

Green's function for 1D superconductive systems

Yury Holubeu *

December 24, 2024

This draft is not aimed for distribution.

Green's function for 1D systems is discussed in details.

Contents

1	Preface and main motivation	9
I	Green's function for 1D systems in a Nutshell	10
1.1	Main Formulas and Ideas (?)	10
1.1.1	PPS BGf method	10
1.1.2	Caroli, Comberscot's et al. Tunneling I, II: Formalism, Free electrons	22
1.1.3	Caroli, Comberscot's et al. Tunneling III, IV: impurities, Electron-phonons	33
1.1.4	The Keldysh method	44
2	Overview of other methods	44
2.0.1	BGf for spinful Majorana NW by Alvarado et. al.	44
II	Very Useful and Famous Methods	54
3	PPS BGf Theory	54
3.1	Overview	54
3.1.1	Conclusions and Outlook	54
3.1.2	BGf glossary	55
3.1.3	Introduction	55
3.1.4	Algorithm of usage of interface Green's functions	57
3.1.5	Checking if all is fine at middle stages of computation	57
3.1.6	How to figure out this bGf method (???)	57
3.2	Interface Green's functions for discrete models	58
3.2.1	Model and problem formulation	58
3.2.2	Single-barrier setup	61
3.2.3	Double-barrier setup	67
3.2.4	Multiple barriers	69
3.2.5	Easy illustrative examples	75
3.3	Interface Green's functions for continuum models	75
3.3.1	Boundary Green's functions	75
3.3.2	Easy illustrative examples	80
3.4	Main examples	80
3.4.1	Basic Josephson model by PPS BGf	80
3.4.2	Random non-magnetic impurities in JJ by PPS BGf	87
3.4.3	The effect of the local Coulomb interaction on JJ by PPS BGf	89
3.4.4	Junction of two Majorana wires by PPS BGf	93
4	Basic Theory of Multiterminal Junctions	104

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

5	Green's functions for a multi-terminal model	104
5.1	Model's description	104
5.2	Green's function	104
5.3	Global DOS and the bound state equation	107
5.4	Josephson currents	109
5.5	Continuum model	111
5.5.1	Three quantum dots	112
5.5.2	Modelling of the hopping amplitudes	114
5.5.3	Evaluation of $\text{Det } D$	116
5.5.4	Bound state equation	117
5.5.5	Continuum limit everywhere	119
5.6	Necessary condition for existence of zero-energy solutions	122
5.7	Generalization from equilateral to isosceles triangular geometry of contacts	124
5.7.1	Relation to the scattering matrix approach	125
5.8	Microwave drive	125
5.9	Rest theory	128
5.9.1	Model Hamiltonian for general terminal case	128
5.9.2	Formalism for one SC terminal	135
5.9.3	Formalism for several SC terminals	138
6	Quasiclassical GF approach to mesoscopic SC by Belzig et al.	140
6.1	Introduction	140
6.2	Quasiclassical formalism	142
6.2.1	Green's functions	142
6.2.2	The quasiclassical approximation	143
6.2.3	Kinetics and time-ordering	144
6.2.4	The dirty limit	146
6.2.5	Boundary Conditions	149
6.2.6	Parameterizations	150
6.2.7	Solution strategies	152
6.3	Spectral quantities and equilibrium properties	152
6.3.1	Density of states	152
6.3.2	Magnetic response of a proximity structure	155
6.3.3	Equilibrium supercurrent in a diffusive SNS junction	160
6.4	Applications to nonequilibrium transport problems	162
6.4.1	Supercurrent under nonequilibrium conditions: The SNS Transistor	162
6.4.2	Dissipative current in a mesoscopic wire	165
6.4.3	Supercurrent and interference effects on the dissipative current	167
6.4.4	Transport through tunneling barriers	169
6.4.5	Four-Point conductance	170
6.5	Conclusion	173
7	BGf of topological insulators and superconductors by Peng	173
7.1	Conclusion	173
7.2	Introduction	174
7.3	Basic considerations	175
7.3.1	Boundary Green function	175
7.3.2	The Su-Schrieffer-Heeger (SSH) model: An illustrative example	176
7.3.3	Topological invariants	178
7.4	Boundary Green function and transfer matrix	179
7.5	Application to 1D and 2D models	182
7.5.1	Perturbation theory	183
7.5.2	Kitaev chain	183
7.5.3	Chern insulator	187

7.6	Higher-order topological insulators	189
7.6.1	Model	189
7.6.2	Effective boundary Hamiltonian	190
7.7	A1: Boundary Green function and reflection matrix	192
7.8	A2: Topological invariants	193
7.8.1	Bott periodicity	194
7.8.2	Topological invariants	197
7.9	Easy illustrative examples	200
7.9.1	Example 1	200
7.9.2	Example 2	200
8	The Non-Equilibrium Green Function (NEGF) Method	200
8.1	Introduction	200
8.1.1	Decoupling NEGF from MBPT	202
8.1.2	Outline	203
8.2	NEGF Equations	203
8.3	NEGF equations from one-electron Schrödinger equation	204
8.3.1	NEGF Eqs. (8.1-8.2)	205
8.3.2	NEGF Eq. (8.3)	205
8.4	A simple example	206
8.4.1	Hamiltonian $\mathbf{H}$	207
8.4.2	Self-energy due to contacts	208
8.4.3	Inscattering from contacts	208
8.4.4	Current	209
8.4.5	Dephasing interactions	210
8.4.6	Diffusive spin transport using NEGF	212
8.4.7	Application to superconducting devices	213
8.5	Frequently asked questions	213
8.5.1	Shouldn't we also consider the Poisson equation for modeling real devices? . . .	213
8.5.2	Can you arbitrarily designate the ends as contacts?	214
8.5.3	Does NEGF give the correct coherent and semiclassical limits?	215
8.5.4	Questions about Kubo formalism and NEGF	215
8.5.5	Questions about many-body processes and NEGF	216
8.5.6	How can we teach NEGF without advanced quantum statistical mechanics? . .	217
8.5.7	But is this the real NEGF?	217
8.5.8	Contact-ing Schrödinger	218
9	Tunneling current I: Formalism	218
9.0.1	1. Introduction	218
9.1	2. Model Hamiltonian and Green function at zero voltage	219
9.1.1	2.1. The model Hamiltonian	219
9.1.2	2.2. The Green functions at zero bias	221
9.1.3	Other basic questions about the model	223
9.2	3. The tunneling current	223
9.2.1	3.1. The current operator	223
9.2.2	3.2. Calculation of the current	223
9.2.3	3.3. Discussion	227
9.2.4	Appendix	229
9.3	Easy illustrative examples	232
9.3.1	Example 1	232
9.3.2	Example 2	232

10 Tunneling current II: Free electron description by Caroli et. al.	232
10.1 1. Introduction	232
10.2 2. Expression of the current in the continuous representation	234
10.2.1 2.1. The Green function at zero bias	234
10.2.2 2.2 Formal expression of the current	238
10.3 3. The current through a square barrier	241
10.4 4. Conclusion	242
10.5 Appendix	243
10.6 Easy illustrative examples	245
10.6.1 Example 1	245
10.6.2 Example 2	245
11 Tunneling current III: Localized impurities in the barrier by Comberscot	245
11.0.1 1. Introduction	245
11.1 2. One dimensional model	246
11.1.1 2.1. Localized functions formalism	246
11.1.2 2.2. Formal expression of the current	249
11.1.3 2.3. The transmission coefficient	251
11.2 3. Generalization to three dimensions	253
12 Tunneling current IV: Electron-phonon interaction effects by Caroli et. al.	257
12.0.1 1. Introduction	257
12.0.2 2. Perturbation formalism	258
12.0.3 3. Electron-phonon interaction in the electrodes of an MIM junction	264
12.0.4 4. Electron-phonon interaction in the barrier	271
12.0.5 Appendix	275
12.1 Easy illustrative examples	277
12.1.1 Example 1	277
12.1.2 Example 2	277
III Examples and Solved Problems	278
13 Illustrative examples	278
13.1 Superconductive ring	278
13.1.1 Three-terminal ring	278
13.1.2 Two-terminal SC ring	286
13.1.3 Two-terminal SC ring with $H = p^2/2m + \dots$	286
14 Examples for practice	286
IV Other Methods	287
15 BGf approach for spinful Majorana nanowires by Alvarado et. al.	287
15.1 Concluding remarks	287
15.2 Introduction	288
15.3 Boundary Green's function	289
15.4 Spinful single-channel hybrid nanowires	291
15.5 Two-channel class-D nanowire	295
15.6 TRITOPS nanowires	297
15.6.1 Single-channel case	297
15.6.2 Two-channel case	299
15.7 Phase-biased topological Josephson junctions	300
15.7.1 Three-terminal TS junctions	301

15.7.2	TRITOPS-TS junction	304
15.8	On the spinful single-channel model	306
15.9	On TRITOPS wires	308
15.10	Easy illustrative examples	309
15.10.1	Example 1	309
15.10.2	Example 2	309
16	Conductance calculations: mode matching and GF by by Khomyakov et al	309
16.1	Summary and Conclusions	309
16.2	Introduction	310
16.3	Hamiltonian	311
16.4	Mode matching approach	312
16.4.1	Bloch matrices	312
16.4.2	Transmission matrix	313
16.4.3	Green function matrix	314
16.5	Green function approach	315
16.5.1	Partitioning	315
16.5.2	Transmission matrix	316
16.6	Mode matching versus Green functions	317
16.6.1	Green functions of ideal wires in terms of Bloch matrices	317
16.6.2	Equivalence of mode matching and Green function approaches	318
16.6.3	Invariance of transmission probability	319
16.6.4	The Caroli expression	320
16.6.5	Transmission matrix: a compact expression	321
16.7	Examples	322
16.7.1	A simple analytical model	322
16.7.2	Fe vacuum Fe tunnel junction	324
16.8	Velocity matrix	328
16.9	Projector matrix	329
16.10	Easy illustrative examples	330
16.10.1	Example 1	330
16.10.2	Example 2	330
17	Thermal transport in one-dimensional spin heterostructures	330
17.1	Introduction	330
17.2	The formalism	331
17.2.1	Model	331
17.2.2	Energy balance	333
17.2.3	Solving Dyson's equations	334
17.2.4	Heat currents and transmission functions	337
17.3	Results	339
17.4	Summary and discussion	341
17.5	Green's functions for an open chain with p -wave superconductivity	343
17.6	Easy illustrative examples	344
17.6.1	Example 1	344
17.6.2	Example 2	344
18	Generalized Fisher-Lee relation	344
18.1	Conclusion	345
18.2	Introduction	346
18.3	Wave function matching approach	348
18.3.1	Ideal leads	349
18.3.2	The scattering region	352
18.3.3	Transmission	354

