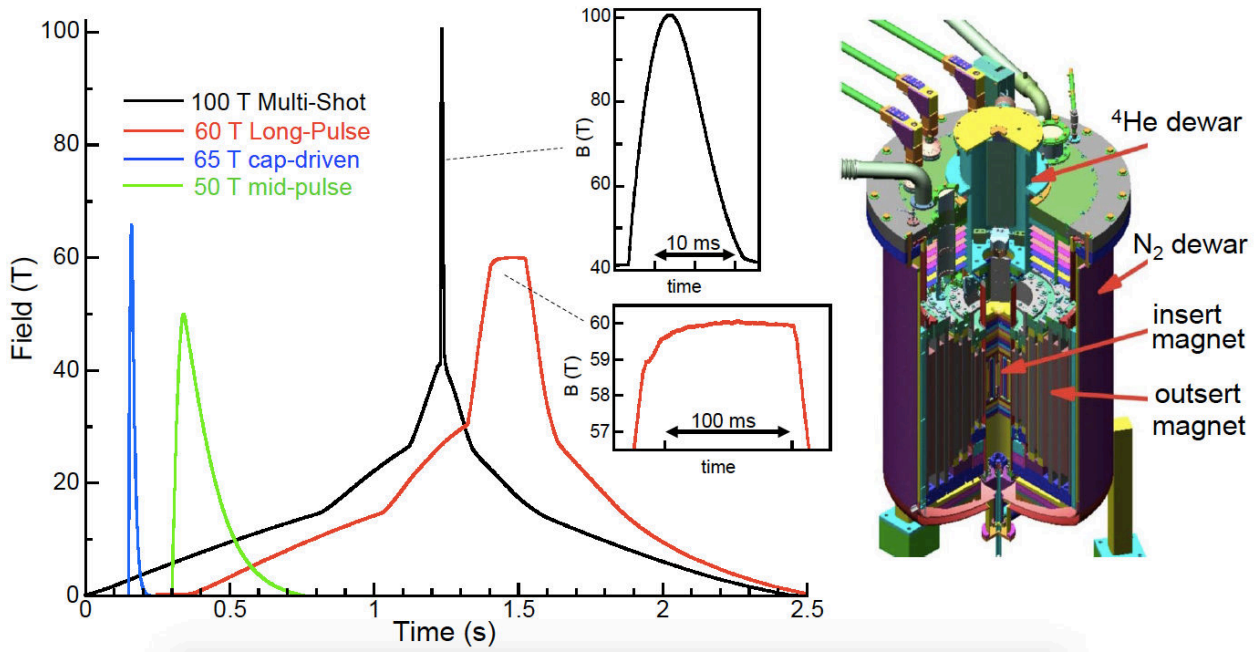


Figure 375: Measured field profiles of the various nondestructive pulsed magnets at the NHMFL-Los Alamos. The 65 T "standard-user" magnets are powered with a 4 MJ capacitor bank, as is the larger-inductance 50 T "mid-pulse" magnet. The 60 T Long Pulse magnet, which can maintain 60 T for 100 ms (or, for example, 50 T for 200 ms), has a user-defined waveform and is driven with a 1.4 GW motor-generator. The 100 T Multi-Shot magnet, also depicted in the diagram on the right, has a generator-driven $\approx$ 40 T outsert magnet and a $\approx$ 60 T insert magnet that is driven with a 2.2 MJ capacitor bank; it achieved fields $>$ 100.7 T in early 2012. The total height of the magnet assembly is $\approx$ 2 m.

Pulsed magnets at NHMFL Los Alamos

The 60 T Long-Pulse Magnet, which came online in 1998, can maintain a constant peak field of 60 T for up to 100 ms, in a 32 mm bore. This magnet is driven with a 1.4 GW motor-generator as mentioned above. The overall pulse length (see the field profile in Fig. 375) is about 2 s. Such a long duration at peak field in comparison with typical capacitor-driven pulsed magnets, for which time at peak field is of order milliseconds, allows for sensitive experiments that benefit from significant signal averaging or extended photon collection, such as the case in many optics and laser experiments.

The 100 T Multi-Shot Magnet, which consists of a large generator-driven resistive outsert magnet (40 T peak field in a large 225-mm bore) and a capacitor-driven insert magnet ($\approx$ 60 T), achieved a peak field of 100.75 T in 2012. The time at peak field is of order milliseconds (see field profiles in Fig. 375), while the overall pulse duration is $\approx$ 2.5 s.

Pulsed magnets developed at LNCMI Toulouse

The pulsed magnets at LNCMI are cooled with liquid nitrogen and produce long pulses with a duration on the order of 100 ms at a repetition rate of 1 pulse/hour. The typical sample-space dimensions are $5 - 20$ mm. The LNCMI 100 T magnet is depicted in Fig. 377. A test of the system up to 98.8 T is described in [165].

Timing diagrams for several representative LNCMI pulsed magnets is presented in Fig. 378.

In LNCMI pulsed magnets, the dissipated energies are in the MJ range, corresponding to

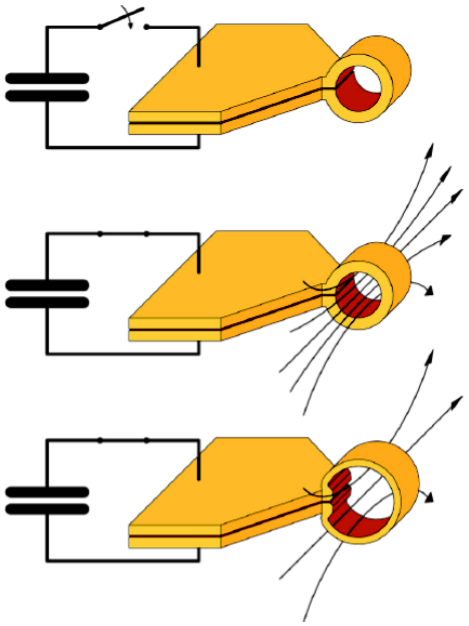


Figure 376: A schematic of a single-turn pulsed coil. The rise time of the field is determined by the capacitance and inductance of the entire circuit, which includes the capacitor bank, cabling, and the coil itself. Once the current flows through the coil, magnetic field is generated and its pressure explodes the coil. The disintegration time of the coil is determined by its mass (inertia) and the applied magnetic force. The circuit impedance can be adjusted to make the rise time of the field shorter than the disintegration time of the coil. Typical values of the field achieved with this technique are 300 T in a 5 mm diameter or 100 T in a 20 mm diameter, both with ≈ 5 μs duration. Figure courtesy LNCMI [158].

peak powers in the GW range. The mechanical stresses on the coils are in excess of 3 GPa (30 000 atm). These features necessitate extremely careful mechanical design and thermal management in order to ensure long lifetime of the magnet components and avoid potentially violent failures. The ability to dissipate the heat generated by a field pulse determines how frequently the magnet can be pulsed.

In many cases, it is desirable to combine high magnetic fields with other unique facilities, necessitating the development of "portable" high-field magnets. An example is a high-field magnet (Fig. 379) developed by LNCMI for x-ray studies at the European Synchrotron Radiation Facility (ESRF) [166]. This is a horizontal-field, 30 T magnet system with a conical bore optimized for synchrotron-radiation powder diffraction. The magnet provides a wide-angle optical access downstream of the sample, which allows to measure many Debye rings enabling accurate crystal-structure analysis.

Another example is a pulsed 40 T magnet operating at the neutron-diffraction facility at the Institut Laue-Langevin (ILL). The magnet also features a wide-angle conical access to a 8-mm diameter sample whose temperature may be adjusted within the $1.5 - 300$ K range. The magnet produces 100-ms long pulses every 6 minutes corresponding to the total "field-on" time of about 20 s/day. Recently, this device was used to study magnetically ordered phases in an enigmatic material URu_2Si_2 with strongly correlated electrons [167]. In a strong magnetic field, the material undergoes a phase transition into magnetically ordered phases. Neutron diffraction under pulsed magnetic fields was used to identify the field-induced phases as a spin-density-wave state.

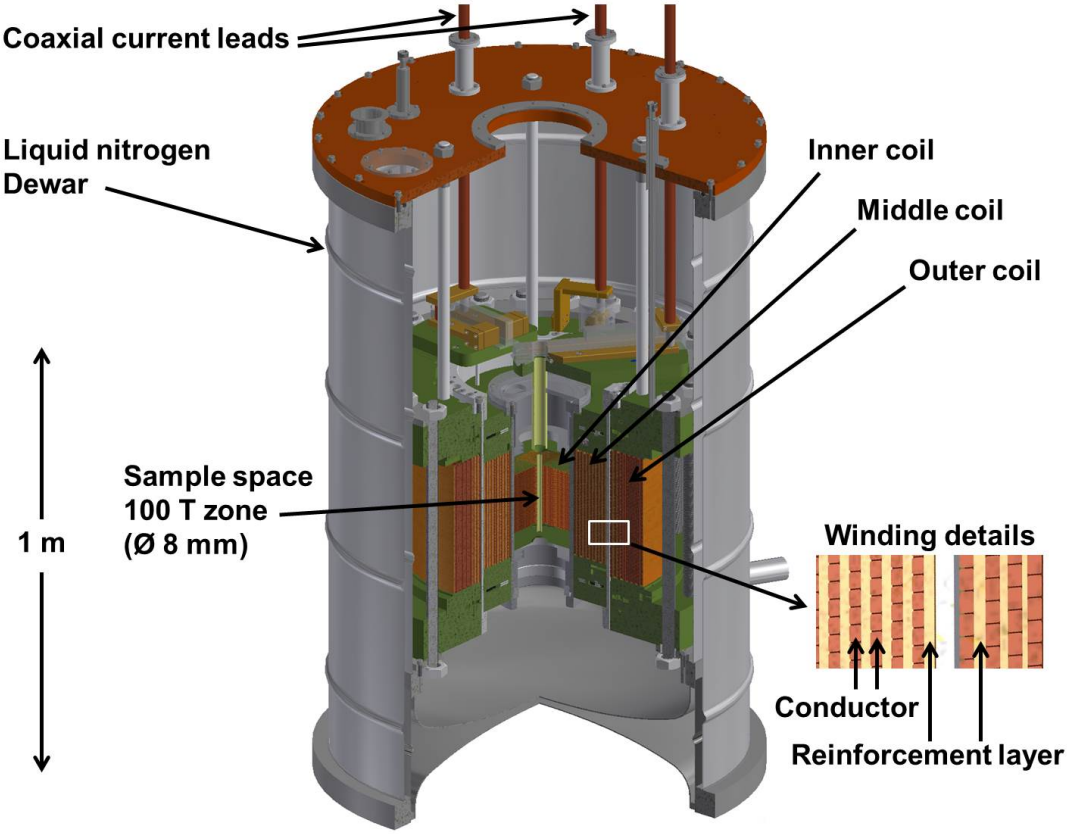


Figure 377: The LNCMI 100 T magnet. The magnet is composed of three concentric coils powered with three independent capacitor banks. Before a pulse, the magnet is cooled with liquid nitrogen. Winding details show alternating conductor and reinforcement layers. Conductors are insulated wires made of high-strength copper alloys or micro-composites and reinforcement is made of Zylon polymer fibers. Figure courtesy LNCMI−Toulouse.

18.5 Techniques for measuring high magnetic fields

Most experiments using magnets do not only require generation of a field with particular sets of parameters, the experimentalist also needs to know what the values of the parameters are. In many cases, direct measurements are required, sometimes referred to as "magnetic-field metrology." We briefly discuss a selection of measurement techniques in this section.

18.5.1 DC fields

NMR is frequently the technique of choice for DC-field measurements. Commercial "NMR precision teslameters" cover the range of fields to $\approx$ 10 T with proton-based NMR probes (the proton gyromagnetic ratio is 42.5774806(10) MHz/T) and to $\approx$ 23 T with deuterium based probes (the deuteron gyromagnetic ratio is 6.53569(2) MHz/T). Typical precision of such devices is in the several parts-per-billion (ppb) range, while relative accuracy is in the parts-per-million (ppm) range.