18.4	AGF formalism	355
18.4.1	A brief review of the conventional AGF formalism	355
18.4.2	Equivalence of WFM and AGF	356
18.4.3	Mode-resolved transmissions	358
18.5	Applications	359
18.6	Velocity matrices	363
18.7	Injection matrix	363
18.8	Eigenvector matrices from the Green's functions	364
19	Green's function by Wimmer	364
19.1	Chapter 2 Green's function formalism for transport	364
19.1.1	2.1 Introduction	364
19.1.2	2.2 Basic definitions	365
19.1.3	2.3 Green's functions of non-interacting systems	369
19.1.4	2.4 Non-equilibrium perturbation theory	373
19.1.5	2.4.1 The need for a non-equilibrium theory	373
19.1.6	2.4.2 Contour-ordered Green's function theory	375
19.1.7	2.4.3 Real-time formulation	379
19.1.8	2.5 Transport in tight-binding models	381
19.1.9	2.5.1 Green's functions in tight-binding	381
19.1.10	2.5.2 Transport equations	382
19.1.11	2.5.3 Connection with the scattering formalism	387
19.1.12	2.6 Challenges for a numerical implementation	390
19.2	Chapter 3 Lead Green's functions	391
19.2.1	3.1 Introduction	391
19.2.2	3.2 A general expression for the lead Green's function	393
19.2.3	3.2.1 Green's function of the infinite wire	393
19.2.4	3.2.2 Zeroes of $(\mathbf{E} + i\eta - \mathbf{H}_0)\mathbf{z} - \mathbf{H}_1\mathbf{z}^2 - \mathbf{H}_{-1}$	395
19.2.5	3.2.3 The role of degenerate modes	397
19.2.6	3.2.4 Solving the contour integral	398
19.2.7	3.3 A more stable algorithm based on the Schur decomposition	400
19.2.8	3.3.1 Failure of the eigendecomposition based algorithm	400
19.2.9	3.3.2 Surface Green's function from the Schur decomposition	403
19.2.10	3.3.3 Numerical Stability	405
19.2.11	3.3.4 Examples	407
19.2.12	3.4 Summary	407
19.3	Chapter 4 Optimal block-tridiagonalization of matrices for quantum transport	408
19.3.1	4.1 Introduction	408
19.3.2	4.2 Definition of the problem	411
19.3.3	4.2.1 Definition of the matrix reordering problem	411
19.3.4	4.2.2 Mapping onto a graph partitioning problem	411
19.3.5	4.3 Optimal Matrix reordering by graph partitioning	413
19.3.6	4.3.1 A local approach—breadth first search	413
19.3.7	4.3.2 A global approach—recursive bisection	414
19.3.8	4.3.3 Computational complexity	418
19.3.9	4.4 Examples: Charge transport in two-dimensional systems	419
19.3.10	4.4.1 Ballistic transport in two-terminal devices	419
19.3.11	4.4.2 Multi-terminal structures	424
19.3.12	4.5 Summary	425
19.4	Chapter 5 Magnetic field effects on tunneling anisotropic magnetoresistance	426
19.4.1	5.1 Introduction	426
19.4.2	5.2 Magnetic field dependence of TAMR	428
19.4.3	5.2.1 Model	428
19.4.4	5.2.2 Numerical results	430

19.4.5	5.2.3 A qualitative picture	431
19.4.6	5.3 Summary	435
19.5	Chapter 6 The graphene edge state	435
19.5.1	6.1 Introduction	435
19.5.2	6.2 Fundamentals of graphene	436
19.5.3	6.2.1 Graphene Hamiltonian	436
19.5.4	6.2.2 Symmetries of the graphene Hamiltonian	438
19.5.5	6.3 Characterization of the edge state	440
19.5.6	6.3.1 Edge state at a single zigzag edge	440
19.5.7	6.3.2 Edge state in a zigzag nanoribbon	443
19.5.8	6.3.3 Effects of a staggered potential	445
19.5.9	6.3.4 Effects of next-nearest neighbor hopping	447
19.5.10	6.4 Edge state transport in zigzag nanoribbons	449
19.5.11	6.4.1 Model-dependence of edge state transport-some peculiar examples	449
19.5.12	6.4.2 Where does the current flow?	453
19.5.13	6.4.3 Effects of perturbations on the current flow	455
19.5.14	6.4.4 Scattering from edge defects	460
19.5.15	6.5 Summary	462
19.6	Chapter 7 Edge state based spintronics in graphene	463
19.6.1	7.1 Introduction	463
19.6.2	7.2 Mean-field model for edge-state magnetism	464
19.6.3	7.3 Spin currents in rough graphene nanoribbons	466
19.6.4	7.3.1 Basic ideas	466
19.6.5	7.3.2 Spin conductance of rough graphene nanoribbons	467
19.6.6	7.3.3 Universal spin conductance fluctuations	471
19.6.7	7.3.4 All-electrical detection of edge magnetism	474
19.6.8	7.4 Edge-state induced Spin Hall effect	475
19.6.9	7.4.1 Spin-dependent deflection of the edge state	475
19.6.10	7.4.2 Pure spin currents in graphene micro-bridges	477
19.6.11	7.5 Summary	479
19.7	Chapter 8 Summary and perspectives	480
19.7.1	8.1 Summary	480
19.7.2	8.2 Outlook	481
19.8	Appendix A Observables in tight-binding	482
19.8.1	A. 1 Single-particle operators in tight-binding	482
19.8.2	A. 2 Observables in terms of Green's functions	484
19.9	Appendix B The recursive Green's function method	486
19.9.1	B. 1 Basic principles	486
19.9.2	B. 2 Summary of algorithms	489
19.9.3	B. 3 Computational complexity and implementation	490
19.10	Appendix C Generalized Fisher-Lee relation	491
19.10.1	C. 1 Landauer-Büttiker formalism	491
19.10.2	C. 2 Generalized orthogonality relations	492
19.10.3	C. 3 Fisher-Lee relation	495
19.10.4	C. 4 Equivalence of scattering and non-equilibrium Green's function formalism	498
19.11	Appendix D Details of the derivation of the lead Green's function	499
19.11.1	D. 1 $\lambda(E + i\eta)$ for propagating modes	499
19.11.2	D. 2 Derivation of Eq. (3.27)	501
19.11.3	D. 3 Summary of the numerical algorithms	502
19.11.4	D.3.1 Eigendecomposition based algorithms	502
19.11.5	D.3.2 Schur decomposition based algorithms	503
19.12	Appendix E The Fiduccia-Mattheyses algorithm	505
19.12.1	E. 1 Graphs and hypergraphs	505

19.12.2 E. 2 Fiduccia-Mattheyses bisection	505
19.13 Appendix F The method of finite differences	505
19.13.1 F. 1 Basic ideas	505
19.13.2 F. 2 Example	506
19.13.3 F. 3 Finite differences form of the Hamiltonian	507
19.13.4 F. 4 Bloch's theorem and periodic boundary conditions	509
19.14 Appendix G Tight-binding model for graphene	510
19.14.1 G. 1 Lattice structure	510
19.14.2 G. 2 Electronic structure	512
19.14.3 G.2.1 Tight-binding model	512
19.14.4 G.2.2 Band structure in tight-binding approximation	515
19.14.5 G.2.3 Effective Hamiltonian	518
19.14.6 G. 3 Graphene nanoribbons	521
19.14.7 G.3.1 Band structure	521
19.14.8 G.3.2 Boundary conditions	522
20 Diagram technique for nonequilibrium processes by Keldysh	523
21 Hamiltonian approach to the transport of SC point contacts by Cuevas et al.	535
21.0.1 I. INTRODUCTION	536
21.0.2 II. MODEL AND METHOD	537
21.0.3 III. $N - N$ AND $N - S$ CONTACTS	540
21.0.4 IV. THE S-S CONTACT	542
21.0.5 A. An efficient algorithm for evaluating the ac current	542
21.0.6 B. Analysis of ac and dc $I - V$ characteristics	544
21.0.7 C. Small bias regime	546
21.0.8 V. CONCLUDING REMARKS	551
21.0.9 APPENDIX A	551
21.0.10 APPENDIX B	553
21.0.11 APPENDIX C	554
21.0.12 1. Linear regime ($\eta \gg eV$)	554
21.0.13 2. Nonlinear regime ($n \ll eV$)	555
22 An intro to the diffusive proximity-coupled SC systems by Chandrasekhar	556
22.1 Transport equations in the diffusive approximation	556
22.2 The Keldysh Green's Functions	560
22.3 The quasiclassical approximation	567
22.4 Non-equilibrium Green's functions for superconducting systems	571
22.5 Quasiclassical Superconducting Green's Functions	574
22.6 The dirty limit: the Usadel equation	578
22.7 Parametrization of the quasiclassical Green's function	585
22.8 Applications of the quasiclassical equations to proximity-coupled systems	588
22.8.1 Proximity-coupled wire	588
22.8.2 Superconductor-metal bilayer	593
22.8.3 The SNS junction and Andreev interferometers	594
22.9 Summary	596
V Adds	597