Significantly higher relative accuracies of better then a part in 10^{12} were demonstrated with NMR magnetometers based on gaseous ^{3}He [168] (absolute accuracy is limited to roughly a part in 10^8 by the finite accuracy of fundamental constants). In these magnetometers, nuclear

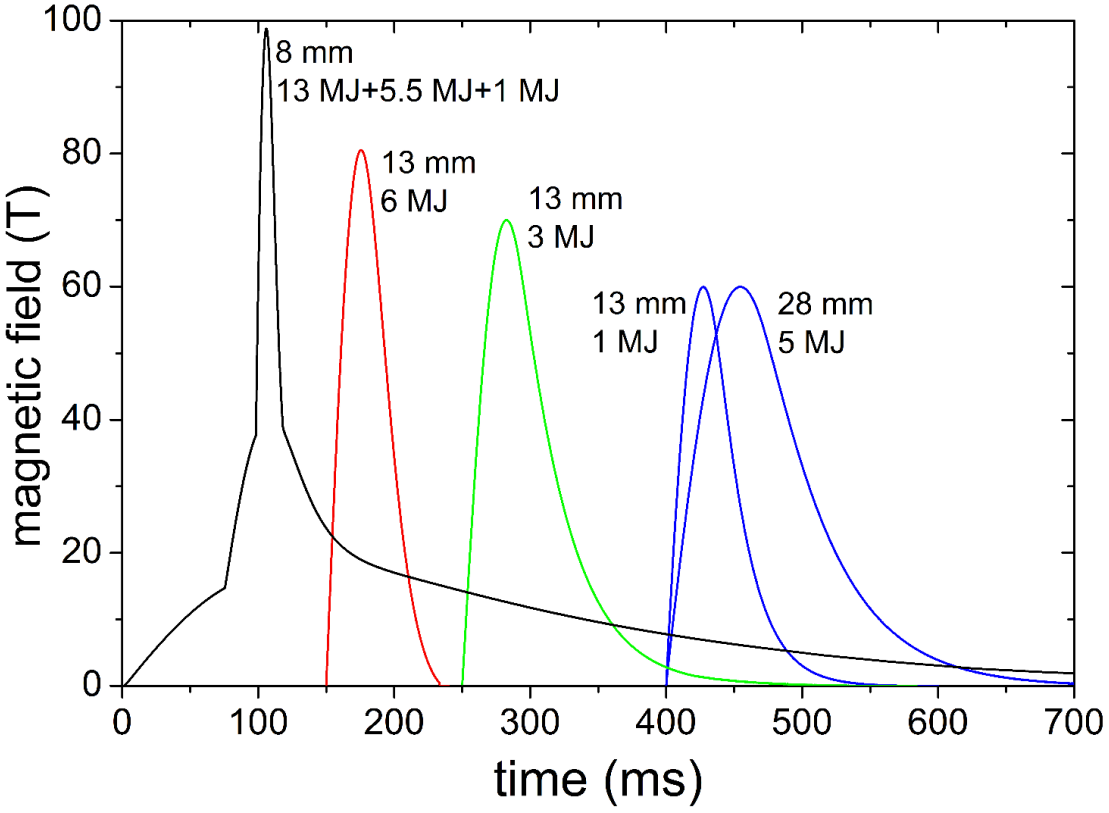


Figure 378: Temporal profiles of magnetic pulses realized by the solenoids available at LNCMI. The diameter of the free bore and the magnetic energy stored at the maximum magnetic field are indicated. The 100 T magnet is composed of three concentric solenoids fed by three independent capacitor banks. The three different energies indicated are the three magnetic energies for the sub-coils from the outside to the inside. Figure courtesy LNCMI−Toulouse.

spins are polarized via optical pumping. Since the atomic transitions of helium are deep in the ultraviolet, the atoms are excited into a metastable state in a discharge. There are convenient near-infrared transitions from the metastable excited state, so the electronic and nuclear spins can be polarized by a repeated absorption/de-excitation cycle. A large fraction of nuclei in the gas cell are eventually polarized where the atoms in the metastable states transfer their excitation and nuclear polarization to the ground-state atoms in a process known as "metastability exchange." The efficiency of optical pumping decreases at high magnetic fields where the Zeeman energy exceeds the hyperfine-interaction energy and the hyperfine structure is "decoupled." This has limited the operation range of such magnetometers to about 12 T. The magnetometry protocol consists in measuring the frequency of the free-induction decay of the ^{3}He spins after excitation with a resonant radio-frequency pulse. The coherent spin-precession times in such magnetometers are on the order of minutes, which is achieved due to averaging of magnetic-field inhomogeneities by helium atoms sampling the entire volume of the cell during the free precession and by ensuring accurate spherical shape of the cell [168].

A particular challenge is metrology for the highest-field DC magnets (up to 45 T), such as the resistive Bitter-type magnets and hybrid magnets described in Sec. 18.4.1. One technique, widely employed at the NHMFL-Tallahasee, is careful determination of the field-current cali-

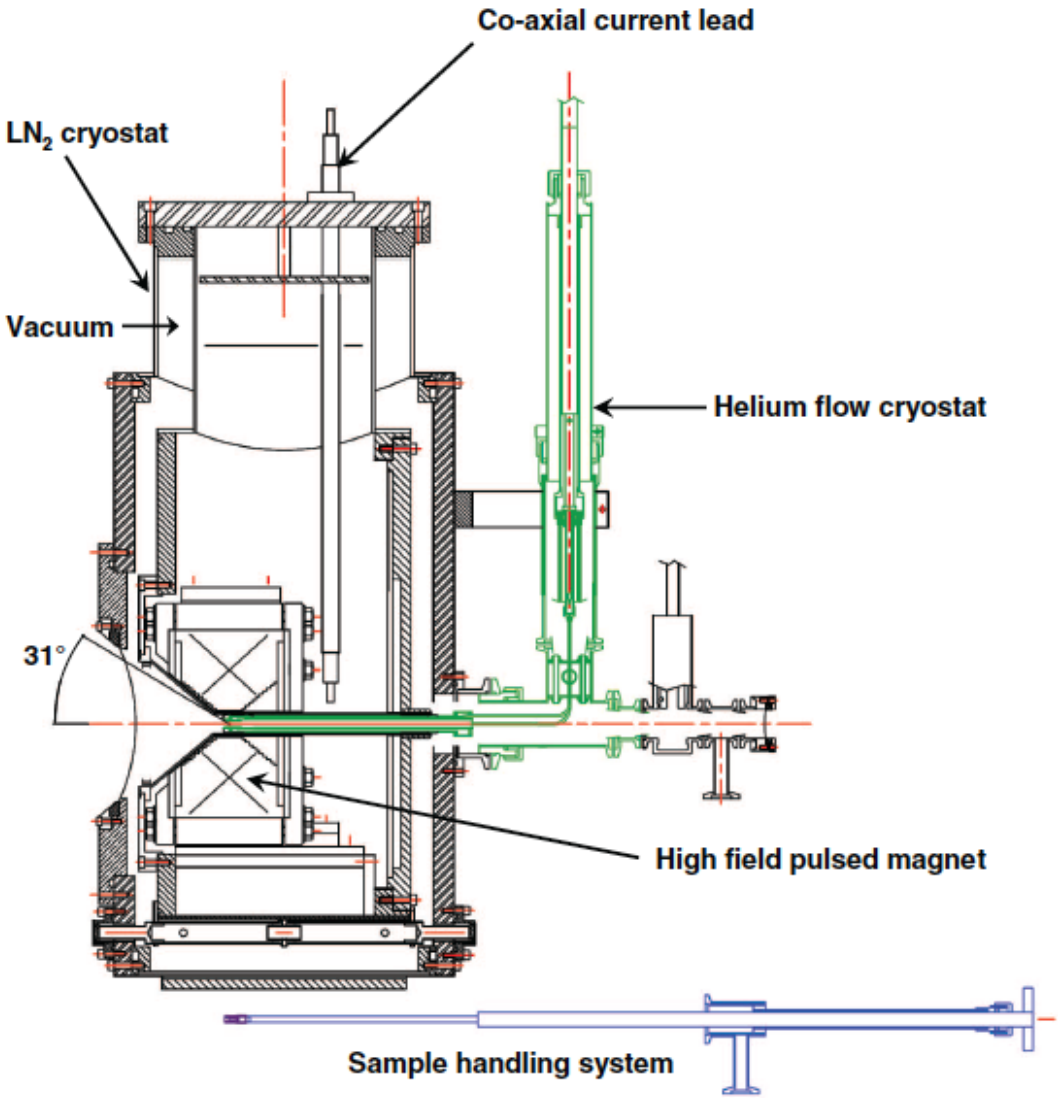


Figure 379: A 30 T pulsed magnet for X-ray studies operating at ESRF. The solenoid coil is 16 mm in diameter and is maintained at 77 K. The total pulse duration is up to 60 ms with a repetition rate of 6 pulses/hr at maximum field. The sample can be maintained at a temperature in the range of $1.5 - 300$ K in a liquid-helium cryostat. Figure from [166].

bration curve (B versus I) for each magnet using the following method: First, a measure of the magnetic field is obtained from a pickup coil in the magnet's bore. By Faraday's law of induction, the voltage induced in the pickup coil is related to the time-derivative of the magnetic field as the current is ramped from zero up to its peak value. Integrating the signal from the pickup coil therefore gives the (uncalibrated) magnetic field versus current. This curve is then calibrated by scaling it with respect to known specific fields (and therefore currents) that are separately identified with fixed-frequency NMR probes. This method provides field-measurement accuracy better than one-tenth of one percent, which is found to be superior to that obtained using four-wire Hall effect probes (owing largely to the temperature- and slightly nonlinear field-dependence of the high-field Hall effect in most materials).

18.5.2 Pulsed fields

For pulsed high-field magnets with millisecond time scales of the field evolution, a typical goal for magnetic field metrology is 0.1% accuracy with 1 kHz bandwidth, while pulsed-coil

diagnostics requires even higher bandwidth on the order of 100 kHz and higher accuracy may be required for particular applications.

There are a number of constraints typically imposed on a magnetic-field measurement in a strong-field magnet, including the need for linear response over a wide range of fields, small (typically, several mm) sensor size, construction out of nonmagnetic and mostly nonconductive materials, and the need to operate in a wide temperature range from 400 K down to 0.1 K.

A typical implementation of the measurement device is a pick-up coil, in which the induced voltage is proportional, according to Faraday's law of induction, to the rate of the magnetic-flux change through the coil. An immediate concern for such a device is that it sensitive not only to the rate of change of magnetic induction but also to vibrations that can change the effective area of the coil. A pick-up unit is usually constructed by winding several turns of copper enameled wire around a ceramic or low-thermal-expansion plastic bobbin. The device needs to be calibrated, carefully positioned and aligned and mounted to ensure vibration damping. In a carefully implemented device of this kind, overall relative accuracy of $\sim 2 \cdot 10^{-3}$ is achieved with standards laboratory equipment, while better than 10^{-4} was reported with metrology-grade equipment.

The pickup coils used in the pulsed magnets at the NHMFL Los Alamos are typically calibrated using the de Haas - van Alphen (dHvA) quantum oscillations in copper. The dHvA effect is a quantum mechanical phenomenon wherein the magnetization of a material oscillates as a function of magnetic field due to the Landau quantization of electron energy. The oscillations are periodic in $1/B$, and occur at a well-known frequencies in simple cubic-lattice metals such as copper. Counterwound copper pickup coils, typically used for measuring dHvA signals in other samples of interest, also show clear quantum oscillations due to the copper wires themselves. The actual magnetic field is calibrated by comparing the measured dHvA frequency with the known dHvA frequency for copper.

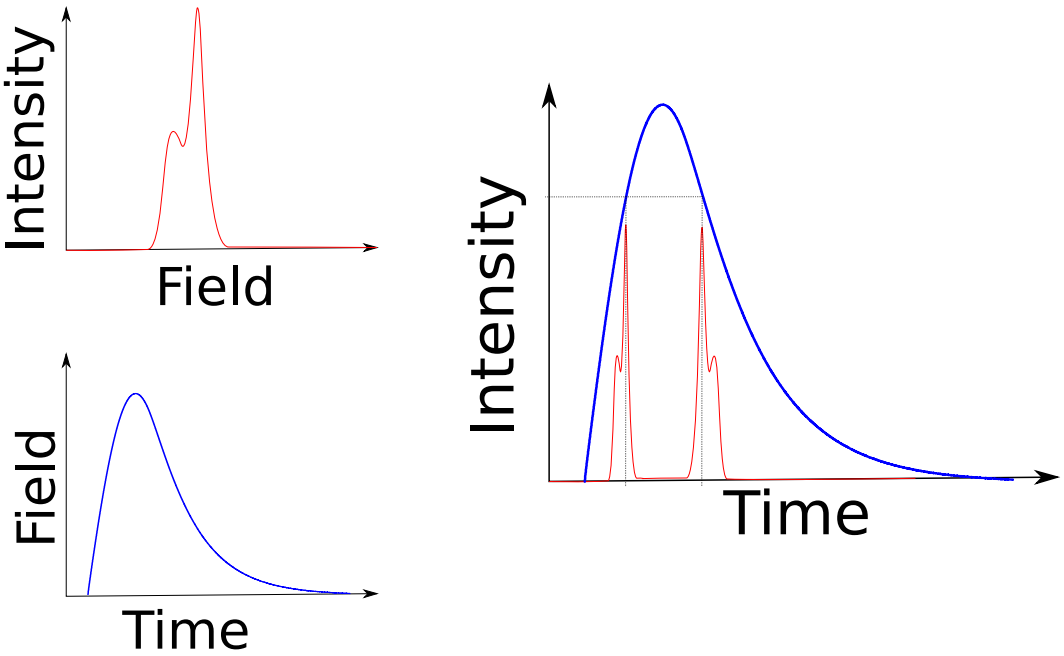


Figure 380: Resonant marking of time-variable magnetic field is based on sharp dependence of some spectroscopic signal, for instance, atomic absorption or NMR, on the applied field (upper left). During the temporal evolution of the pulsed magnetic field (lower left), the field passes through the resonant value, typically, twice, leading to the time-dependent signals marking the times the resonance is achieved (right).

Another technique for pulsed-field metrology–resonant marking–allows to pinpoint the precise times during the magnetic-field pulse when the resonant value of the field is reached (Fig. 380).

A variety of nonzero-spin nuclei have been used for NMR marking. Most commonly, these are ^{1}H protons and ^{2}H deuterons, but ^{63}Cu has been a choice for some of the recent work. The magnetic moment of ^{63}Cu is about a factor of four lower than that of the proton, allowing to keep NMR frequencies in the sub-GHz range and the magnitude of the magnetic moment is known with a 10^{-5} accuracy. Moreover, the Knight shift of the NMR frequency due to the conduction electrons in a metal is also at the same relative level of 10^{-5}. Additional advantages of copper include the fact that it remains solid in a broad range of temperatures and that it is a simple metal in that there are no field-induced transitions at low temperatures and high magnetic fields. All of these factors make ^{63}Cu a good reference material for high-field measurements, particularly those at low temperature.