A	Дополнения	597
A.1	Useful mathematical formulas	597
A.1.1	Useful mathematical formulas for SC rings	597
A.2	Tips about mathematical transformations	597
A.3	Acknowledgements	597
A.4	Literature	597
A.5	Codes for Green's functions	597
A.6	Overview of applications (???)	597
B	Bibliography	598

1 Preface and main motivation

Amazing facts

(I'll reveal it later)

Puzzles for motivation

(I'll reveal it later)

Part I

Green's function for 1D systems in a Nutshell

1.1 Main Formulas and Ideas (?)

1.1.1 PPS BGf method

Continuum PPS BGf Formalism

Algorithm

$$\mathcal{H} := \int_{-\infty}^{\infty} dx \hat{\psi}^\dagger(x) \left[p \frac{1}{2m(x)} p + \frac{1}{2} \{A(x), p\} + V(x) \right] \hat{\psi}(x) + \frac{1}{2} \int_{-\infty}^{\infty} dx \left[\hat{\psi}^\dagger(x) \hat{\Delta}(x) (\hat{\psi}^\dagger(x))^T + \text{h.c.} \right],$$

$$\begin{aligned} \{\hat{\psi}_\sigma(x), \hat{\psi}_{\sigma'}(x')\} &= 0, \\ \{\hat{\psi}_\sigma(x), \hat{\psi}_{\sigma'}^\dagger(x')\} &= \delta_{\sigma,\sigma'} \delta(x - x'). \end{aligned}$$

$$\begin{pmatrix} m(x) \\ A(x) \\ V(x) \\ \Delta(x) \end{pmatrix} = \Theta(-x) \begin{pmatrix} m_L \\ A_L \\ V_L \\ \Delta_L e^{i\varphi_L} \end{pmatrix} + \Theta(W > x > 0) \begin{pmatrix} m_C \\ A_C \\ V_C \\ 0 \end{pmatrix} + \Theta(x - W) \begin{pmatrix} m_R \\ A_R \\ V_R \\ \Delta_R e^{i\varphi_R} \end{pmatrix}.$$

$$A_\lambda = A_\lambda^{(0)} + \sum_{j=x,y,z} A_\lambda^{(j)} \sigma_j \equiv A_\lambda^{(0)} + \vec{A}_\lambda \cdot \vec{\sigma}; \quad V_\lambda = V_\lambda^{(0)} + \sum_{j=x,y,z} V_\lambda^{(j)} \sigma_j \equiv V_\lambda^{(0)} + \vec{V}_\lambda \cdot \vec{\sigma}$$

$$\hat{\Psi}(x) = \begin{pmatrix} \hat{\psi}(x) \\ i\sigma_y (\hat{\psi}^\dagger(x))^T \end{pmatrix},$$

$$\mathcal{H} = \frac{1}{2} \int_{-\infty}^{\infty} dx \hat{\Psi}^\dagger(x) H \hat{\Psi}(x); \quad H = \begin{pmatrix} h^{(0)} & \Delta(x) \\ \Delta^*(x) & -\sigma_y h^{(0)*} \sigma_y \end{pmatrix}; \quad h^{(0)} = p \frac{1}{2m(x)} p + \frac{1}{2} \{A(x), p\} + V(x).$$

$$H = \tau_z \left[p \frac{1}{2m(x)} p + \frac{1}{2} \{ \vec{A}(x) \cdot \vec{\sigma}, p \} + V^{(0)}(x) \right] + \tau_0 \left[\frac{1}{2} \{ A^{(0)}(x), p \} + \vec{V}(x) \cdot \vec{\sigma} \right] + \Delta(x) \tau_+ + \Delta^*(x) \tau_-.$$

$$\begin{aligned} p_0 &= \mathcal{U}_1 + \mathcal{L}_L - \mathcal{L}_{L \rightarrow C}, \\ p_W &= \mathcal{U}_2 - \mathcal{L}_R + \mathcal{L}_{R \rightarrow C}, \end{aligned}$$

$$\begin{aligned}
\mathcal{L}_{L \rightarrow C} &= \frac{\mathcal{A}_C}{2} G_1^{(0,C)}(0^-, 0) [G^{(0,C)}(0, 0)]^{-1} + \frac{i}{2} \mathcal{B}_C, \\
\mathcal{L}_{R \rightarrow C} &= \frac{\mathcal{A}_C}{2} G_1^{(0,C)}(0^+, 0) [G^{(0,C)}(0, 0)]^{-1} + \frac{i}{2} \mathcal{B}_C \\
&= \mathcal{L}_{L \rightarrow C} + [G^{(0,C)}(0, 0)]^{-1}.
\end{aligned}$$

$$G_X^m(x, x') := G^{(0,m)}(x, x') - G^{(0,m)}(x, X) [G^{(0,m)}(X, X)]^{-1} G^{(0,m)}(X, x'),$$

$$\begin{aligned}
G_{X,Y}^m(x, x') &:= G^{(0,m)}(x, x') - \\
&\quad - G^{(0,m)}(x, X) [G_Y^m(X, X)]^{-1} G_Y^m(X, x') - G^{(0,m)}(x, Y) [G_X^m(Y, Y)]^{-1} G_X^m(Y, x').
\end{aligned}$$

Observables

$$\mathcal{A}_\lambda = \frac{\tau_z}{m_\lambda}, \quad \mathcal{U}_{1,2} = 0.$$

$$J(\varphi) = \frac{2e}{\hbar} \frac{dF}{d\varphi}; \quad \delta\rho(\omega, \varphi) = -\frac{1}{\pi} \text{Im} \frac{\partial}{\partial\omega} \ln \det d(\omega + i0^+, \varphi).$$

$$J(\varphi) = -\frac{2e}{\hbar\beta} \int_0^\infty d\omega \ln \left(2 \cosh \frac{\beta\omega}{2} \right) \frac{\partial}{\partial\varphi} \delta\rho(\omega, \varphi) = -\frac{1}{\Phi_0} \text{Im} \int_0^\infty d\omega \operatorname{th} \frac{\beta\omega}{2} \frac{\partial}{\partial\varphi} \ln \det d(\omega + i0^+, \varphi),$$

$$\Phi_0 = \frac{\hbar}{2e}, \quad \frac{\partial}{\partial\varphi} \ln \det d(z, \varphi) = \text{tr} \left\{ [d(z, \varphi)]^{-1} \frac{\partial d(z, \varphi)}{\partial\varphi} \right\}$$

$$J(\varphi) = -\frac{\pi}{\Phi_0\beta} \sum_{i\omega_n} \text{tr} \left\{ [d(i\omega_n, \varphi)]^{-1} \frac{\partial d(i\omega_n, \varphi)}{\partial\varphi} \right\} \xrightarrow{T \rightarrow 0} -\frac{1}{2\Phi_0} \int_{-\infty}^{\infty} d\omega \text{tr} \left\{ [d(i\omega, \varphi)]^{-1} \frac{\partial d(i\omega, \varphi)}{\partial\varphi} \right\}$$

Derivation of GF in a nutshell

(!!! what is the main idea of the derivation?????)

$$H_m = \frac{1}{2} \mathcal{A}_m p^2 + \mathcal{B}_m p + \mathcal{C}_m,$$

$$\begin{aligned}
H &= \frac{1}{2} p \mathcal{A}(x) p + \frac{1}{2} \{ \mathcal{B}(x), p \} + \mathcal{C}(x) + \mathcal{U}(x), \\
\begin{Bmatrix} \mathcal{A}(x) \\ \mathcal{B}(x) \\ \mathcal{C}(x) \end{Bmatrix} &= \sum_{m=0}^M \Theta(x_m < x < x_{m+1}) \begin{Bmatrix} \mathcal{A}_m \\ \mathcal{B}_m \\ \mathcal{C}_m \end{Bmatrix}, \\
\mathcal{U}(x) &= \sum_{m=1}^M \mathcal{U}_m \delta(x - x_m),
\end{aligned}$$

$$\Psi(x_m^-) = \Psi(x_m^+) \equiv \Psi(x_m), \quad x_m^\pm = x_m \pm 0^+,$$

$$\frac{\mathcal{A}_m \Psi'(x_m^+) - \mathcal{A}_{m-1} \Psi'(x_m^-)}{2} = \left[\mathcal{U}_m - i \frac{\mathcal{B}_m - \mathcal{B}_{m-1}}{2} \right] \Psi(x_m).$$

$$[z - H]G(x, x'; z) = \delta(x - x')$$

$$[z - H_m]G^{(0,m)}(x, x'; z) = \delta(x - x')$$

$$G^{(0,m)}(x, x'; z) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \frac{e^{ik(x-x')}}{z - h_m(k)},$$

$$h_m(k) = H_m(p \rightarrow k).$$

Multiple barriers in continuum model

$$G(x, x') = \sum_{m=0}^M G^m(x, x') + \sum_{m, m'=1}^M F_m(x) (d^{-1})_{m, m'} \bar{F}_{m'}(x'),$$

$$d_{m, m} = -\mathcal{U}_m - \frac{1}{8} \mathcal{A}_{m-1} \lim_{x \rightarrow x_m^-} \frac{d^2}{dx^2} G^{m-1}(x, x) \mathcal{A}_{m-1} - \frac{1}{8} \mathcal{A}_m \lim_{x \rightarrow x_m^+} \frac{d^2}{dx^2} G^m(x, x) \mathcal{A}_m$$

$$d_{m, m+1} = \frac{1}{4} \mathcal{A}_m G_{12}^m(x_m^+, x_{m+1}^-) \mathcal{A}_m,$$

$$d_{m+1, m} = \frac{1}{4} \mathcal{A}_m G_{12}^m(x_{m+1}^-, x_m^+) \mathcal{A}_m$$

$$F_m(x) = \begin{cases} - \left[G_2^{(0, m-1)}(x, x_m^+) + G_2^{m-1}(x, x_m^-) - G_2^{(0, m-1)}(x, x_m^-) \right] \frac{\mathcal{A}_{m-1}}{2}, & x_{m-1} < x < x_m \\ \left[G_2^{(0, m)}(x, x_m^-) + G_2^m(x, x_m^+) - G_2^{(0, m)}(x, x_m^+) \right] \frac{\mathcal{A}_m}{2}, & x_m < x < x_{m+1} \end{cases}$$