The progress in RF electronics, NMR methodology and magnet technology make it now possible to obtain reasonable signal-to-noise ratios in NMR measurements using pulsed field magnets with typical pulse durations of tens of milliseconds, on certain classes of samples. These samples should have high abundances of nuclei with large gyromagnetic ratios (see Table 34), and should have rather short nuclear relaxation times, preferably shorter, but definitely not much longer, than the field pulse duration. This is generally true for systems containing paramagnetic ions, conduction electrons or quadrupolar nuclei. Although this technique cannot compete with NMR in DC fields in terms of sensitivity or resolution, it can give access to NMR data in the field range above 45 T. The current field limit of pulsed-field NMR experiments is 80 T, and there is no reason to believe that it could not be implemented in the highest field pulsed magnets, which currently generate up to 100 T.

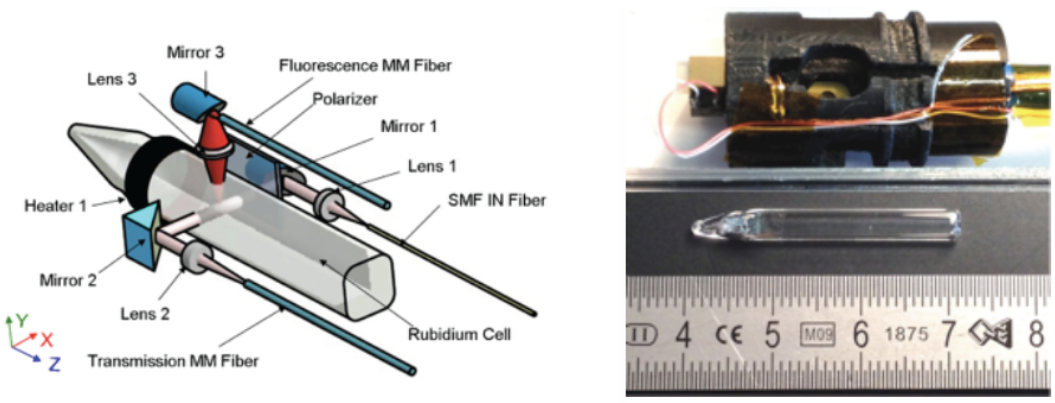


Figure 381: High-magnetic-field probe based on Rb optical spectroscopy [170]. Left: optical measurement head; MM and SM stand for multi-mode and single-mode (optical fibers), respectively. Right: Rb vapor cell and 3D-printed holder.

A technique to measure pulsed magnetic fields based on the use of rubidium in gas phase as a metrological standard was demonstrated in Ref. [170]. The idea is that the magnetic field "tunes" the energies of the Zeeman sublevels in the upper and the lower state and brings the transition on resonance with light of a fixed and precisely measured frequency, increasing absorption and fluorescence. The authors of Ref. [170] developed an instrument based on laser inducing transitions at about 780 nm (D2 line) in rubidium gas contained in a cell of 3×3 mm^2 cross section (Fig. 381). A temperature-stabilized fiberized probe was used to insert the cell into a high-field pulsed magnet. Transition frequencies for light linearly polarized parallel and perpendicular to the magnetic field were measured with a wavemeter. Both the light transmission and fluorescence (from a volume of 0.13 mm^3) were monitored. The sensor was

Nucleus	Relative sensitivity	Natural abundance	f_{NMR} at 40 T (MHz)
^{1}H	100	100	1703.10
^{7}Li	29.3	100	661.89
^{11}B	16.5	80.4	546.42
^{19}F	83.3	100	1602.51
^{27}Al	20.6	100	443.77
^{31}P	6.6	100	689.43
^{45}Sc	30.1	100	413.71
^{51}V	38.2	99.8	447.97
^{55}Mn	17.5	100	422.18
^{59}Co	27.7	100	404.10
^{65}Cu	11.4	30.9	483.74
^{71}Ga	14.2	60.4	519.39
^{93}Nb	48.2	100	416.85
^{115}In	34.7	95.7	373.19
^{121}Sb	23.9	100	407.56
^{141}Pr	29.3	100	521.49
^{151}Eu	17.8	47.8	423.39
^{165}Ho	18.1	100	363.44
^{205}Tl	19.2	70.5	981.56
^{209}Bi	13.7	100	273.68

Table 34: Nuclei suitable for pulsed-field NMR. Relative sensitivity is it the product of gyromagnetic ratio and nuclear moment, arbitrarily normalized to 100 for the proton. It gives the NMR signal strength one can expect with respect to proton NMR at the same concentration and field.

operated up to fields of 58 T and allowed precise scaling of the temporal profile of the magnetic field obtained with a pick-up-coil. In this way, the magnetic field strength during the entire duration of the pulse was determined with an accuracy of $\approx 2 \cdot 10^{-4}$, better than with the calibrated pick-up coil by more then an order of magnitude. Interestingly, the absolute accuracy is not limited by the experimental method but rather by the uncertainty in the Landé g-factor of the 5p Rb excited state. Once the g-factor is known with high precision, the overall accuracy can be improved to the 10^{-5} level [171].

In the highest-field pulsed magnets such as single-turn and flux-compression magnets, the field is often measured via magneto-optical (Faraday) rotation. In this method, linearly polarized light propagates in a medium (e.g., glass) in the direction of the magnetic field. The polarization plane of the light rotates by an angle given by a product of the length of the sample, the strength of the magnetic and a so-called Verdet constant, which is a characteristic of the material. Major strengths of this method are that it allows for measurements that are largely immune to electromagnetic pick-up accompanying the magnetic-field pulse, the ability to measure the field evolution on fast (μs) scales, and a large dynamic range. Faraday-rotation magnetometry is used for measuring the strongest laboratory fields [169].

18.6 Conclusions and outlook

In this paper, we have provided an overview of modern technologies for generating and measuring both DC and pulsed magnetic fields and discussed a representative sampling of currently ongoing and planned fundamental-physics experiments using high magnetic fields. There is certainly impressive progress in the field and no lack of important measurements to carry out with

strong magnetic fields and a number of novel ideas and experimental concepts. We conclude with some of the highlights and and outlook of forthcoming developments in high magnetic field technologies.

- Both DC and pulsed resistive magnets have been working reliably for decades. Fields continue to rise via more sophisticated engineering and larger power supplies. Magnets with geometries adapted to specific experimental requirements are possible and exist, for example, at NHMFL- Tallahassee (25 T split) and Berlin (25 T conical) for scattering experiments, as well as at LNMCI-Toulouse and in Tokyo, where these magnets are used for vacuum-birefringence experiments, see Sec. 18.2.1.

- High-Tc superconducting magnets producing DC fields up to 32 T already exist and will become generally available in the next decade but will require significant investments and dedicated operating environment. Lower field values but with geometries adapted to specific experimental requirements (related to the figures-of-merit such as those listed in Table 28) are feasible but require considerable engineering effort.

- The NHMFL all-superconducting magnet that reached 32 T has a thoroughly tested quench-protection system that allows this to be the first SC magnet for physics experiments to operate above 24 T. This 8 T leap in field (33%) is a harbinger of a revolution in SC magnets for a variety of fields of science. To put this in perspective, the previous 8 T increase (from 16 T to 24 T) took 42 years (1975 – 2017, Fig. 371). The step from 24 T to 32 T took about a year (24 T became available in January 2017).

- There is a "quiet revolution" going on in the field of superconducting magnets, owing to recent advances and development of high-Tc superconductors made from, for example, YBCO. Whereas the critical current of conventional low-Tc superconducting wire (NbTi, Nb_3Sn) limits the maximum achievable field to about 23.5 T, high-Tc materials such as YBCO remain superconducting even in magnetic fields exceeding 45 T (see Table 33). Already, recent developments described above have boosted the peak field in a YBCO-based all-superconducting magnet to 32 T at NHMFL- Tallahassee. As better and stronger wires and tapes are made from these high-Tc materials, the maximum fields that are achievable with all-superconducting magnets will increase correspondingly. Similarly, the physical size and design complexity of the magnets can also increase, which is important for many magnets designed for fundamental science, where a combination of field strength and magnet size (length, volume) are key parameters (Table 28).

- Next-generation SC magnet technology should allow both higher fields than before as well as to allow experiments previously conducted in resistive magnets to move to SC magnets where there will be less field ripple, less acoustic noise, lower operating costs, and more time available.

- Significant advances in superconducting magnets for a variety of applications have been demonstrated by commercial companies using low-temperature superconducting materials. This includes the high-field MRI magnets (9.4 T with 900 mm bore) as well the high-field NMR and research magnets up to 23 T, which are narrow-bore systems using NbTi and Nb_3Sn materials. Recently, a new class of compact high-field, wide-bore magnets developed by commercial companies is facilitating all-superconducting magnets for over 30 T using LTS and HTS materials. The 32 T all superconducting magnet at NHMFL used a 17 T HTS insert developed by NHMFL which was integrated into a 15 T/250 mm LTS outsert developed by Oxford Instruments.

- Hybrid magnets producing 45 T will become more available in the coming years. The next-generation of hybrid systems that are currently discussed are expected to come online around 2030.

- Pulsed magnets producing 100 T with a few millisecond pulse duration are already here (in Los Alamos since 2012) and will become more available during the next years. Lower field values but with geometries adapted to specific experimental requirements are quite feasible but require engineering efforts.

High magnetic fields have enabled important fundamental-physics discoveries. One example involving small-scale condensed-matter physics is the discovery of the Quantum Hall Effect [172] by Klaus von Klitzing, at the time of discovery working with a high-field magnet at LNCMI, for which he was awarded the 1985 Nobel Prize in Physics. We hope that this review has convinced the reader that other fundamental-physics discoveries will likely be forthcoming due to the remarkable progress of high-magnetic-field technologies and the abundance of ideas for new experiments. But there is also a reason for us to be humble: in spite of the impressive progress in strong-field generation, even the strongest fields we can imagine producing in macroscopic volumes in the laboratory still come many orders of magnitude short of the strongest fields found elsewhere in the universe, for example, in neutron stars where magnetic fields are expected to be as high as 10^{11} T.

18.7 Glossary

Some of the acronyms used in this review are presented in Table 35.

Table 35: Abbreviations and their meanings.

Abbreviation	Meaning
AMO	atomic, molecular and optical physics
ADMX	axion dark matter experiment (UW, Seattle)
ALPs	axion-like particles
BSCCO	bismuth-strontium-calcium-copper-oxide (HTS)
CAST	CERN axion solar telescope
CAPP	Center for Axion Precision Physics (Korea)
CULTASK	CAPP's ultralow-temperature axion search in Korea
CERN	European Organization for Nuclear Research
CHMFL	High Magnetic Field Laboratory of the Chinese Academy of Sciences
DM	dark matter
EMFC	electromagnetic flux compression
ESRF	European Synchrotron Radiation Facility
FOM	figure of merit
HLD	Hochfeld-Magnetlabor Dresden
HTS	high-temperature superconductor
IAXO	International axion observatory
ILL	Institut Laue–Langevin (Grenoble)
LHC	Large Hadron Collider at CERN
LNCMI	Laboratoire National des Champs Magnétiques Intenses
LTS	low-temperature superconductor
NHMFL	US National High Magnetic Field Laboratory
NLED	nonlinear electrodynamics
SC	superconductor, superconducting
QCD	quantum chromodynamics
QED	quantum electrodynamics
TML	Tsukuba Magnet Laboratory
YBCO	yttrium barium copper oxide (HTS)
XFEL	X-ray Free-Electron Laser

Part VII
Appendix

A Introduction

A.1 Other Motivation

(потом раскрою)

A.2 Some General Ideas of Important Discoveries in Magnetism

Как следовало понимать магнетизм, чтобы догадаться до магнитных носителей информации, которые стали одним из важнейших элементов компьютеров и другой электроники?

(важное соображение)
(????)

A.3 Literature

A.3.1 Main Literature

Basic educational literature

[1] Kittel, Charles Introduction to Solid State Physics
Очень хорошо написано введение в магнетизм, поэтому пройти сперва эту книгу стоит того.

[2] Blundell S., Magnetism in Condensed Matter;
250 pages of a good introductory theory, famous book for a typical course of magnetism. Also has interesting exercises.
(?? У Тонга по магнетизму что-то есть, кстати, мб это можно основным считать? почитаю потом)

[3] Spaldin N. Magnetic Materials: Fundamentals and Applications
Небольшая книга с простыми и обзорными темами, но и продумать темы из нее - не малое занятие. Наличие немногих задач и их решений в конце добавляет книге пользы. Также в ней имеются опечатки.
L. Paolasini Lectures on magnetism in ESRF
Хорошие лекции, очень приятно читать их, узнал немного очень интересного. Нашел в гугле, пару картинок использовал.

Books With Many Solved Problems

(пока таких чисто задачников не нашел, но потом мб добавлю их)

Статьи и книги про математический магнетизм

[4] Michel TALAGRAND Mean Field Models for Spin Glasses Volume I: Basic Examples

Куча математики и формул, можно месяцв 3 только это изучать. Очень много специфики, книга доказывает, что спиновые стекла - огромная сфера исследований.

[5] Michel Talagrand Spin Glasses: A Challenge for Mathematicians
Открыл, полистал - действительно, куча математических формул и сложностей.

Статьи про свойства магнитных материалов

[6] Zaanen, J. and Sawatzky, G. A. and Allen, J. W. Band gaps and electronic structure of transition-metal compounds
Известная в узких кругах статья с известной диаграммой.