$$\bar{F}_{m'}(x') = \begin{cases} - \frac{\mathcal{A}_{m'-1}}{2} \left[G_1^{(0, m'-1)}(x_{m'}^+, x') + G_1^{m'-1}(x_{m'}^-, x') - G_1^{(0, m'-1)}(x_{m'}^-, x') \right], & x_{m'-1} < x' < x_{m'} \\ \frac{\mathcal{A}_{m'}}{2} \left[G_1^{(0, m')}(x_{m'}^-, x') + G_1^{m'}(x_{m'}^+, x') - G_1^{(0, m')}(x_{m'}^+, x') \right], & x_{m'} < x' < x_{m'+1} \end{cases}$$

$$F_m(x_{m'}) = -\delta_{m, m'},$$

$$\bar{F}_{m'}(x_m) = -\delta_{m', m},$$

$$G(x_m, x_{m'}) = (d^{-1})_{m, m'},$$

$$\rho(\omega) = \sum_{m=0}^M \rho^m(\omega) - \frac{1}{\pi} \text{Im} \frac{\partial}{\partial \omega} \ln \det d(\omega + i0^+),$$

$$d_{\text{tw bar}} = \begin{pmatrix} G(0,0) & G(0,W) \\ G(W,0) & G(W,W) \end{pmatrix}^{-1} = \begin{pmatrix} G^{(C,0)}(0,0) & G^{(C,0)}(0,W) \\ G^{(C,0)}(W,0) & G^{(C,0)}(W,W) \end{pmatrix}^{-1} - \begin{pmatrix} p_0 & 0 \\ 0 & p_W \end{pmatrix}.$$

Single barrier examples in continuum model

Discrete PPS BGf Formalism

Single barrier of a discrete model

$$H = H_L + H_R + |0\rangle W_0 \langle 0| - |0\rangle t_L \langle -1| - |-1\rangle t_L^\dagger \langle 0| - |1\rangle t_R \langle 0| - |0\rangle t_R^\dagger \langle 1|,$$

$$H_L = - \sum_{n=n_0+1}^{-1} \left(|n\rangle t_L \langle n-1| + |n-1\rangle t_L^\dagger \langle n| \right) + \sum_{n=n_0}^{-1} |n\rangle W_L \langle n|,$$

$$H_R = - \sum_{n=1}^{n_2-1} \left(|n+1\rangle t_R \langle n| + |n\rangle t_R^\dagger \langle n+1| \right) + \sum_{n=1}^{n_2} |n\rangle W_R \langle n|,$$

$$G_{n,n'} = G_{n,n'}^L + G_{n,n'}^C + G_{n,n'}^R - \left(G_{n,-1}^L t_L^\dagger + G_{n,1}^R t_R \right) G_{0,n'} - \delta_{n,0} G_{0,0}^C \left(t_L G_{-1,n'} + t_R^\dagger G_{1,n'} \right)$$

$$G^L = \frac{1}{z-H_L}, \quad G^C = |0\rangle \frac{1}{z-W_0} \langle 0|, \quad G^R = \frac{1}{z-H_R}$$

$$G_{n,n'} = G_{n,n'}^L + G_{n,n'}^R + F_n D^{-1} \bar{F}_{n'},$$

$$D := z - W_0 - t_L G_{-1,-1}^L t_L^\dagger - t_R^\dagger G_{1,1}^R t_R$$

$$F_n := -\delta_{n,0} + G_{n,-1}^L t_L^\dagger + G_{n,1}^R t_R,$$

$$\bar{F}_{n'} := -\delta_{0,n'} + t_L G_{-1,n'}^L + t_R^\dagger G_{1,n'}^R.$$

$$F_0 = \bar{F}_0 = -1,$$

$$G_{0,0} = D^{-1}.$$

$$z = \omega + i0^+,$$

$$\sum_{n=n_0}^{n_2} \text{tr} [G_{n,n} - G_{n,n}^L - G_{n,n}^R] = \text{tr} \left[\frac{1}{D} \right] + \sum_{n=n_0}^{-1} \text{tr} \left[\frac{1}{D} t_L G_{-1,n}^L G_{n,-1}^L t_L^\dagger \right] + \sum_{n=1}^{n_2} \text{tr} \left[t_R^\dagger G_{1,n}^R G_{n,1}^R t_R \frac{1}{D} \right],$$

$$\sum_{n=n_0}^{-1} G_{-1,n}^L G_{n,-1}^L = -\frac{\partial G_{-1,-1}^L}{\partial \omega},$$

$$\sum_{n=1}^{n_2} G_{1,n}^R G_{n,1}^R = -\frac{\partial G_{1,1}^R}{\partial \omega},$$

$$\text{tr} [D^{-1} \frac{\partial D}{\partial \omega}] = \frac{\partial}{\partial \omega} \ln \det D,$$

$$\rho(\omega) = \rho^L(\omega) + \rho^R(\omega) - \underbrace{\frac{1}{\pi} \text{Im} \frac{\partial}{\partial \omega} \ln \det D(\omega + i0^+)}_{\text{tunneling}}.$$

Multiple barriers in discrete model

Other important about PPS BGf method

Connection of discrete and continuum models

$$t_m + t_m^\dagger = \frac{\mathcal{A}_m}{a^2},$$

$$i(t_m - t_m^\dagger) = \frac{\mathcal{B}_m}{a},$$

$$W_m = \frac{\mathcal{A}_m}{a^2} + \mathcal{C}_m,$$

$$\mathcal{A}_m, \mathcal{B}_m, \mathcal{C}_m = O(a^0).$$

$$W_0 = \frac{\mathcal{A}_L + \mathcal{A}_R}{2a^2} + \frac{\mathcal{U}_1}{a}.$$

$$G(x, x') = \lim_{a \rightarrow 0} \frac{1}{a} G_{n, n'},$$

$$G(x, x') = G^L(x, x') + G^R(x, x') + F(x) d^{-1} \bar{F}(x'),$$

$$F(x) = \lim_{a \rightarrow 0} F_n = \begin{cases} - \left[G_2^{(0,L)}(x, 0^+) + G_2^L(x, 0^-) - G_2^{(0,L)}(x, 0^-) \right] \frac{\mathcal{A}_L}{2}, & x < 0 \\ \left[G_2^{(0,R)}(x, 0^-) + G_2^R(x, 0^+) - G_2^{(0,R)}(x, 0^+) \right] \frac{\mathcal{A}_R}{2}, & x > 0 \end{cases}$$

$$\bar{F}(x') = \lim_{a \rightarrow 0} \bar{F}_{n'} = \begin{cases} -\frac{\mathcal{A}_L}{2} \left[G_1^{(0,L)}(0^+, x') + G_1^L(0^-, x') - G_1^{(0,L)}(0^-, x') \right], & x' < 0 \\ \frac{\mathcal{A}_R}{2} \left[G_1^{(0,R)}(0^-, x') + G_1^R(0^+, x') - G_1^{(0,R)}(0^+, x') \right], & x' > 0 \end{cases}$$

$$d = \lim_{a \rightarrow 0} a D = -\mathcal{U}_1 - \frac{1}{8} \mathcal{A}_L \lim_{x \rightarrow 0^-} \frac{d^2}{dx^2} G^L(x, x) \mathcal{A}_L - \frac{1}{8} \mathcal{A}_R \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G^R(x, x) \mathcal{A}_R,$$

$$\begin{aligned}
G(x, x') &= G^{L+R}(x, x') + \langle x | G^{L+R} l^\dagger d^{-1} l G^{L+R} | x' \rangle \\
&= G^{L+R}(x, x') + \int dy \int dy' G^{L+R}(x, y) T(y, y') G^{L+R}(y', x'),
\end{aligned}$$

$$\begin{aligned}
\langle x | G^{L+R} l^\dagger | y \rangle &= i\delta(y) F(x), \\
\langle y' | l G^{L+R} | x' \rangle &= -i\delta(y') \bar{F}(x'),
\end{aligned}$$

$$T(y, y') = \langle y | l^\dagger d^{-1} l | y' \rangle$$

$$\begin{aligned}
F(0^+) &= F(0^-) \equiv F(0) = -1, \\
\bar{F}(0^+) &= \bar{F}(0^-) \equiv \bar{F}(0) = -1,
\end{aligned}$$

$$\begin{aligned}
[G_2^{(0,m)}(0, 0^+) - G_2^{(0,m)}(0, 0^-)] \frac{\mathcal{A}_m}{2} &= 1, \\
\frac{\mathcal{A}_m}{2} [G_1^{(0,m)}(0^+, 0) - G_1^{(0,m)}(0^-, 0)] &= 1
\end{aligned}$$

$$G(0,0) = d^{-1}.$$

$$\rho(\omega) = \rho^L(\omega) + \rho^R(\omega) - \frac{1}{\pi} \text{Im} \frac{\partial}{\partial \omega} \ln \det d(\omega + i0^+).$$

$$\rho^m(\omega) = -\frac{1}{\pi} \text{Im} \int_{x_m}^{x_{m+1}} dx \text{tr} G^m(x, x; \omega + i0^+).$$

$$\det d(\omega) = 0$$

$$d = [G^{(0,L)}(0,0)]^{-1} - p_0,$$

$$\det \left[1 - G^{(0,L)}(0,0;\omega) p_0(\omega) \right] = 0.$$

$$p_0 = \mathcal{U}_1 + \frac{1}{8} \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} \left[\mathcal{A}_R G^R(x, x) \mathcal{A}_R - \mathcal{A}_L G_0^L(x, x) \mathcal{A}_L \right]$$

$$[G^{(0,L)}(0)]^{-1} = -\frac{1}{8} \mathcal{A}_L \left[\lim_{x \rightarrow 0^-} \frac{d^2}{dx^2} G^L(x, x) + \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G_0^L(x, x) \right] \mathcal{A}_L,$$

$$\frac{\mathcal{A}_R}{2} [G_1^R(0^+, x') + F'(0^+)d^{-1}\bar{F}(x')] - \frac{\mathcal{A}_L}{2} [G_1^L(0^-, x') + F'(0^-)d^{-1}\bar{F}(x')] = - [\mathcal{U}_1 - i\frac{\mathcal{B}_R - \mathcal{B}_L}{2}] d^{-1}\bar{F}(x'),$$

$$\bar{F}(x') = \frac{1}{2}\mathcal{A}_R G_1^R(0^+, x') - \frac{1}{2}\mathcal{A}_L G_1^L(0^-, x'),$$

$$d = -\mathcal{U}_1 - \frac{\mathcal{A}_R}{2}F'(0^+) + \frac{\mathcal{A}_L}{2}F'(0^-) + i\frac{\mathcal{B}_R - \mathcal{B}_L}{2}.$$

$$d = -\mathcal{U}_1 + \mathcal{L}_R - \mathcal{L}_L,$$

$$\mathcal{L}_R = \frac{\mathcal{A}_R}{2}G_1^{(0,R)}(0^+, 0)[G^{(0,R)}(0, 0)]^{-1} + \frac{i}{2}\mathcal{B}_R,$$

$$\mathcal{L}_L = \frac{\mathcal{A}_L}{2}G_1^{(0,L)}(0^-, 0)[G^{(0,L)}(0, 0)]^{-1} + \frac{i}{2}\mathcal{B}_L.$$

Josephson model in in PPS BGF

Basic Josephson model

$$\hat{\Psi}(x) = (\hat{\psi}_\uparrow(x), \hat{\psi}_\downarrow^\dagger(x))^T$$

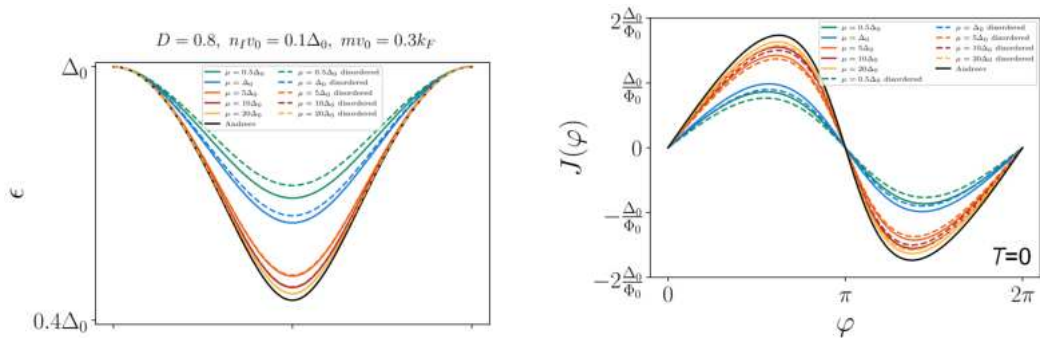
$$H_\lambda = \begin{pmatrix} \frac{p^2}{2m} - \mu & \Delta_\lambda \\ \Delta_\lambda^* & -\frac{p^2}{2m} + \mu \end{pmatrix}, \quad \lambda = R, L,$$

$$|\Delta_\lambda| = \Delta_0 > 0.$$

$$\varphi = \varphi_R - \varphi_L$$

$$\Delta_\lambda = \Delta_0 e^{i\varphi_\lambda}, \quad \varphi_\lambda = \lambda \frac{\varphi}{2}.$$

$$\mathcal{U}(x) = \delta(x)V_0\tau_z.$$



For $\text{Im } z > 0$

$$d(z, \varphi) = -V_0\tau_z + i\frac{k^{(+)}}{m}\tau_z + i\frac{k^{(-)}}{m}\frac{z - \Delta_0 \cos \frac{\varphi}{2}\tau_x}{z\sqrt{1 - (\frac{\Delta_0}{z})^2}},$$

$$k^{(\pm)} = k_F \frac{\sqrt{1 + \frac{z}{\mu}\sqrt{1 - (\frac{\Delta_0}{z})^2}} \mp \sqrt{1 - \frac{z}{\mu}\sqrt{1 - (\frac{\Delta_0}{z})^2}}}{2},$$

$$k_F = \sqrt{2m\mu}.$$

$$0 = \frac{m^2}{k_F^2} \det d(\omega + i0^+, \varphi) = \left(\frac{k^{(+)} }{ik_F} + \frac{mV_0}{k_F} \right)^2 + \frac{k^{(-)2}}{k_F^2} \left(1 - \frac{\Delta_0^2}{\Delta_0^2 - \omega^2} \sin^2 \frac{\varphi}{2} \right),$$

$$D = 1/[1 + (mV_0/k_F)^2]$$

$$z \sqrt{1 - \left(\frac{\Delta_0}{z} \right)^2} \Big|_{z=\omega+i0^+} = i \sqrt{\Delta_0^2 - \omega^2},$$

$$J_{\text{ABS}}(\varphi) = \frac{2\pi}{\Phi_0} \text{th} \frac{\beta\omega_A}{2} \frac{\partial_\varphi \det d(\omega_A, \varphi)}{\partial_\omega \det d(\omega_A, \varphi)} = -\frac{2\pi}{\Phi_0} \text{th} \frac{\beta\omega_A(\varphi)}{2} \frac{d\omega_A(\varphi)}{d\varphi}$$

$$J_{\text{cont}}(\varphi) = -\frac{\sin \varphi}{\Phi_0} \int_{\Delta_0}^{\infty} d\omega \text{th} \frac{\beta\omega}{2} \frac{\Delta_0^2}{\omega^2 - \Delta_0^2} \text{Im} \frac{k^{(-)2}}{m^2 \det d(\omega + i0^+, \varphi)}.$$

$$\omega_A(\varphi) = \Delta_0 \sqrt{1 - D \sin^2 \frac{\varphi}{2}}, \quad (\text{Andreev states})$$

Proof: (????)

$$d(z, \varphi) \approx -V_0 \tau_z + \frac{ik_F}{m} \frac{z - \Delta_0 \cos \frac{\varphi}{2} \tau_x}{z \sqrt{1 - \left(\frac{\Delta_0}{z} \right)^2}},$$

$$J(\varphi) = \frac{\pi \Delta_0}{2\Phi_0} \frac{D \text{th} \left[\frac{\beta \Delta_0}{2} \sqrt{1 - D \sin^2 \frac{\varphi}{2}} \right]}{\sqrt{1 - D \sin^2 \frac{\varphi}{2}}} \sin \varphi.$$

Proof: (???)

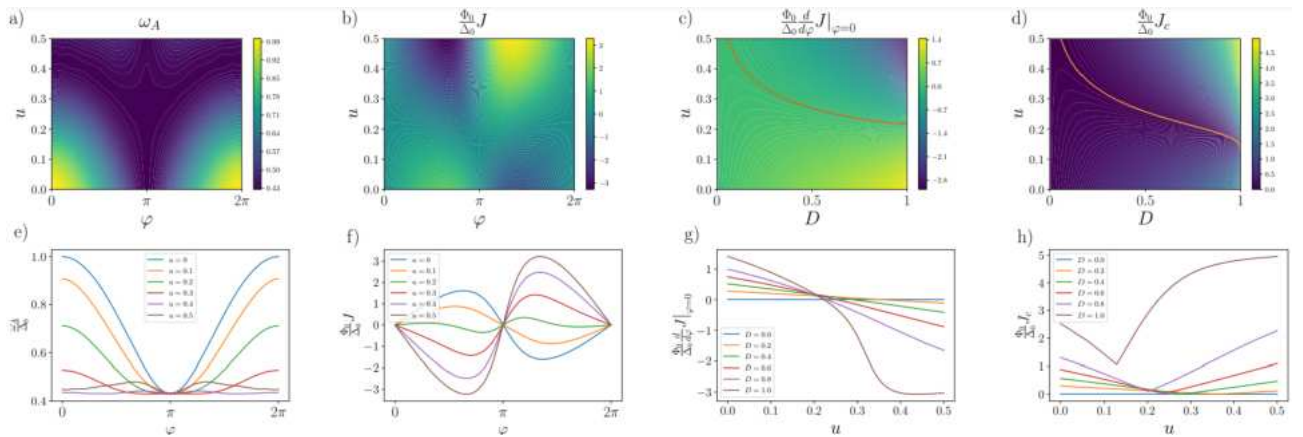


Figure 1: a) and e): $\mu = 3\Delta_0$, $D = 0.9$; b) and f): $\mu = 3\Delta_0$, $D = 0.9$; c) and g): $\mu = 3\Delta_0$, orange line - sign change of $\frac{d}{d\varphi} J(\varphi = 0)$; d) and h): $\mu = 3\Delta_0$, orange line - cusp in $J(u)$.