A.3.2 Extra Literature

Другая обучающая литература

[7] Mugiraneza, Sam and Hallas, Alannah M. Tutorial: a beginner's guide to interpreting magnetic susceptibility data with the Curie-Weiss law
Статья 2022го года, где пояснено про эксперименты много интересного и не сложного. Взял из нее картинку про правила Хунда. Выглядит очень интересной, как раз про то, как считать на практике восприимчивость, можно разобраться там, наверное.

Профессиональная по магнетизму

[8] Е.С.Боровик, В.В.Еременко, А.С. Мильнер, Лекции по магнетизму, М., Физматлит, 2005.
Очень большая, подробная, хорошая книга, где много деталей и экспериментов. Не уверен в актуальности многих тем, ну и слишком она подробная для чтения не специалистам.
А.Г.Гуревич, Г.А.Мелков, Магнитные колебания и волны, М., Наука, 1994.
Рекомендована кафедрой ПТФ Ландау, значит, видимо, хорошая книга.
Д. Маттис, Теория магнетизма, М., Мир. 1967.
Рекомендована кафедрой ПТФ Ландау, значит, видимо, хорошая книга.
В.С.Львов, Нелинейные спиновые волны, М., Наука, 1987.
Рекомендована кафедрой ПТФ Ландау, значит, видимо, хорошая книга.
Р.Уайт, Квантовая теория магнетизма, М., Мир, 1985.
Рекомендована кафедрой ПТФ Ландау, значит, видимо, хорошая книга.
А. И. Ахиезер, С. В. Пелетминский, В. Г. Барьяхтар, Спиновые волны;
Рекомендована кафедрой ПТФ Ландау, значит, видимо, хорошая книга.

[9] Hal Tasaki Physics and Mathematics of Quantum Many-Body Systems 2020
Много общей теории, в том чилсе про спиновые модели, много математики, не специалистам она не самая важная для чтения. Качается по https://link.springer.com/book/10.1007/978-3-030-41265-4.

[10] Sorai, Michio and Nakazawa, Yasuhiro and Nakano, Motohiro and Miyazaki, Yuji Update 1 of: Calorimetric Investigation of Phase Transitions Occurring in Molecule-Based Magnets
Примерно 80 страниц разных графиков магнетизма разных веществ. Может быть полезен как справочник.

Литература по основам статфиза, квантмех и электродианмики для магнетизма (????)

(тоже потом многое добавлю.)

Е.М.Лифшиц, Л.П.Питаевский, Статистическая физика.Часть2. М., Физматлит, 2002.
Рекомендована кафедрой ПТФ Ландау, значит, видимо, хорошая книга.

[11] Tomonaga, S., The Story of Spin

Если Вы хотите понять, что же такое спин, который везде в магнетизме, то тут ответ. Однако, на начальных этапах разумнее представить, что вы это знаете и заниматься магнетизмом, если выбрали изучать его.

А.З.Паташинский, В.Л.Покровский, Флуктуационная теория фазовых переходов, М., Наука, 1982.

Рекомендована кафедрой ПТФ Ландау, значит, видимо, хорошая книга.

Статьи о моделировании

https://github.com/ubermag/tutorials

В этом репозитории есть примеры кодов, а также каринки, так что что-то годное их коды делают. Если бы думал про моделирование, начал изучать это бы, но это проект минимум на пару недель.

Статьи о теоретических методах

[12] Bernatska, J and Holod, P A generalized Landau–Lifshitz equation for an isotropic SU(3) magnet

Ineteresting article about applications of algebra in magnets. Personally I don't understand now how to write such models, so there is a lot to learn in this article and in refs. from it. For now I need to finish basics of magnetism and algebra to start thinking about it.

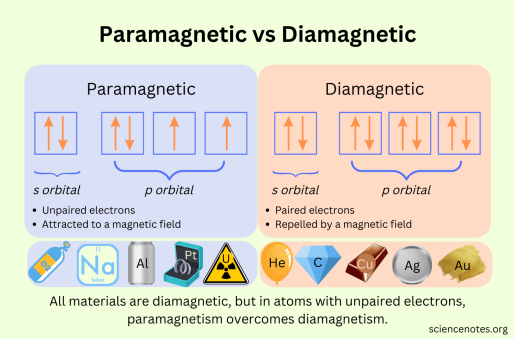
Литература о приложениях магнетизма

(про магнитные приборы, т.п.)

A.4 General Introduction and Overview

(потом раскрою)

A.4.1 Popular Introduction



(потом когда нечего будет делать - напишу)

A.4.2 General Overview

(I'll write it very later)

Связи магнетизма с другими науками

Обсудим связи с разделами
(потом раскрою)

Важность некоторых тем основ квантовой механики для магнетизма

(раскрою мысль, что иногда имеет смысл просто квантмех додумывать долго, потом магнетизм продолжить. очень много на него отсылок.)

Обзор теоретических подходов

такие-то есть, такие полезные, такие - нет.

A.5 Puzzles for Different Cases

(??? I'll write them later)

Бытовые головоломки

Принципиальные головоломки

Головоломки о деталях

A.6 History

Обсудим вкратце историю развития...

B Mathematics, Physics, Programming

(??? I'll write what is really needed, very very later)

B.1 On general physics and mathematics concepts

(всякая необходимая для понимания физика и матем тут указана будет.)

C Bibliography

References

[1] Kittel, Charles: *Introduction to Solid State Physics*. Wiley, 8th edition, 2004, ISBN 9780471415268. http://www.amazon.com/Introduction-Solid-Physics-Charles-Kittel/dp/047141526X/ref=dp_ob_title_bk.

[2] Blundell, S.: *Magnetism in Condensed Matter*. Oxford Master Series in Condensed Matter Physics. OUP Oxford, 2001, ISBN 9780198505914. https://books.google.be/books?id=OGhGmgEACAAJ.

[3] Spaldin, Nicola A.: *Magnetic Materials: Fundamentals and Applications*. Cambridge University Press, 2nd edition, 2010.

[4] Talagrand, Michel: *Mean field models for spin glasses. Volume I: Basic examples. 2nd revised and enlarged ed*. January 2011, ISBN 978-3-642-15201-6.

[5] Talagrand, Michel: *Spin glasses: A challenge for mathematicians*. Springer, 2003. https://api.semanticscholar.org/CorpusID:137383333.

[6] Zaanen, J., Sawatzky, G. A., and Allen, J. W.: *Band gaps and electronic structure of transition-metal compounds*. Phys. Rev. Lett., 55:418–421, Jul 1985. https://link.aps.org/doi/10.1103/PhysRevLett.55.418.

[7] Mugiraneza, Sam and Hallas, Alannah M.: *Tutorial: a beginner's guide to interpreting magnetic susceptibility data with the curie-weiss law*. Communications Physics, 5(1), April 2022, ISSN 2399-3650. http://dx.doi.org/10.1038/s42005-022-00853-y.

[8] Боровик Е.С., Еременко В.В., Мильнер А.С: *Лекции по магнетизму*. Физматлит, 3-е изд., перераб. и доп. edition, 2005, ISBN 5-9221-0577-9.

[9] Tasaki, Hal: *Physics and Mathematics of Quantum Many-Body Systems*. January 2020, ISBN 978-3-030-41264-7.

[10] Sorai, Michio, Nakazawa, Yasuhiro, Nakano, Motohiro, and Miyazaki, Yuji: *Update 1 of: Calorimetric investigation of phase transitions occurring in molecule-based magnets*. Chemical Reviews, 113(1):PR41–PR122, 2013. https://doi.org/10.1021/cr300156s, PMID: 23270520.

[11] Tomonaga, S.: *The Story of Spin*. A Pacific Basin Institute book. University of Chicago Press, 1997, ISBN 9780226807942. https://books.google.bg/books?id=he9sANipCj8C.

[12] Bernatska, J and Holod, P: *A generalized landau–lifshitz equation for an isotropic su(3) magnet*. Journal of Physics A: Mathematical and Theoretical, 42(7):075401, January 2009, ISSN 1751-8121. http://dx.doi.org/10.1088/1751-8113/42/7/075401.

[13] Arovas, Daniel: *Lecture notes on condensed matter physics*, 2023. https://courses.physics.ucsd.edu/2020/Spring/physics239/LECTURES/CONDMAT.pdf.

books from Critical Entropy of Quantum Heisenberg Magnets on Simple-Cubic Lattices by Stefan Wessel

References

[1] I. Bloch, J. Dalibar, and W. Zwerger, Rev. Mod. Phys. **80**, 885 (2008).

[2] M. Greiner, O. Mandel, T. Esslinger, T.W. Hänsch, and I. Bloch, Nature **415**, 39 (2002).

[3] R. Jördens, N. Strohmaier, K. Günter, H. Moritz and T. Esslinger, Nature **455**, 204 (2008).

[4] U. Schneider, L. Hackermüller, S. Will, Th. Best, I. Bloch, T. A. Costi, R. W. Helmes, D. Rasch, and A. Rosch, Science **322**, 1520 (2008).

[5] F. Werner, O. Parcollet, A. Georges, and S. R. Hassan, Phys. Rev. Lett. **95**, 056401 (2005).

[6] A. Koetsier, R. A. Duine, I. Bloch, and H. T. C. Stoof, Phys. Rev. A **77**, 023623 (2008).

[7] A.-M. Daré, L. Raymond, G. Albinet, and A.-M. S. Tremblay, Phys. Rev. B **76**, 064402 (2007).

[8] L. De Leo, C. Kollath, A. Georges, M. Ferrero, and O. Parcollet, Phys. Rev. Lett. **101**, 210403 (2008).

[9] T.-L. Ho and Q. Zhou, PNAC **106**, 6916 (2009).

[10] J. B. Bernier, C. Kollath, A. Georges, L. De Leo, F. Gerbier, C. Salomon, and M. Köhl, Phys. Rev. A **79**, 061601 (2009).

[11] C. J. M. Mathy and D. A. Huse, Phys. Rev. A **79**, 063412 (2009).

[12] T. Gottwald and P. G. J. van Dongen, Phys. Rev. A **80**, 033603 (2009).

[13] R. Staudt, M. Dzierzawa, and A. Muramatsu, Eur. Phys. J. B **17**, 411 (2000).

[14] A. Auerbach, *Interacting Electrons and Quantum Magnetism* (Springer, New York, 1994).

[15] A. W. Sandvik, Phys. Rev. Lett. **80**, 5196 (1998).

[16] M. Troyer, F. Alet, and S. Wessel, Braz. J. of Physics **34**, 377 (2004).

[17] S. Trebst, D. A. Huse, and M. Troyer, Phys. Rev. E **70**, 046701 (2004).

[18] S. Wessel, N. Stoop, E. Gull, S. Trebst, and M. Troyer, J. Stat. Mech. P12005 (2007).

[19] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).

[20] M. Troyer, S. Wessel, and F. Alet, Phys. Rev. Lett. **90**, 120201 (2003).

[21] A. W. Sandvik, Phys. Rev. B **59**, R14157 (1999).

[22] P. Henelius and A. W. Sandvik, Phys. Rev. B **62**, 1102 (2000).

[23] M. Campostrini, M. Hasenbusch, A. Pelissetto, P. Rossi, and E. Vicari, Phys. Rev. B **65**, 144520 (2002).

Refs from spin ice

References

[1] See, e.g., Milton, K. A. Rep. Progr. Phys.**69**, 1637-1712 (2006).

[2] Laughlin, R. B., Nobel Lecture: Fractional quantization, *Rev. Mod. Phys.* **71**, 863-874 (1999).

[3] Harris, M. J., Bramwell, S. T., McMorrow, D. F., Zeiske, T., and Godfrey, K. W., Geometrical frustration in the ferromagnetic pyrochlore $Ho_2Ti_2O_7$, *Phys. Rev. Lett.* **79**, 2554-2557 (1997).

[4] For a review of spin ice, see Bramwell, S. T., and Gingras, M. J. P., Spin ice state in frustrated magnetic pyrochlore materials. *Science* **294**, 1495-1501 (2001).

[5] Ramirez, A. P., Hayashi, A., Cava, R. J., Siddharthan, R. B., and Shastry, S., Zero-point entropy in 'spin ice'. *Nature* **399**, 333-336 (1999).

[6] Aoki, H., Sakakibara, T., Matsuhira, K., and Hiroi, Z., Magnetocaloric effect study on the pyrochlore spin ice compound $Dy_2Ti_2O_7$ in a [111] magnetic field. *J. Phys. Soc. Jpn.* **73**, 2851-2856 (2004).

[7] Higashinaka, R., Fukazawa, H., Deguchi, K., and Maeno, Y., Low temperature specific heat of $Dy_2Ti_2O_7$ in the kagome ice state. *J. Phys. Soc. Jpn.* **73**, 2845-2850 (2004).

[8] Cabrera, B., First results from a superconductive detector for moving magnetic monopoles. *Phys. Rev. Lett.* **48**, 1378-1381 (1982).

[9] Siddharthan, R. et al., Ising Pyrochlore Magnets: Low-Temperature Properties, Ice Rules, and Beyond, *Phys. Rev. Lett.* **83**, 1854-1857 (1999).

[10] Ruff, J. P. C., Melko, R. G, and Gingras, M. J. P., Finite-temperature transitions in dipolar spin ice in a large magnetic field. *Phys. Rev. Lett.* **95**, 097202 (2005).

[11] Huse, D. A., Krauth, W., Moessner, R., and Sondhi, S. L., Coulomb and liquid dimer models in three dimensions. *Phys. Rev. Lett.* **91**, 167004-4 (2003).