Disorder

$$\Sigma^\lambda(z) = U_\lambda \Sigma(z) U_\lambda^\dagger, \quad U_\lambda = e^{\frac{i}{4} \tau_z \lambda \varphi},$$

$$\Sigma(z) = n_I v_0 \tau_z \frac{1}{1 - G^{(0)}(0, 0; z) v_0 \tau_z} + O(n_I^2),$$

$$G_k^{(0, \lambda)}(z) = U_\lambda G_k^{(0)}(z) U_\lambda^\dagger \rightarrow \tilde{G}_k^{(0, \lambda)}(z) = U_\lambda \tilde{G}_k^{(0)}(z) U_\lambda^\dagger,$$

$$\tilde{G}_k^{(0)}(z) = \frac{1}{[G_k^{(0)}(z)]^{-1} - \Sigma(z)},$$

$$\tilde{G}_k^{(0)}(z) = \frac{1}{\tilde{z}(z) - \tau_z [\frac{k^2}{2m} - \tilde{\mu}(z)] - \tau_x \tilde{\Delta}_0(z)}$$

$$\tilde{z}(z) = z \left(1 - n_I v_0 k_{1,+} k_{2,+} + \frac{imv_0 k^{(-)}}{(mv_0 k^{(-)})^2 + (k_{1,+} k_{2,+} + imv_0 k^{(+)})^2} \frac{1}{z \sqrt{1 - (\frac{\Delta_0}{z})^2}} \right),$$

$$\tilde{\Delta}_0(z) = \Delta_0 \left(1 - n_I v_0 k_{1,+} k_{2,+} + \frac{imv_0 k^{(-)}}{(mv_0 k^{(-)})^2 + (k_{1,+} k_{2,+} + imv_0 k^{(+)})^2} \frac{1}{z \sqrt{1 - (\frac{\Delta_0}{z})^2}} \right),$$

$$\tilde{\mu}(z) = \mu - n_I v_0 k_{1,+} k_{2,+} + \frac{k_{1,+} k_{2,+} + imv_0 k^{(+)}}{(mv_0 k^{(-)})^2 + (k_{1,+} k_{2,+} + imv_0 k^{(+)})^2},$$

$$\frac{z}{\Delta_0} = \frac{\tilde{z}(z)}{\tilde{\Delta}_0(z)},$$

$$\frac{\tilde{k}^{(\pm)}}{k_F} = \sqrt{\frac{\tilde{\mu}(z)}{\mu}} \frac{\sqrt{1 + \frac{\tilde{z}(z)}{\tilde{\mu}(z)} \sqrt{1 - (\frac{\Delta_0}{z})^2}} \mp \sqrt{1 - \frac{\tilde{z}(z)}{\tilde{\mu}(z)} \sqrt{1 - (\frac{\Delta_0}{z})^2}}}{2}.$$

Coulomb interaction

$$V_U = U \left\{ \hat{\psi}_\uparrow^\dagger(0) \hat{\psi}_\uparrow(0) - \langle n_\uparrow(0) \rangle \right\} \left\{ \hat{\psi}_\downarrow^\dagger(0) \hat{\psi}_\downarrow(0) - \langle n_\downarrow(0) \rangle \right\}.$$

$$\begin{aligned} V_U &\simeq U \langle \hat{\psi}_\downarrow(0) \hat{\psi}_\uparrow(0) \rangle \hat{\psi}_\uparrow^\dagger(0) \hat{\psi}_\downarrow^\dagger(0) \\ &\quad + U \langle \hat{\psi}_\uparrow^\dagger(0) \hat{\psi}_\downarrow^\dagger(0) \rangle \hat{\psi}_\downarrow(0) \hat{\psi}_\uparrow(0) \\ &\quad - U \langle \hat{\psi}_\uparrow^\dagger(0) \hat{\psi}_\downarrow^\dagger(0) \rangle \langle \hat{\psi}_\downarrow(0) \hat{\psi}_\uparrow(0) \rangle \\ &= \Delta_{\text{loc}} \{ \hat{\psi}_\uparrow^\dagger(0) \hat{\psi}_\downarrow^\dagger(0) + \hat{\psi}_\downarrow(0) \hat{\psi}_\uparrow(0) \} - \frac{\Delta_{\text{loc}}^2}{U}, \end{aligned}$$

$$\Delta_{\text{loc}} = U \langle \hat{\psi}_\downarrow(0) \hat{\psi}_\uparrow(0) \rangle = U \langle \hat{\psi}_\uparrow^\dagger(0) \hat{\psi}_\downarrow^\dagger(0) \rangle$$

$$\begin{aligned} d_U &= [G_U(0,0)]^{-1} \\ &= [G(0,0)]^{-1} - \Delta_{\text{loc}} \tau_x = d - \Delta_{\text{loc}} \tau_x. \end{aligned}$$

$$\begin{aligned} J(\varphi) &= -\frac{2\pi}{\Phi_0 \beta} \sum_{i\omega_n} \text{tr} \left\{ [d_U(i\omega_n, \varphi)]^{-1} \frac{\partial d_U(i\omega_n, \varphi)}{\partial \varphi} \right\} \\ &\quad - \frac{2\pi}{\Phi_0} \frac{d}{d\varphi} \frac{\Delta_{\text{loc}}^2}{U}. \end{aligned}$$

$$\Delta_{\text{loc}} = \frac{U}{2\beta} \sum_{i\omega_n} \text{tr} \{ [d_U(i\omega_n, \varphi)]^{-1} \tau_x \}$$

$$J(\varphi) = -\frac{2\pi}{\Phi_0 \beta} \sum_{i\omega_n} \text{tr} \left\{ [d_U(i\omega_n, \varphi)]^{-1} \frac{\partial d(i\omega_n, \varphi)}{\partial \varphi} \right\}.$$

$$\Delta_{\text{loc}}(\varphi) = \frac{k_F}{m} \bar{\Delta}_{\text{loc}}(\varphi) \text{ by its}$$

$$\det d_U(\omega, \varphi) = 0.$$

Majorana wires in PPS BGF

Junction of two Majorana wires: spectral properties

$$\begin{aligned} H_\lambda &= \begin{pmatrix} h_p^{(0)} + \alpha p \sigma_z - B \sigma_x & \Delta_\lambda \\ \Delta_\lambda^* & -h_p^{(0)} - \alpha p \sigma_z - B \sigma_x \end{pmatrix}, \quad (\text{BdG}) \\ h_p^{(0)} &= \frac{p^2}{2m} - \mu; \quad \lambda = R, L \equiv +, -. \end{aligned}$$

$$\mathcal{U}(x) = \delta(x) V_0 \tau_z \sigma_0,$$

$$G^{(0,\lambda)}(x,x') = U_\lambda g(x-x') U_\lambda^\dagger, \quad U_\lambda = e^{\frac{i}{4} \tau_z \lambda \varphi},$$

$$d = \frac{\tau_z}{2m} \left\{ -2mV_0 + U_+ g'(0^+) g^{-1}(0) U_- - U_- g'(0^-) g^{-1}(0) U_+ \right\}.$$

$$\partial_\varphi d = \frac{i\tau_z}{8m} \left\{ U_+ [\tau_z g'(0^+) g^{-1}(0)] U_- + U_- [\tau_z g'(0^-) g^{-1}(0)] U_+ \right\}.$$

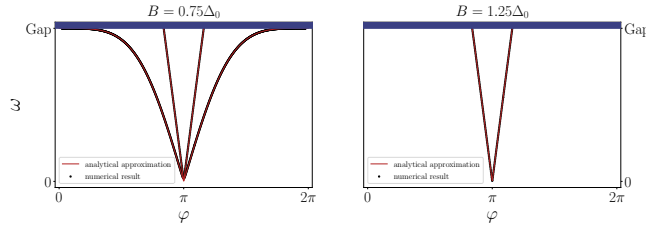


Figure 2: Short Majorana junction, $\mu = 0$, $E_{SO} = \frac{m\alpha^2}{2} = 60\Delta_0$. Left - non-topological ($B = 0.75\Delta_0 < \Delta_0$), right - topological ($B = 1.25\Delta_0 > \Delta_0$) regimes.

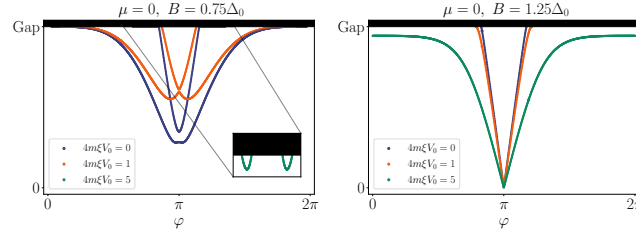


Figure 3: Short Majorana junction, $\mu = 0$, $E_{SO} = \frac{m\alpha^2}{2} = 4\Delta_0$. Left - non-topological ($B < \Delta_0$), right - topological ($B > \Delta_0$) regimes. $\xi := 1/\sqrt{2m\Delta_0}$.

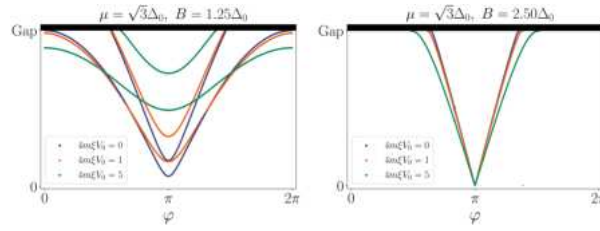


Figure 4: Short Majorana junction, $\mu = \sqrt{3}\Delta_0$, $E_{SO} = \frac{m\alpha^2}{2} = 4\Delta_0$. Left - the non-topological ($B < \sqrt{\Delta_0^2 + \mu^2}$), right - topological ($B > \sqrt{\Delta_0^2 + \mu^2}$) regimes.

Junction of two Majorana wires: Josephson current

(пока это вообще не нужно)

$$\lim_{T \rightarrow 0^+} \lim_{\omega \rightarrow 0^+} \text{th} \left[\frac{\omega}{2k_B T} \right] = 0,$$

$$\lim_{\omega \rightarrow 0^+} \lim_{T \rightarrow 0^+} \text{th} \left[\frac{\omega}{2k_B T} \right] = 1.$$

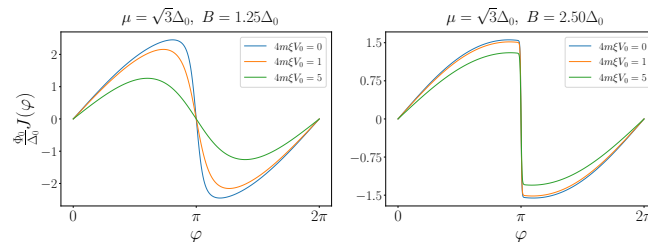


Figure 5: $T = 0$, $E_{SO} = \frac{m\alpha^2}{2} = 4\Delta_0$; left - non-topological ($B = 1.25\Delta_0 < \sqrt{\Delta_0^2 + \mu^2} = 2\Delta_0$), right - topological ($B = 2.50\Delta_0 > 2\Delta_0$), phases.

Long junction of two Majorana wires

$$H_C = \tau_z h_p^{(0,C)} + \tau_z \alpha_C p \sigma_z - B_C \sigma_x,$$

$$h_p^{(0)} = \frac{p^2}{2m_C} - \mu_C; \quad \lambda = R, L \equiv +, -.$$

$$\det \begin{pmatrix} 1 - \hat{p}_0 g^C(0) & -\hat{p}_0 g^C(-W) e^{\frac{i}{2} \tau_z \varphi} \\ -\hat{p}_W g^C(W) e^{-\frac{i}{2} \tau_z \varphi} & 1 - \hat{p}_W g^C(0) \end{pmatrix} = 0,$$

$$\hat{p}_{0,W} = \pm \frac{\tau_z}{2m} g'(0^\mp) [g(0)]^{-1} \mp \frac{\tau_z}{2m} g^{C'}(0^\mp) [g^C(0)]^{-1} \pm \frac{i}{2} (\alpha - \alpha_C) \tau_z \sigma_z$$

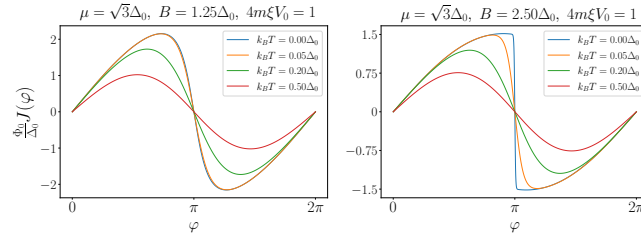


Figure 6: For $E_{\text{SO}} = \frac{m\alpha^2}{2} = 4\Delta_0$. Left - non-topological ($B = 1.25\Delta_0 < \sqrt{\Delta_0^2 + \mu^2} = 2\Delta_0$), right - topological phases ($B = 2.50\Delta_0 > 2\Delta_0$), respectively.

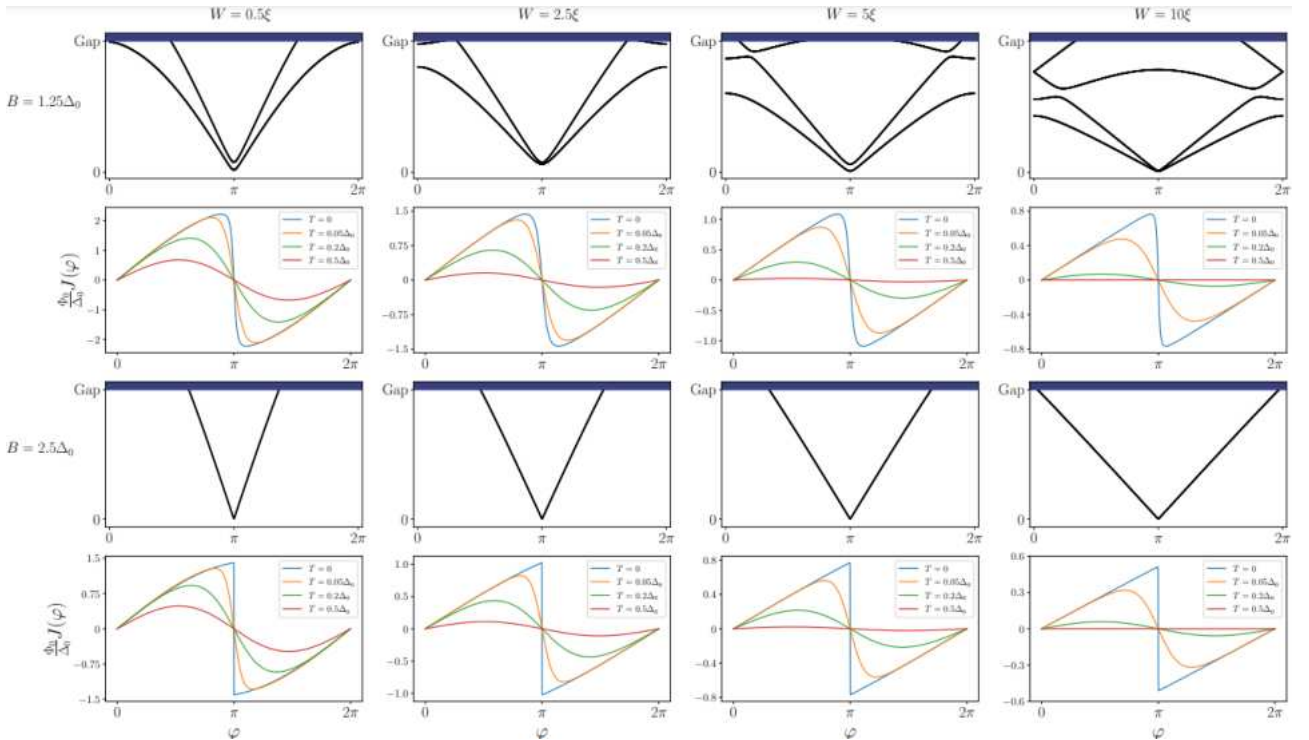


Figure 7: $\alpha_C = \alpha = 4\Delta_0\xi$, $\mu_C = \mu = \sqrt{3}\Delta_0$, $B_C = B$; $B = 1.25\Delta_0 < \sqrt{\Delta_0^2 + \mu^2} = 2\Delta_0$ - non-topological phase, $B = 2.5\Delta_0 > 2\Delta_0$ - topological.

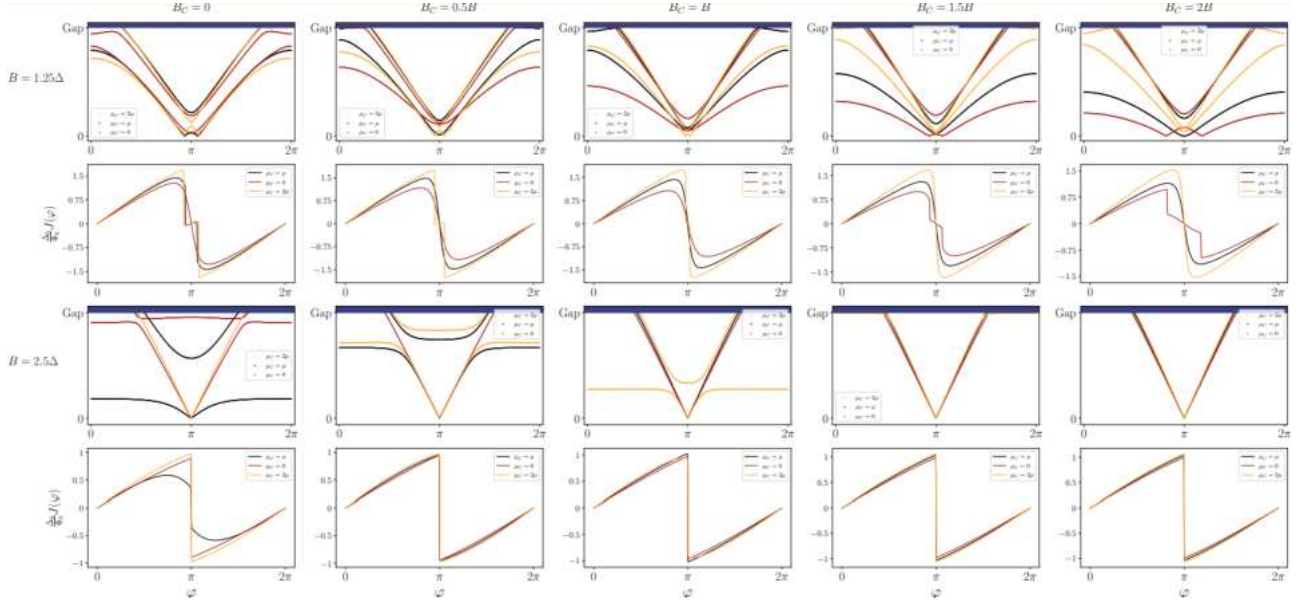


Figure 8: $W = 2.5\xi$; $\alpha_C = \alpha = 4\Delta_0\xi$; $B = 1.25\Delta_0$ - non-topological phase, $B = 2.5\Delta_0$ - topological.

1.1.2 Caroli, Comberscot's et al. Tunneling I, II: Formalism, Free electrons

Main formulas

Caroli's Formalism

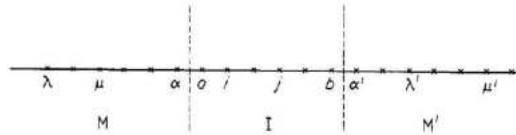
1. Introduction

$$H_{\text{eff}} = H_L + H_R + \mathbf{T} \quad (1)$$

2. Model Hamiltonian and Green function at zero voltage

2.1. The model Hamiltonian

$$H = \sum_{l,l'} T_{ll'}^{nn'} C_l^{(n)+} C_{l'}^{(n')} \quad (2)$$



$$H = \sum_{i\mu} T_{i\mu} C_i^+ C_\mu + \sum_{\lambda\mu'} T_{\lambda\mu'} C_{\lambda'}^+ C_{\mu'} + \sum_{ij} T_{ij} C_i^+ C_j + (T_{\alpha\mu} C_\alpha^+ C_\alpha + \text{hc}) + (T_{\alpha'b} C_{\alpha'}^+ C_b + \text{hc}) \quad (3)$$