[12] Hermele, M., Fisher, M. P. A., Balents, L., Pyrochlore photons: The U(1) spin liquid in a S=1/2 three-dimensional frustrated magnet. *Phys. Rev. B* **69**, 064404-21 (2004).

[13] Henley, C. L., Power-law spin correlations in pyrochlore antiferromagnets. *Phys. Rev. B* **71**, 014424-11 (2005).

[14] Fulde, P., Penc, K., and Shannon, N., Fractional charges in pyrochlore lattices. *Ann. Phys.* **11**, 892-900 (2002).

[15] Levin, M., and Wen, X. G., Quantum ether: Photons and electrons from a rotor model, *Phys. Rev. B* **73**, 035122-10 (2006).

[16] Nagle, J. F., Theory of the dielectric constant of ice. *Chem. Phys.* **43**, 317-328 (1979).

[17] Dirac, P. A. M., Quantised singularities in the electromagnetic field. *Proc. Roy. Soc. A* **133**, 60 (1931).

[18] Jackson, J. D., Classical Electrodynamics (Wiley, New York, 1975).

[19] Saitoh, E., Miyajima, H., Yamaoka, T., and Tatara,G., Current-induced resonance and mass determination of a single magnetic domain wall. *Nature* **432**, 203-206 (2004).

[20] Rajaraman, R., Fractional charge. *http://arxiv.org/abs/cond-mat/0103366v1* (2001).

[21] Kobelev, V., Kolomeisky, A. B., and Fisher, M. E., Lattice models of ionic systems. *J. Chem. Phys.* **116**, 7589-7598 (2002).

[22] Heilmann, O. J., and, Lieb, E. H., Theory of monomer-dimer systems. *Comm. Math. Phys.* **25**, 190-232 (1972).

[23] Isakov, S. V., Raman, K. S., Moessner, R., and Sondhi, S. L., Magnetization curve of spin ice in a [111] magnetic field. *Phys. Rev. B* **70**, 104418-12 (2004).

[24] Fang, Z. *et al.*, The anomalous Hall effect and magnetic monopoles in momentum space. *Science* **302**, 92-95 (2003).

[25] Kitaev, A. Yu., Fault tolerant quantum computing. *Ann. Phys.* **303**, 2-30 (2003).

Refs from Nagaoka

References

[1] W. J. Heisenberg, Z. Phys. **49**, 619 (1928).

[2] F. Bloch, Z. Phys. **57**, 545 (1929).

[3] J. Kanamori, Prog. Theor. Phys. **30**, 275 (1963).

[4] M. C. Gutzwiller, Phy. Rev. Lett. **10**, 159 (1963).

[5] J. Hubbard, Proc. Roy. Soc. (London) **A276**, 238 (1963).

[6] J. C. Slater, H. Statz, and G. F. Koster, Phy. Rev. **91**, 1323 (1953).

[7] H. Tasaki, Phy. Rev. Lett. **69**, 1608 (1992).

[8] A. Mielke and H. Tasaki, Commun. Math Phys. **158**, 341 (1993), cond-mat/9305026.

[9] D. J. Thouless, Proc. Phys. Soc. London **86**, 893 (1965).

[10] Y. Nagaoka, Phy. Rev. **147**, 392 (1966).

[11] E. H. Lieb, Phy. Rev. Lett. **62**, 1201 (1989).

[12] S. Q. Shen, Z. M. Qiu, and G. S. Tian, Phy. Rev. Lett. **72**, 1280 (1994).

[13] A. Mielke, J. Phys. **A24**, L73 (1991).

[14] A. Mielke, J. Phys. **A24**, 3311 (1991).

[15] A. Mielke, J. Phys. **A25**, 4335 (1992).

[16] H. Tasaki, Phy. Rev. Lett. **73**, 1158 (1994).

[17] H. Tasaki, J. Stat. Phys. **84**, 535 (1996).

[18] H. Tasaki, Phy. Rev. Lett. **75**, 4678 (1995), cond-mat/9509063.

[19] H. Tasaki, in prepration (1998).

[20] E. H. Lieb, in *Advances in Dynamical Systems and Quantum Physics* (World Scientific, 1995), cond-mat/9311033.

[21] H. Tasaki, J. Phys. Cond. Matt. to appear (1998), cond-mat/9512169.

[22] P. Fazekas, Philosophical Magazine B to appear (1997), cond-mat/9612090.

[23] D. Vollhardt, N. Blümer, K. Held, M. Kollar, J. Schlipf, and M. Ulmke, Z. Phys. B **103**, 283 (1997), cond-mat/9701150.

[24] T. Hanisch, G. S. Uhrig, and E. Mueller-Hartmann, preprint (1997), cond-mat/9707286.

[25] S.-Q. Shen, Int. J. Mod. Phys. B to appear (1997).

[26] P. Pieri, S. Daul, D. Baeriswyl, M. Dzierzawa, and P. Fazekas, Phys. Rev. B **45**, 9250 (1996), cond-mat/9603163.

[27] B. S. Shastry, H. R. Krishnamurthy, and P. W. Anderson, Phy. Rev. **B41**, 2375 (1990).

[28] L. M. Roth, J. Phys. Chem. Solids **28**, 1549 (1967).

[29] E. H. Lieb and D. Mattis, Phy. Rev. **125**, 164 (1962).

[30] H. Tasaki, Phy. Rev. **B40**, 9192 (1989).

[31] E. H. Lieb, in *Phase Transitions* (Wiley, Interscience, 1971), p. 45.

[32] U. Brandt and A. Giesekus, Phy. Rev. Lett. **68**, 2648 (1992).

[33] H. Tasaki, Phy. Rev. **B49**, 7763 (1994).

[34] B. Simon, *The statistical mechanics of lattice gases. Vol. 1* (Princeton University Press, Princeton, New Jersey, 1993).

[35] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics (Non-relativisitic Theory)* (Pergamon Press, 1977).

[36] B. Douçot and X. G. Wen, Phy. Rev. **B40**, 2719 (1989).

[37] B. Tóth, Lett. Math. Phys. **22**, 321 (1991).

[38] A. Sütő, Commun. Math Phys. **140**, 43 (1991).

[39] T. Hanisch and E. Müller-Hartmann, Ann. Physik **2**, 381 (1993).

[40] S. A. Trugman, Phys. Rev. B **42**, 6612 (1990).

[41] G. S. Tian, Phys. Rev. B **44**, 4444 (1991).

[42] A. Barbieri, J. A. Riera, and A. P. Young, Phys. Rev. B **41**, 11697 (1990).

[43] W. O. Putikka, M. U. Luchini, and M. Ogata, Phy. Rev. Lett. **69**, 2288 (1992).

[44] S. Liang and H. Pang, Europhys. Lett. **32**, 173 (1995), cond-mat/9404067.

[45] K. Kusakabe and H. Aoki, Phys. Rev. B **50**, 12991 (1994), cond-mat/9407036.

[46] M. Yamanaka, S. Honjo, Y. Hatsugai, and M. Kohmoto, J. Stat. Phys. **84**, 1133 (1996), cond-mat/9512065.

[47] A. Mielke, Phys. Lett. **A174**, 443 (1993).

[48] K. Kusakabe and H. Aoki, Physica **B194–B196**, 215 (1994).

[49] K. Kusakabe and H. Aoki, Phy. Rev. Lett. **72**, 144 (1994).

[50] K. Penc, H. Shiba, F. Mila, and T. Tsukagoshi, Phys. Rev. B **54**, 4056 (1996), cond-mat/9603042.

[51] F. Mizuno, H. Masuda, I. Hirabayashi, S. Tanaka, M. Hasegawa, and U. Mizutani, Nature **345**, 788 (1990).

[52] H. Masuda, F. Mizunoand, I. Hirabayashi, and S. Tanaka, Physical Review B **43**, 7871 (1991).

[53] F. Mizuno, H. Masuda, and I. Hirabayashi, in *Studies of High Temperature Superconductors, vol. 10*, edited by A. Narlikar (Nova Science Publisher, Commack, NY, 1993).

[54] S. Feldkemper, W. Weber, J. Schulenburg, and J. Richter, Phys. Rev. B **52**, 313 (1995).

[55] I. V. Paukov, M. N. Popova, and B. V. Mill', Phys. Lett. A **157**, 306 (1991).

[56] I. V. Golosovsky, P. Böni, and P. Fischer, Phys. Lett. A **182**, 161 (1993).

[57] A. Salinas-Sánchez and R. Sáez-Puche, Solid State Ionics **63-65**, 927 (1993).

[58] M. W. Pieper, E.-G. Caspary, P. Adelmann, and G. Roth, J. Mag. Mag. Mat. **153**, 319 (1994).

[59] V. Eyert, K.-H. Höck, and P. S. Riseborough, Europhys. Lett. **31**, 385 (1995).

[60] I. Hirabayashi, private communication.

[61] S. Watanabe, M. Ichimura, T. Onogi, Y. A. Ono, T. Hashizume, and Y. Wada, Jpn. J. Appl. Phys. **36**, L929 (1997).

[62] M. Ichimura, K. Kusakabe, S. Watanabe, and T. Onogi, in preparation (1998).

[63] H. Sakamoto and K. Kubo, J. Phys. Soc. Jpn. **65**, 3732 (1996).

[64] Y. Watanabe and S. Miyashita, J. Phys. Soc. Jpn. **66**, 2123 (1997).

[65] Y. Watanabe and S. Miyashita, J. Phys. Soc. Jpn. **66**, to appear (1997).

[66] M. W. Long, C. W. M. Castleton, and C. A. Hayward, J. Phys. Condens. Matter **6**, 481 (1994).

[67] K. Ueda, T. Nishino, and H. Tsunetsugu, Phys. Rev. B **50**, 612 (1994).

[68] E. Müller-Hartmann, J. Low Temp. Phys. **99**, 349 (1995), cond-mat/9502104.

[69] S. Daul and R. Noack, preprint (1996), cond-mat/9612056.

[70] M. Kohno, Phys. Rev. B to appear (1997), cond-mat/9709170.

[71] T. Kennedy and H. Tasaki, Commun. Math Phys. **147**, 431 (1992).

[72] K. Kubo, J. Phys. Soc. Jpn. **51**, 782 (1982).

[73] K. Kusakabe and H. Aoki, Phys. Rev. B **52**, R8684 (1995), cond-mat/9509034.

[74] R. Arita, Y. Shimoi, K. Kuroki, and H. Aoki, preprint (1997), cond-mat/9801242.

[75] O. Bratteli and D. W. Robinson, *Operator algebras and quantum statistical mechanics. Vol. 2* (Springer-Verlag, New York, Heidelberg, Berlin, 1997).

From High magnetic fields for fundamental physics

References

[1] National Research Council, High Magnetic Field Science and Its Application in the United States: Current Status and Future Directions, National Academies Press, 2013.

[2] W. Heisenberg, H. Euler, Consequences of Dirac's theory of positrons, Z. Phys. 98 (1936) 714–732.

[3] V. Weisskopf, The electrodynamics of the vacuum based on the quantum theory of the electron, Kong. Dan. Vid. Sel. Mat. Fys. Med. 14N6 (1936) 1–39.

[4] J. Schwinger, On gauge invariance and vacuum polarization, Phys. Rev. 82 (1951) 664–679.

[5] H. Euler, B. Kockel, Über die streuung von licht an licht nach der diracschen theorie, Naturwissenschaften 23 (1935) 246–247.

[6] R. Karplus, M. Neuman, The scattering of light by light, Phys. Rev. 83 (1951) 776–784.

[7] R. Karplus, M. Neuman, Non-Linear Interactions between Electromagnetic Fields, Phys. Rev. 80 (1950) 380–385.

[8] J. S. Toll, The Dispersion relation for light and its application to problems involving electron pairs, Ph.D. thesis, Princeton U., 1952.

[9] R. Baier, P. Breitenlohner, The vacuum refraction index in the presence of external fields, Il Nuovo Cimento B (1965-1970) 47 (1967) 117–120.

[10] R. Baier, P. Breitenlohner, Photon propagation in external fields, Acta Physica Austria 25 (1967) 212–223.

[11] M. Fouché, R. Battesti, C. Rizzo, Limits on nonlinear electrodynamics, Phys. Rev. D 93 (2016) 093020.

[12] R. Battesti, C. Rizzo, Magnetic and electric properties of a quantum vacuum, Reports on Progress in Physics 76 (2013) 016401.

[13] M. Born, L. Infeld, Foundations of the new field theory, Proc. Roy. Soc. Lond. A144 (1934) 425–451.

[14] K. Scharnhorst, Photon-photon scattering and related phenomena. Experimental and theoretical approaches: The early period (2017).

[15] W. Dittrich, M. Reuter, EFFECTIVE LAGRANGIANS IN QUANTUM ELECTRODYNAMICS, Lect. Notes Phys. 220 (1985) 1–244.

[16] W. Dittrich, H. Gies, Probing the quantum vacuum. Perturbative effective action approach in quantum electrodynamics and its application, Springer Tracts Mod. Phys. 166 (2000) 1–241.

[17] H. Gies, F. Karbstein, An Addendum to the Heisenberg-Euler effective action beyond one loop, JHEP 03 (2017) 108.

[18] R. P. Mignani, V. Testa, D. G. Caniulef, R. Taverna, R. Turolla, S. Zane, K. Wu, Evidence for vacuum birefringence from the first optical-polarimetry measurement of the isolated neutron star RX J1856.5-3754, Mon. Not. Roy. Astron. Soc. 465 (2017) 492–500.

[19] L. M. Capparelli, A. Damiano, L. Maiani, A. D. Polosa, A note on polarized light from Magnetars, Eur. Phys. J. C77 (2017) 754.

[20] E. Iacopini, E. Zavattini, Experimental method to detect the vacuum birefringence induced by a magnetic field, Physics Letters B 85 (1979) 151 – 154.

[21] R. Cameron, et al., Search for nearly massless, weakly coupled particles by optical techniques, Phys. Rev. D47 (1993) 3707–3725.

[22] E. Zavattini, G. Zavattini, G. Ruoso, G. Raiteri, E. Polacco, E. Milotti, V. Lozza, M. Karuza, U. Gastaldi, G. Di Domenico, F. Della Valle, R. Cimino, S. Carusotto, G. Cantatore, M. Bregant (PVLAS Collaboration)), New pvlas results and limits on magnetically induced optical rotation and ellipticity in vacuum, Phys. Rev. D 77 (2008) 032006.

[23] F. D. Valle, U. Gastaldi, G. Messineo, E. Milotti, R. Pengo, L. Piemontese, G. Ruoso, G. Zavattini, Measurements of vacuum magnetic birefringence using permanent dipole magnets: the pvlas experiment, New Journal of Physics 15 (2013) 053026.

[24] A. Cadène, P. Berceau, M. Fouché, R. Battesti, C. Rizzo, Vacuum magnetic linear birefringence using pulsed fields: status of the BMV experiment, The European Physical Journal D 68 (2014) 16.

[25] A. Ejlli, PROGRESS TOWARDS A FIRST MEASUREMENT OF THE MAGNETIC BIREFRINGENCE OF VACUUM WITH A POLARIMETER BASED ON A FABRY-PEROT CAVITY, Ph.D. thesis, Dipartimento di Fisica e Scienze della Terra Università di Ferrara, Via Saragat 1, Blocco C 44122, Ferrara, Italy, 2017.

[26] F. Della Valle, A. Ejlli, U. Gastaldi, et al., The PVLAS experiment: measuring vacuum magnetic birefringence and dichroism with a birefringent fabry‚äìperot cavity, Eur. Phys. J. C 76 (2016) 24. Doi:10.1140/epjc/s10052-015-3869-8.

[27] X. Fan, S. Kamioka, T. Inada, T. Yamazaki, T. Namba, S. Asai, J. Omachi, K. Yoshioka, M. Kuwata-Gonokami, A. Matsuo, K. Kawaguchi, K. Kindo, H. Nojiri, The oval experiment: a new experiment to measure vacuum magnetic birefringence using high repetition pulsed magnets, The European Physical Journal D 71 (2017) 308.

[28] P. Pugnat, M. Král, A. Siemko, L. Duvillaret, M. Finger, J. Zicha, Feasibility study of an experiment to measure the vacuum magnetic birefringence, Czechoslovak Journal of Physics 55 (2005) A389–A397.

[29] R. Ballou, G. Deferne, M. Finger, M. Finger, L. Flekova, J. Hosek, S. Kunc, K. Macuchova, K. A. Meissner, P. Pugnat, M. Schott, A. Siemko, M. Slunecka, M. Sulc, C. Weinsheimer, J. Zicha (OSQAR Collaboration), New exclusion limits on scalar and pseudoscalar axionlike particles from light shining through a wall, Phys. Rev. D 92 (2015) 092002.

[30] M. T. Hartman, A. Rivère, R. Battesti, C. Rizzo, Noise characterization for resonantly enhanced polarimetric vacuum magnetic-birefringence experiments, Review of Scientific Instruments 88 (2017) 123114.

[31] S. Batut, J. Mauchain, R. Battesti, et al., A Transportable Pulsed Magnet System for Fundamental Investigations in Quantum Electrodynamics and Particle Physics, IEEE Trans. Appl. Supercond. 8 (2008) 600. Doi:10.1109/TASC.2008.922518.

[32] T. Yamazaki, T. Inada, T. Namba, et al., Repeating pulsed magnet system for axion-like particle searches and vacuum birefringence experiments, Nucl. Instrum. Methods Phys. Res., Sect. A 833 (2016) 122. Doi:10.1016/j.nima.2016.07.031.

[33] G. Zavattini, F. Della Valle, A. Ejlli, G. Ruoso, A polarisation modulation scheme for measuring vacuum magnetic birefringence with static fields, The European Physical Journal C 76 (2016) 294.

[34] G. Zavattini, F. Della Valle, A. Ejlli, G. Ruoso, Erratum to: A polarisation modulation scheme for measuring vacuum magnetic birefringence with static fields, The European Physical Journal C 77 (2017) 873.

[35] B. Dobrich, H. Gies, Interferometry of light propagation in pulsed fields, EPL 87 (2009) 21002.

[36] G. Zavattini, E. Calloni, Probing for new physics and detecting non-linear vacuum qed effects using gravitational wave interferometer antennas, The European Physical Journal C 62 (2009) 459–466.

[37] H. Grote, On the possibility of vacuum qed measurements with gravitational wave detectors, Phys. Rev. D 91 (2015) 022002.

[38] A. Di Piazza, C. Müller, K. Hatsagortsyan, C. Keitel, Extremely high-intensity laser interactions with fundamental quantum systems, Reviews of Modern Physics 84 (2012) 1177.

[39] T. Heinzl, B. Liesfeld, K.-U. Amthor, H. Schwoerer, R. Sauerbrey, A. Wipf, On the observation of vacuum birefringence, Opt. Commun. 267 (2006) 318–321.

[40] B. Marx, I. Uschmann, S. Höfer, R. Lötzsch, O. Wehrhan, E. Förster, M. Kaluza, T. Stöhlker, H. Gies, C. Detlefs, T. Roth, J. Härtwig, G. G. Paulus, Determination of high-purity polarization state of X-rays, Optics Communications 284 (2011) 915–918.

[41] B. Marx, K. S. Schulze, I. Uschmann, T. Kämpfer, R. Lötzsch, O. Wehrhan, W. Wagner, C. Detlefs, T. Roth, J. Härtwig, E. Förster, T. Stöhlker, G. G. Paulus, High-precision x-ray polarimetry, Phys. Rev. Lett. 110 (2013) 254801.

[42] F. Karbstein, H. Gies, M. Reuter, M. Zepf, Vacuum birefringence in strong inhomogeneous electromagnetic fields, Phys. Rev. D92 (2015) 071301.

[43] H.-P. Hans-Peter Schlenvoigt, T. Heinzl, U. Schramm, T. E. Cowan, R. Sauerbrey, Detecting vacuum birefringence with x-ray free electron lasers and high-power optical lasers: a feasibility study, Physica Scripta 91 (2016) 023010.

[44] F. Karbstein, C. Sundqvist, Probing vacuum birefringence using x-ray free electron and optical high-intensity lasers, Phys. Rev. D94 (2016) 013004.

[45] A. Di Piazza, K. Z. Hatsagortsyan, C. H. Keitel, Light diffraction by a strong standing electromagnetic wave, Phys. Rev. Lett. 97 (2006) 083603.

[46] V. Dinu, T. Heinzl, A. Ilderton, M. Marklund, G. Torgrimsson, Vacuum refractive indices and helicity flip in strong-field QED, Phys. Rev. D89 (2014) 125003.

[47] V. Dinu, T. Heinzl, A. Ilderton, M. Marklund, G. Torgrimsson, Photon polarization in light-by-light scattering: Finite size effects, Phys. Rev. D90 (2014) 045025.

[48] G. L. Kotkin, V. G. Serbo, Variation in polarization of high-energy gamma quanta traversing a bunch of polarized laser photons, Phys. Lett. B413 (1997) 122–129.

[49] Y. Nakamiya, K. Homma, T. Moritaka, K. Seto, Probing vacuum birefringence under a high-intensity laser field with gamma-ray polarimetry at the GeV scale, Phys. Rev. D96 (2017) 053002.

[50] A. Ilderton, M. Marklund, Prospects for studying vacuum polarisation using dipole and synchrotron radiation, J. Plasma Phys. 82 (2016) 655820201.

[51] B. King, N. Elkina, Vacuum birefringence in high-energy laser-electron collisions, Phys. Rev. A94 (2016) 062102.

[52] S. Bragin, S. Meuren, C. H. Keitel, A. Di Piazza, High-Energy Vacuum Birefringence and Dichroism in an Ultrastrong Laser Field, Phys. Rev. Lett. 119 (2017) 250403.

[53] M. Marklund, J. Lundin, Quantum Vacuum Experiments Using High Intensity Lasers, Eur. Phys. J. D55 (2009) 319–326.

[54] G. V. Dunne, New Strong-Field QED Effects at ELI: Nonperturbative Vacuum Pair Production, Eur. Phys. J. D55 (2009) 327–340.

[55] T. Heinzl, A. Ilderton, Exploring high-intensity QED at ELI, Eur. Phys. J. D55 (2009) 359–364.

[56] A. Di Piazza, C. Muller, K. Z. Hatsagortsyan, C. H. Keitel, Extremely high-intensity laser interactions with fundamental quantum systems, Rev. Mod. Phys. 84 (2012) 1177.

[57] B. King, T. Heinzl, Measuring vacuum polarization with high-power lasers, High Power Laser Science and Engineering 4 (2016) e5.

[58] F. Karbstein, The quantum vacuum in electromagnetic fields: From the Heisenberg-Euler effective action to vacuum birefringence, in: Proceedings, Quantum Field Theory at the Limits: from Strong Fields to Heavy Quarks (HQ 2016): Dubna, Russia, July 18-30, 2016, 2017, pp. 44–57. URL: https://inspirehep.net/record/1501006/files/arXiv:1611.09883.pdf. arXiv:1611.09883.

[59] T. Inada, T. Yamazaki, T. Yamaji, Y. Seino, X. Fan, S. Kamioka, T. Namba, S. Asai, Probing Physics in Vacuum Using an X-ray Free-Electron Laser, a High-Power Laser, and a High-Field Magnet, Applied Sciences 7 (2017) 671.

[60] P. Arias, D. Cadamuro, M. Goodsell, J. Jaeckel, J. Redondo, A. Ringwald, Wispy cold dark matter, Journal of Cosmology and Astroparticle Physics 2012 (2012) 013.

[61] K. Baker, et al., The quest for axions and other new light particles, Annalen Phys. 525 (2013) A93–A99.

[62] R. Essig, et al., Working Group Report: New Light Weakly Coupled Particles, in: Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013, 2013. URL: http://inspirehep.net/record/1263039/files/arXiv:1311.0029.pdf. arXiv:1311.0029.

[63] L. Marzola, M. Raidal, F. R. Urban, Oscillating spin-2 dark matter, Phys. Rev. D 97 (2018) 024010.

[64] K. Van Bibber, N. R. Dagdeviren, S. E. Koonin, A. K. Kerman, H. N. Nelson, Proposed experiment to produce and detect light pseudoscalars, Phys. Rev. Lett. 59 (1987) 759–762.

[65] J. Redondo, A. Ringwald, Light shining through walls, Contemporary Physics 52 (2011) 211–236.

[66] K. Ehret, et al., New ALPS Results on Hidden-Sector Lightweights, Phys. Lett. B689 (2010) 149–155.

[67] C. Robilliard, R. Battesti, M. Fouche, J. Mauchain, A.-M. Sautivet, F. Amiranoff, C. Rizzo, No light shining through a wall, Phys. Rev. Lett. 99 (2007) 190403.

[68] A. S. Chou, W. C. Wester, III, A. Baumbaugh, H. R. Gustafson, Y. Irizarry-Valle, P. O. Mazur, J. H. Steffen, R. Tomlin, X. Yang, J. Yoo (GammeV (T-969)), Search for axion-like particles using a variable baseline photon regeneration technique, Phys. Rev. Lett. 100 (2008) 080402.

[69] R. Ballou, et al. (OSQAR), New exclusion limits on scalar and pseudoscalar axionlike particles from light shining through a wall, Phys. Rev. D92 (2015) 092002.

[70] R. Bähre, et al., Any light particle search II —Technical Design Report, JINST 8 (2013) T09001.

[71] R. Battesti, M. Fouche, C. Detlefs, T. Roth, P. Berceau, F. Duc, P. Frings, G. L. J. A. Rikken, C. Rizzo, A Photon Regeneration Experiment for Axionlike Particle Search using X-rays, Phys. Rev. Lett. 105 (2010) 250405.

[72] T. Inada, et al., Search for Two-Photon Interaction with Axionlike Particles Using High-Repetition Pulsed Magnets and Synchrotron X Rays, Phys. Rev. Lett. 118 (2017) 071803.

[73] CAST Collaboration, New cast limit on the axion-photon interaction, Nat Phys 13 (2017) 584–590.

[74] E. Armengaud, F. T. Avignone, M. Betz, P. Brax, P. Brun, G. Cantatore, J. M. Carmona, G. P. Carosi, F. Caspers, S. Caspi, S. A. Cetin, D. Chelouche, F. E. Christensen, A. Dael, T. Dafni, M. Davenport, A. V. Derbin, K. Desch, A. Diago, B. Döbrich, I. Dratchnev,

A. Dudarev, C. Eleftheriadis, G. Fanourakis, E. Ferrer-Ribas, J. Galán, J. A. García, J. G. Garza, T. Geralis, B. Gimeno, I. Giomataris, S. Gninenko, H. Gómez, D. González-Díaz, E. Guendelman, C. J. Hailey, T. Hiramatsu, D. H. H. Hoffmann, D. Horns, F. J. Iguaz, I. G. Irastorza, J. Isern, K. Imai, A. C. Jakobsen, J. Jaeckel, K. Jakovčić, J. Kaminski, M. Kawasaki, M. Karuza, M. Krčmar, K. Kousouris, C. Krieger, B. Lakić, O. Limousin, A. Lindner, A. Liolios, G. Luzón, S. Matsuki, V. N. Muratova, C. Nones, I. Ortega, T. Papaevangelou, M. J. Pivovaroff, G. Raffelt, J. Redondo, A. Ringwald, S. Russenschuck, J. Ruz, K. Saikawa, I. Savvidis, T. Sekiguchi, Y. K. Semertzidis, I. Shilon, P. Sikivie, H. Silva, H. ten Kate, A. Tomas, S. Troitsky, T. Vafeiadis, K. van Bibber, P. Vedrine, J. A. Villar, J. K. Vogel, L. Walckiers, A. Weltman, W. Wester, S. C. Yildiz, K. Zioutas, Conceptual design of the international axion observatory (IAXO), Journal of Instrumentation 9 (2014) T05002.

[75] I. G. Irastorza, et al., Gaseous time projection chambers for rare event detection: Results from the T-REX project. II. Dark matter, JCAP 1601 (2016) 034. [Erratum: JCAP1605,no.05,E01(2016)].

[76] I. Shilon, A. Dudarev, H. Silva, H. Kate, Conceptual Design of a New Large Superconducting Toroid for IAXO, the New International AXion Observatory, IEEE Trans. Appl. Supercond. 23 (2012).

[77] H. H. J. ten Kate, The ATLAS superconducting magnet system at the Large Hadron Collider, Physica C 468 (2008) 2137–2142.

[78] H. H. J. ten Kate, ATLAS Magnet System Nearing Completion, IEEE Trans. Appl. Supercond. 18 (2008) 352–355.

[79] D. M. Lazarus, et al., A search for solar axions, Phys. Rev. Lett. 69 (1992) 2333–2336.

[80] S. Moriyama, et al., Direct search for solar axions by using strong magnetic field and X-ray detectors, Phys. Lett. B434 (1998) 147.

[81] Y. Inoue, et al., Search for solar axions with mass around 1 eV using coherent conversion of axions into photons, Phys. Lett. B668 (2008) 93–97.

[82] K. Zioutas, et al., A decommissioned LHC model magnet as an axion telescope, Nucl. Instrum. Meth. A425 (1999) 480–489.

[83] K. Zioutas, et al. (CAST), First results from the CERN Axion Solar Telescope (CAST), Phys. Rev. Lett. 94 (2005) 121301.

[84] E. Arik, et al. (CAST), Probing eV-scale axions with CAST, JCAP 0902 (2009) 008.

[85] E. Arik, et al. (CAST), Search for sub-eV mass solar axions by the cern axion solar telescope with ^{3}He buffer gas, Phys. Rev. Lett. 107 (2011) 261302.

[86] M. Arik, S. Aune, K. Barth, A. Belov, S. Borghi, H. Bräuninger, G. Cantatore, J. M. Carmona, S. A. Cetin, J. I. Collar, E. Da Riva, T. Dafni, M. Davenport, C. Eleftheriadis, N. Elias, G. Fanourakis, E. Ferrer-Ribas, P. Friedrich, J. Galán, J. A. García, A. Gardikiotis, J. G. Garza, E. N. Gazis, T. Geralis, E. Georgiopoulou, I. Giomataris, S. Gninenko, H. Gómez, M. Gómez Marzoa, E. Gruber, T. Guthörl, R. Hartmann, S. Hauf, F. Haug, M. D. Hasinoff, D. H. H. Hoffmann, F. J. Iguaz, I. G. Irastorza, J. Jacoby, K. Jakovčić, M. Karuza, K. Königsmann, R. Kotthaus, M. Krčmar, M. Kuster, B. Lakić, P. M. Lang,

J. M. Laurent, A. Liolios, A. Ljubičić, G. Luzón, S. Neff, T. Niinikoski, A. Nordt, T. Papaevangelou, M. J. Pivovaroff, G. Raffelt, H. Riege, A. Rodríguez, M. Rosu, J. Ruz, I. Savvidis, I. Shilon, P. S. Silva, S. K. Solanki, L. Stewart, A. Tomás, M. Tsagri, K. van Bibber, T. Vafeiadis, J. Villar, J. K. Vogel, S. C. Yildiz, K. Zioutas (CAST Collaboration), Search for solar axions by the cern axion solar telescope with ^{3}He buffer gas: Closing the hot dark matter gap, Phys. Rev. Lett. 112 (2014) 091302.

[87] I. G. Irastorza, Talk at the Physics Beyond Colliders working group meeting at CERN, 2017.

[88] I. G. Irastorza, F. Avignone, S. Caspi, J. Carmona, T. Dafni, et al., Towards a new generation axion helioscope, JCAP 1106 (2011) 013.

[89] P. Sikivie, Experimental tests of the "invisible" axion, Phys. Rev. Lett. 51 (1983) 1415–1417.

[90] P. Sikivie, Detection rates for "invisible"-axion searches, Phys. Rev. D 32 (1985) 2988–2991.

[91] R. H. Dicke, Measurement of thermal radiation at ultra-high frequencies, Rev. Sci. Instrum. 17 (1946) 268—275.

[92] S. A. Kenany, M. Anil, K. Backes, B. Brubaker, S. Cahn, G. Carosi, Y. Gurevich, W. Kindel, S. Lamoreaux, K. Lehnert, S. Lewis, M. Malnou, D. Palken, N. Rapidis, J. Root, M. Simanovskaia, T. Shokair, I. Urdinaran, K. van Bibber, L. Zhong, Design and operational experience of a microwave cavity axion detector for the 20–100 μeV range, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 854 (2017) 11 – 24.

[93] B. M. Brubaker, L. Zhong, Y. V. Gurevich, S. B. Cahn, S. K. Lamoreaux, M. Simanovskaia, J. R. Root, S. M. Lewis, S. Al Kenany, K. M. Backes, I. Urdinaran, N. M. Rapidis, T. M. Shokair, K. A. van Bibber, D. A. Palken, M. Malnou, W. F. Kindel, M. A. Anil, K. W. Lehnert, G. Carosi, First results from a microwave cavity axion search at 24 μeV, Phys. Rev. Lett. 118 (2017) 061302.

[94] G. Rybka, Laboratory searches for qcd axion dark matter, Journal of Physics G: Nuclear and Particle Physics 44 (2017) 124002.

[95] B. T. McAllister, G. Flower, E. N. Ivanov, M. Goryachev, J. Bourhill, M. E. Tobar, The organ experiment: An axion haloscope above 15 ghz, Physics of the Dark Universe 18 (2017) 67 – 72.

[96] Petrakou, Eleni for CAPP/IBS, Haloscope searches for dark matter axions at the center for axion and precision physics research, EPJ Web Conf. 164 (2017) 01012.

[97] P. Pugnat, R. Ballou, P. Camus, F. Caspers, B. R. Ko, N. Roch, Y. K. Semertzidis, Preliminary study for a new axion dark-matter haloscope, in: Y. Kim, A. Lindner, Y. K. Semertzidis (Eds.), 12th "Patras" Workshop on Axions, WIMPs and WISPs, PATRAS 2016, DESY-PROC-2016-03, Verlag Deutsches Elektronen-Synchrotron Hamburg, 2017, pp. 129–132.

[98] O. K. Baker, M. Betz, F. Caspers, J. Jaeckel, A. Lindner, A. Ringwald, Y. Semertzidis, P. Sikivie, K. Zioutas, Prospects for searching axionlike particle dark matter with dipole, toroidal, and wiggler magnets, Phys. Rev. D 85 (2012) 035018.

[99] J. Choi, H. Themann, M. J. Lee, B. R. Ko, Y. K. Semertzidis, First axion dark matter search with toroidal geometry, Phys. Rev. D 96 (2017) 061102.

[100] J.-C. Vallet, Why not include tokamaks in the quest of exotic particles?, in: 7th Patras Workshop on Axions, WIMPs and WISPs (PATRAS 2011), 2012.

[101] S. Chaudhuri, P. W. Graham, K. Irwin, J. Mardon, S. Rajendran, Y. Zhao, Radio for hidden-photon dark matter detection, Phys. Rev. D 92 (2015) 075012.

[102] Y. Kahn, B. R. Safdi, J. Thaler, Broadband and resonant approaches to axion dark matter detection, Phys. Rev. Lett. 117 (2016) 141801.

[103] P. W. Graham, S. Rajendran, New observables for direct detection of axion dark matter, Phys. Rev. D 88 (2013) 035023.

[104] D. Budker, P. W. Graham, M. Ledbetter, S. Rajendran, A. O. Sushkov, Proposal for a cosmic axion spin precession experiment (CASPEr), Physical Review X 4 (2014) 021030.

[105] A. Garcon, D. Aybas, J. W. Blanchard, G. Centers, N. L. Figueroa, P. W. Graham, D. F. J. Kimball, S. Rajendran, M. G. Sendra, A. O. Sushkov, L. Trahms, T. Wang, A. Wickenbrock, T. Wu, D. Budker, The cosmic axion spin precession experiment (CASPEr): a dark-matter search with nuclear magnetic resonance, Quantum Science and Technology 3 (2018) 014008.

[106] R. Barbieri, C. Braggio, G. Carugno, C. Gallo, A. Lombardi, A. Ortolan, R. Pengo, G. Ruoso, C. Speake, Searching for galactic axions through magnetized media: The QUAX proposal, Physics of the Dark Universe 15 (2017) 135 – 141.

[107] J. Angrik, T. Armbrust, A. Beglarian, U. Besserer, J. Blumer, J. Bonn, R. Carr, B. Bornschein, L. Bornschein, T. Burritt, M. Charlton, A. Davies, C. Day, P. Doe, L. Dorr, O. Dragoun, G. Drexlin, F. Eichelhardt, K. Eitel, K. Essig, A. Felden, B. Flatt, J. Formaggio, B. Freudiger, R. Gehring, H. Gemmeke, E. Geraskin, F. Gluck, M. Glugla, S. Grohmann, R. Gumbsheimer, F. Habermehl, G. Harper, G. Hellriegel, J. Herbert, T. Hohn, M. Howe, H. Hucker, O. Ivanov, K. Jungst, J. Kaspar, O. Kazachenko, N. Kernert, P. Komarek, A. Kovalik, C. Kraus, H. Krause, M. Leber, R. Lewis, V. Lobashev, X. Luo, K. Maier, O. Malyshev, M. Mark, I. Meshkov, K. Muller, S. Mutterer, G. Myneni, M. Neuberger, H. Neumann, M. Noe, H. Ortjohann, S. Osipov, A. Osipowicz, B. Ostrick, E. Otten, P. Plischke, A. Povtschinik, M. Prall, R. Reid, K. Rielage, R. Robertson, M. Rysavy, P. Schafer, K. Schlosser, U. Schmitt, F. Schwamm, P. Schwinzer, D. S. Sanchez, F. Sharipov, A. Skasyrskaya, A. Spalek, M. Steidl, M. Susser, E. Syresin, H. Telle, T. Thümmler, N. Titov, V. Usanov, K. Valerius, T. V. Wechel, D. Venos, R. Vianden, C. Weinheimer, A. Weiss, J. Wendel, J. Wilkerson, J. Wolf, S. Wustling, S. Zadoroghny, M. Zboril (KATRIN), KATRIN design report 2004, FZKA Scientific Report 7090 (2005).

[108] B. Monreal, J. A. Formaggio, Relativistic cyclotron radiation detection of tritium decay electrons as a new technique for measuring the neutrino mass, Phys. Rev. D 80 (2009) 051301.

[109] A. A. Esfahani, D. Asner, S. Böser, R. Cervantes, C. Claessens, P. Doe, S. Doeleman, J. Fernandez, M. Fertl, E. Finn, J. Formaggio, D. Furse, M. Guigue, K. Heeger, A. M. Jones, K. Kazkaz, J. Kofron, C. Lamb, B. LaRoque, E. Machado, E. McBride, M. Miller, B. Monreal, P. Mohanmurthy, J. Nikkel, N. Oblath, W. Pettus, H. Robertson, L. Rosenberg, G. Rybka, D. Rysewyk, L. Saldaña, P. Slocum, M. Sternberg, J. Tedeschi,

T. Thuemmler, B. VanDevender, L. Vertatschitsch, L. de Viveiros, M. Wachtendonk, J. Weintroub, N. Woods, A. Young, E. Zayas, Determining the neutrino mass with cyclotron radiation emission spectroscopy - project 8, Journal of Physics G: Nuclear and Particle Physics (2017).

[110] D. M. Asner, R. F. Bradley, L. de Viveiros, P. J. Doe, J. L. Fernandes, M. Fertl, E. C. Finn, J. A. Formaggio, D. Furse, A. M. Jones, J. N. Kofron, B. H. LaRoque, M. Leber, E. L. McBride, M. L. Miller, P. Mohanmurthy, B. Monreal, N. S. Oblath, R. G. H. Robertson, L. J. Rosenberg, G. Rybka, D. Rysewyk, M. G. Sternberg, J. R. Tedeschi, T. Thümmler, B. A. VanDevender, N. L. Woods (Project 8 Collaboration), Single-electron detection and spectroscopy via relativistic cyclotron radiation, Phys. Rev. Lett. 114 (2015) 162501.

[111] M. S. Safronova, D. Budker, D. DeMille, D. F. J. Kimball, A. Derevianko, C. W. Clark, Search for new physics with atoms and molecules, ARXIV:1710.01833 (2017).

[112] V. A. Dzuba, V. V. Flambaum, I. B. Khriplovich, Enhancement of P- and T-nonconserving effects in rare-earth atoms, Zeitschrift für Physik D Atoms, Molecules and Clusters 1 (1986) 243–245.

[113] A. T. Nguyen, D. Budker, D. DeMille, M. Zolotorev, Search for parity nonconservation in atomic dysprosium, Phys. Rev. A 56 (1997) 3453–3463.

[114] A. Kramida, Y. Ralchenko, J. Reader, NIST ASD Team, NIST atomic spectra database (version 5.4), 2016. URL: http://physics.nist.gov/asd.

[115] C.-H. Li, S. M. Rochester, M. G. Kozlov, D. Budker, Unusually large polarizabilities and previously unidentified atomic states in ba, Phys. Rev. A 69 (2004) 042507.

[116] N. A. Leefer, A. Viatkina, M. G. Kozlov, D. Budker, Spectroscopy of complex systems revisited (unpublished), 2016.

[117] Ryabchikova, T., Ryabtsev, A., Kochukhov, O., Bagnulo, S., Rare-earth elements in the atmosphere of the magnetic chemically peculiar star HD 144897 - new classification of the Nd III spectrum, A&A 456 (2006) 329–338.

[118] L. M. Freyhammer, V. G. Elkin, D. W. Kurtz, G. Mathys, P. Martinez, Discovery of 17 new sharp-lined Ap stars with magnetically resolved lines, Monthly Notices of the Royal Astronomical Society 389 (2008) 441–460.

[119] U. Heiter, T. Lebzelter, Gaia spectroscopy overview and comparative spectrum modeling for cool giants, Journal of Physics: Conference Series 328 (2011) 012001.

[120] V. V. Flambaum, A. A. Gribakina, G. F. Gribakin, M. G. Kozlov, Structure of compound states in the chaotic spectrum of the ce atom: Localization properties, matrix elements, and enhancement of weak perturbations, Phys. Rev. A 50 (1994) 267–296.

[121] V. A. Dzuba, V. V. Flambaum, Exponential increase of energy level density in atoms: Th and Th II, Phys. Rev. Lett. 104 (2010) 213002.

[122] A. V. Viatkina, M. G. Kozlov, V. V. Flambaum, Prediction of quantum many-body chaos in the protactinium atom, Phys. Rev. A 95 (2017) 022503.

[123] T. Maier, H. Kadau, M. Schmitt, M. Wenzel, I. Ferrier-Barbut, T. Pfau, A. Frisch, S. Baier, K. Aikawa, L. Chomaz, M. J. Mark, F. Ferlaino, C. Makrides, E. Tiesinga, A. Petrov, S. Kotochigova, Emergence of chaotic scattering in ultracold Er and Dy, Phys. Rev. X 5 (2015) 041029.

[124] V. V. Flambaum, M. G. Kozlov, G. F. Gribakin, Coherent and stochastic contributions of compound resonances in atomic processes: Electron recombination, photoionization, and scattering, Phys. Rev. A 91 (2015) 052704.

[125] M. J. Morrison, A. Derevianko, Proposed search for t-odd, p-even interactions in spectra of chaotic atoms, Phys. Rev. A 86 (2012) 022115.

[126] R. H. Garstang, Atoms in high magnetic fields (white dwarfs), Reports on Progress in Physics 40 (1977) 105.

[127] T. A. Miller, The use of high magnetic fields in the spectroscopy of simple atoms and molecules, Journal of Magnetism and Magnetic Materials 11 (1979) 259 – 268.

[128] A. V. Viatkina, Influence of Magnetic Field on Statistical Properties of Dense Atomic Spectra, Master's thesis, Saint-Petersburg State University, 2016.

[129] S. A. Diddams, L. Hollberg, V. Mbele, Molecular fingerprinting with the resolved modes of a femtosecond laser frequency comb, Nature 445 (2007) 627–630.

[130] L. C. Sinclair, K. C. Cossel, T. Coffey, J. Ye, E. A. Cornell, Frequency comb velocity-modulation spectroscopy, Phys. Rev. Lett. 107 (2011) 093002.

[131] L. Nugent-Glandorf, T. Neely, F. Adler, A. J. Fleisher, K. C. Cossel, B. Bjork, T. Dinneen, J. Ye, S. A. Diddams, Mid-infrared virtually imaged phased array spectrometer for rapid and broadband trace gas detection, Opt. Lett. 37 (2012) 3285–3287.

[132] M. Kozlov, S. Porsev, M. Safronova, I. Tupitsyn, Ci-mbpt: A package of programs for relativistic atomic calculations based on a method combining configuration interaction and many-body perturbation theory, Computer Physics Communications 195 (2015) 199–213.

[133] C. Harabati, V. A. Dzuba, V. V. Flambaum, Identification of atoms that can bind positrons, Physical Review A 89 (2014).

[134] F. Ficek, P. Fadeev, V. V. Flambaum, D. F. J. Kimball, M. G. Kozlov, Y. V. Stadnik, D. Budker, Constraints on exotic spin-dependent interactions between matter and antimatter from antiprotonic helium spectroscopy, arXiv preprint arXiv:1801.00491 (2018).

[135] K. Zioutas, Y. Semertzidis, A new detector scheme for axions, Physics Letters A 130 (1988) 94 – 97.

[136] P. Sikivie, Axion dark matter detection using atomic transitions, Physical Review Letters 113 (2014).

[137] L. Santamaria, C. Braggio, G. Carugno, V. D. Sarno, P. Maddaloni, G. Ruoso, Axion dark matter detection by laser spectroscopy of ultracold molecular oxygen: a proposal, New Journal of Physics 17 (2015) 113025.

[138] A. F. Borghesani, C. Braggio, G. Carugno, F. Chiossi, A. D. Lieto, M. Guarise, G. Ruoso, M. Tonelli, Particle detection through the quantum counter concept in YAG:Er3+, Applied Physics Letters 107 (2015) 193501.

[139] B. Roberts, Y. Stadnik, V. Dzuba, V. Flambaum, N. Leefer, D. Budker, Parity-violating interactions of cosmic fields with atoms, molecules, and nuclei: Concepts and calculations for laboratory searches and extracting limits, Physical Review D 90 (2014).

[140] H. H. Ten Kate, Atlas superconducting toroids—the largest ever built, International Journal of Modern Physics A 25 (2010) 2933–2954.

[141] Webpage, 2018. URL: https://nationalmaglab.org/about/maglab-dictionary/bitter-plate.

[142] Webpage, 2018. URL: https://nationalmaglab.org/news-events/feature-stories/making-resistive-magnets.

[143] Webpage, 2018. URL: https://nationalmaglab.org/news-events/news/new-world-record-magnet-fulfills-superconducting-promise.

[144] Y. Iwasa, S. Hahn, First-cut design of an all-superconducting 100-T direct current magnet, Applied Physics Letters 103 (2013) 253507.

[145] Webpage, 2018. URL: http://www.superpower-inc.com/content/2g-hts-wire.

[146] M. Strongin, P. P. Craig, W. B. Sampson, Advances in superconducting magnets, Scientific American Magazine, (1967).

[147] Webpage, 2016. URL: https://nationalmaglab.org/images/magnet_development/asc/searchable_docs/asc_resources/coated_conductors/2016/io_06_kim.pdf.

[148] Webpage, 2016. URL: https://indico.cern.ch/event/445667/contributions/2562794/attachments/1514915/2363780/Or09-01-Mon-Af-Or9NI.REBCO.Progress_20170828_SH.pdf.

[149] S. Ahn, S. Youn, J. Yoo, D. Kim, J. Jeong, M. Ahn, J. Kim, D. Lee, J. Lee, T. Seong, Y. Semertzidis, Magnetoresistance in copper at high frequency and high magnetic fields, Journal of Instrumentation 12 (2017) P10023.

[150] M. D. Bird, I. R. Dixon, J. Toth, Large, high-field magnet projects at the NHMFL, IEEE Transactions on Applied Superconductivity 25 (2015) 1–6.

[151] A. den Ouden, C. A. Wulffers, N. E. Hussey, G. Laureijs, F. J. P. Wijnen, G. F. A. J. Wulterkens, M. D. Bird, I. R. Dixon, J. A. A. J. Perenboom, Progress in the development of the HFML 45 T hybrid magnet, IEEE Transactions on Applied Superconductivity 26 (2016) 1–7.

[152] Z. Gan, I. Hung, X. Wang, J. Paulino, G. Wu, I. M. Litvak, P. L. Gor'kov, W. W. Brey, P. Lendi, J. L. Schiano, M. D. Bird, I. R. Dixon, J. Toth, G. S. Boebinger, T. A. Cross, NMR spectroscopy up to 35.2 T using a series-connected hybrid magnet, Journal of Magnetic Resonance 284 (2017) 125 – 136.

[153] I. R. Dixon, S. T. Bole, K. R. Cantrell, S. T. Hannahs, J. G. Kynoch, W. S. Marshall, A. A. Powell, J. Toth, M. D. Bird, The 36-T series-connected hybrid magnet system design and integration, IEEE Transactions on Applied Superconductivity 27 (2017) 1–5.

[154] Webpage, 2018. URL: https://nationalmaglab.org/news-events/news/national-maglab-racks-up-another-record.

[155] P. Pugnat et al, Progress in the construction of the 43 T hybrid magnet at LNCMI-Grenoble, IEEE Trans. Appl. Supercond. 28 (2018).

[156] P. Pugnat, H. J. Schneider-Muntau, Hybrid magnets: Past, present, and future, IEEE Transactions on Applied Superconductivity 24 (2014) 1–6.

[157] F. Debray, J. Dumas, C. Trophime, N. Vidal, DC high field magnets at the LNCMI, IEEE Transactions on Applied Superconductivity 22 (2012) 4301804–4301804.

[158] Webpage, 2018. URL: http://www.toulouse.lncmi.cnrs.fr/spip.php?page=article&id_article=102&lang=en.

[159] O. Portugall, N. Puhlmann, H. U. Müller, M. Barczewski, I. Stolpe, M. von Ortenberg, Megagauss magnetic field generation in single-turn coils: new frontiers for scientific experiments, Journal of Physics D: Applied Physics 32 (1999) 2354.

[160] O. Portugall, P. Y. Solane, P. Plochocka, D. K. Maude, R. J. Nicholas, Beyond 100 tesla: Scientific experiments using single-turn coils, Comptes Rendus Physique 14 (2013) 115 – 120. Physics in High Magnetic Fields / Physique en champ magnétique intense.

[161] H. Schneider-Muntau (Ed.), Megagauss Magnetic Field Generation, its application to science and ultra-high pulsed power technology. Proc. VIIIth International conference on Megagauss Magnetic field Generation and Related Topics, Tallahassee, FL., Singapore: World Scientific, 1998.

[162] N. Miura, Physics of semiconductors in high magnetic fields, volume 15, Oxford University Press, 2008.

[163] S. Takeyama, E. Kojima, A copper-lined magnet coil with maximum field of 700 T for electromagnetic flux compression, Journal of Physics D: Applied Physics 44 (2011) 425003.

[164] D. Nakamura, H. Sawabe, S. Takeyama, Note: An approach to 1000 T using the electromagnetic flux compression, Review of Scientific Instruments 89 (2018) 016106.

[165] J. Béard, J. Billette, N. Ferreira, P. Frings, J. M. Lagarrigue, F. Lecouturier, J. P. Nicolin, Design and tests of the 100-T triple coil at LNCMI, IEEE Transactions on Applied Superconductivity 28 (2018) 1–5.

[166] J. Billette, F. Duc, P. Frings, M. Nardone, A. Zitouni, C. Detlefs, T. Roth, W. Crichton, J. E. Lorenzo, G. L. J. A. Rikken, A 30 T pulsed magnet with conical bore for synchrotron powder diffraction, Review of Scientific Instruments 83 (2012) 043904.

[167] W. Knafo, F. Duc, F. Bourdarot, K. Kuwahara, H. Nojiri, D. Aoki, J. Billette, P. Frings, X. Tonon, E. Lelièvre-Berna, J. Flouquet, L. P. Regnault, Field-induced spin-density wave beyond hidden order in URu_2Si_2, Nature Communications 7 (2016) 13075 EP.

[168] A. Nikiel, P. Blümler, W. Heil, M. Hehn, S. Karpuk, A. Maul, E. Otten, L. M. Schreiber, M. Terekhov, Ultrasensitive ^{3}He magnetometer for measurements of high magnetic fields, The European Physical Journal D 68 (2014) 330.

[169] D. Nakamura, H. Sawabe, Y. H. Matsuda, S. Takeyama, Precise measurement of a magnetic field generated by the electromagnetic flux compression technique, Review of Scientific Instruments 84 (2013) 044702.

[170] S. George, N. Bruyant, J. Béard, S. Scotto, E. Arimondo, R. Battesti, D. Ciampini, C. Rizzo, Pulsed high magnetic field measurement with a rubidium vapor sensor, Review of Scientific Instruments 88 (2017) 073102.

[171] D. Ciampini, R. Battesti, C. Rizzo, E. Arimondo, Optical spectroscopy of a microsized Rb vapor sample in magnetic fields up to 58 T, Phys. Rev. A 96 (2017) 052504.

[172] K. von Klitzing, The quantized Hall effect, Reviews of Modern Physics 58 (1986) 519–531.