2.2. The Green functions at zero bias

$$T_{xa} = T \quad T_{\alpha'b} = T' \quad (4)$$

$$\epsilon G_{pq}(\epsilon) - \sum_{p_1} T_{pp_1} G_{p_1 q}(\epsilon) = \delta_{pq} \quad (5)$$

$$G_{\lambda\mu}(\epsilon) = g_{\lambda\mu}(\epsilon) + g_{\lambda\alpha}(\epsilon) T G_{aa}(\epsilon) T g_{x\mu}(\epsilon)$$

$$\begin{aligned} G_{\lambda'\mu'}(\epsilon) &= g_{\lambda'\mu'}(\epsilon) + g_{\lambda'\alpha'}(\epsilon) T' G_{bb}(\epsilon) T' g_{\alpha'\mu'}(\epsilon) \\ G_{\lambda\mu'}(\epsilon) &= g_{\lambda\alpha}(\epsilon) T G_{ab}(\epsilon) T' g_{\alpha'\mu'}(\epsilon) \end{aligned} \quad (6)$$

$$\begin{aligned} \Sigma_{aa} &= \Sigma = T^2 g_{\alpha x} \\ \Sigma_{bb} &= \Sigma' = T'^2 g_{\alpha'\alpha'} \end{aligned} \quad (7)$$

$$G_{ij}(\epsilon) = g_{ij}(\epsilon) + g_{ia}(\epsilon) \Sigma(\epsilon) G_{aj}(\epsilon) + g_{ib}(\epsilon) \Sigma'(\epsilon) G_{bj}(\epsilon) \quad (8)$$

$$\begin{aligned} G_{ab} &= \frac{g_{ab}}{D} \\ G_{aa} &= \frac{g_{aa}(1 - T'^2 g_{bb} g_{\alpha'\alpha'}) + T'^2 g_{ba} g_{ab} g_{\alpha'x'}}{D} \\ G_{bb} &= \frac{g_{bb}(1 - T^2 g_{aa} g_{\alpha x}) + T^2 g_{ba} g_{ab} g_{\alpha x}}{D} \end{aligned} \quad (9)$$

$$D = (1 - T^2 g_{aa} g_{\alpha x}) (1 - T'^2 g_{bb} g_{\alpha'\alpha'}) - T^2 T'^2 g_{ab} g_{ba} g_{\alpha x} g_{\alpha'\alpha'} \quad (10)$$

$$G_{\lambda\lambda'} = \sum_{\mu\mu'} g_{i\mu\mu'} \mathcal{T}_{\mu\mu'} g_{\mu'\lambda'} \quad (11)$$

$$\mathcal{T}_{\mu\mu'} = \frac{T T' G_{ab}}{D} \delta_{\mu\alpha'} \delta_{\mu'\alpha'} \quad (12)$$

$$G_{\lambda\mu} = g_{\lambda\mu} + g_{\lambda\alpha} \mathcal{T} g_{\alpha'\alpha'} \mathcal{T} g_{\alpha\mu} = g_{\lambda\mu} + g_{\lambda\alpha} \frac{T^2 T'^2 g_{ab} g_{ba} g_{\alpha'\alpha'}}{D^2} g_{\alpha\mu} \quad (13)$$

3. The tunneling current

3.1. The current operator

$$J_P = \sum_{\substack{l \geq i+1 \\ m \leq i}} A_{lm} C_l^+ C_m - \sum_{\substack{l \leq i \\ m \geq i+1}} A_{lm} C_l^+ C_m \quad (14)$$

$$J_P - J_{P'} + \frac{\partial \rho_i}{\partial t} = 0 \quad (15)$$

$$\frac{\partial \rho_i}{\partial t} = \frac{1}{i\hbar} [\rho_i, H] = \frac{e}{i\hbar} \sum_m (T_{im} C_i^+ C_m - T_{mi} C_m^+ C_i) \quad (16)$$

$$A_{lm} = \frac{e}{i\hbar} T_{lm} \quad (17)$$

$$J = \langle J_P \rangle = \frac{eT}{i\hbar} (\langle C_\alpha^+ C_a \rangle - \langle C_a^+ C_\alpha \rangle) \quad (18)$$

3.2. Calculation of the current

$$\mu' - \mu = eV \quad (19)$$

$$\sum_i V_i C_i^+ C_i \quad (20)$$

$$J(t) = V(t)G(t)$$

$$\begin{aligned} G_{pq}^a(t, t') &= -i\theta(t - t') \langle [C_p(t), C_q^+(t')]_+ \rangle \\ G_{pq}^r(t, t') &= i\theta(t' - t) \langle [C_p(t), C_q^+(t')]_+ \rangle \\ F_{pq}(t, t') &= -i \langle [C_p(t), C_q^+(t')]_- \rangle \end{aligned} \quad (21)$$

$$J = \frac{eT}{\hbar} \{G_{\alpha a}^c(t, t_+) - G_{u\alpha}^c(t, t_+)\} \quad (22)$$

$$G_{pq}^c(t, t') = -i \langle T \{C_p(t) C_q^+(t')\} \rangle \equiv \frac{1}{2} \{G_{pq}^a(t, t') + g_{pq}^r(t, t') + f_{pq}(t, t')\}$$

$$\begin{aligned} G_{pq}^a(\omega) &= g_{pq_1}^a(\omega) \{ \delta_{q_1 q} + \Sigma_{q_1 q_2} G_{q_2 q}^a(\omega) \} \\ G_{pq}^r(\omega) &= g_{pq_1}^r(\omega) \{ \delta_{q_1 q} + \Sigma_{q_1 q_2}^r G_{q_2 q}(\omega) \} \\ F_{pq}(\omega) &= f_{pq}(\omega) + g_{pq_1}^r(\omega) \Sigma_{q_1 q_2} F_{q_2 q}(\omega) + f_{pq_1}(\omega) \Sigma_{q_1 q_2} G_{q_2 q}^a(\omega) \end{aligned} \quad (24)$$

$$\Sigma_{pq} = T(\delta_{px'} \delta_{aq} + \delta_{pa} \delta_{\alpha q}) + T'(\delta_{pb} \delta_{\alpha' q} + \delta_{px'} \delta_{bq}) \quad (25)$$

$$\begin{aligned} G_{\alpha a}^a &= g_{\alpha\alpha}^a T G_{aa}^a \\ G_{ax}^a &= G_{aa}^a T g_{\alpha\alpha}^a \end{aligned} \quad (26)$$

$$J = \frac{eT}{2\hbar} \{F_{xa}(t, t_+) - F_{ax}(t, t_+)\} \quad (27)$$

$$\begin{aligned} F_{\alpha a} &= f_{\alpha a} T G_{aa}^a + g_{\alpha a}^r T F_{aa} \\ F_{aa} &= f_{aa} + f_{aa} T G_{aa}^a + f_{ab} T' G_{\alpha' a}^a + g_{aa}^r T F_{\alpha a} + g_{ab}^r T' F_{\alpha' a} \\ F_{\alpha' a} &= f_{\alpha' \alpha'} T' G_{ba}^a + g_{\alpha' \alpha'}^a T' F_{ba} \\ F_{ba} &= f_{ba} + f_{bb} T' G_{\alpha' a}^a + f_{ba} T G_{aa}^a + g_{bb}^r T' F_{\alpha' a} + g_{ba}^r T F_{\alpha a}. \end{aligned} \quad (28)$$

$$\begin{aligned} g_{ij}^a &= g_{ij}^r = g_{ij}^c = g_{ij} \\ f_{ij} &= 0 \end{aligned} \quad (29)$$

$$F_{xa} = TT'^2 f_{x'\alpha'} G_{ba}^a G_{ab}^r g_{\alpha\alpha}^r + \frac{T f_{\alpha\alpha}}{D^a D^r} (1 - T'^2 g_{\alpha'\alpha'}^a g_{bb}) (1 - T'^2 g_{bb} g_{\alpha'\alpha'}^r) g_{aa} + \\ + TT'^2 G_{ab}^a G_{ba}^r f_{\alpha\alpha} g_{\alpha'\alpha'}^a (1 - T'^2 g_{x'\alpha'}^r g_{bb}) \quad (30)$$

$$F_{\alpha a}(\omega) - F_{a\alpha}(\omega) = TT'^2 G_{ba}^a(\omega) G_{ab}^r(\omega) \left\{ (g_{\alpha x}^r(\omega) - g_{\alpha x}^a(\omega)) f_{\alpha'\alpha'}(\omega) \right. \\ \left. + f_{\alpha\alpha}(\omega) (g_{\alpha'\alpha'}^a(\omega) - g_{\alpha'x'}(\omega)) \right\} \quad (31)$$

$$G_{pq}^+(t, t') = i \langle C_q^+(t') C_p(t) \rangle \\ G_{pq}^-(t, t') = -i \langle C_p(t) C_q^+(t') \rangle \quad (32)$$

$$g^+ + g^- = f \\ g^+ - g^- = g^r - g^a \quad (33)$$

$$F_{\alpha a}(\omega) - F_{a\alpha}(\omega) = 2TT'^2 G_{ba}^a(\omega) G_{ab}^r(\omega) \{ g_{\alpha\alpha}^+(\omega) g_{\alpha'\alpha'}^-(\omega) - g_{\alpha a}^-(\omega) g_{\alpha'x'}^+(\omega) \} \quad (34)$$

$$g_{zx}^+(\omega) = 2i\theta(\mu - \omega) \text{Im} g_{\alpha x}(\omega) = 2i\pi\theta(\mu - \omega) \rho_{\alpha}(\omega) \\ g_{\alpha x}^-(\omega) = -2i\pi\theta(\omega - \mu) \rho_{\alpha}(\omega) \quad (35)$$

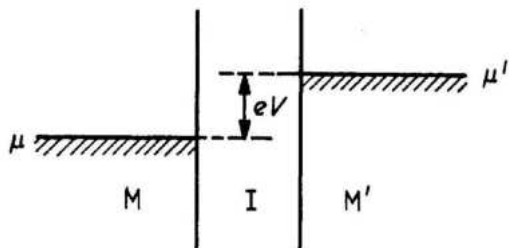
$$J = - \frac{(2\pi)^2 e T^2 T'^2}{\hbar} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} G_{ba}^a(\omega) G_{ab}^r(\omega) \rho_{\alpha}(\omega) \rho_{\alpha'}(\omega) \{ \theta(\omega - \mu) \theta(\mu' - \omega) - \theta(\mu - \omega) \theta(\omega - \mu') \} \quad (36)$$

$$J = - \frac{(2\pi)^2 e T^2 T'^2}{\hbar} \int_{\mu}^{\mu+eV} \frac{d\omega}{2\pi} G_{ba}^a(\omega) G_{ab}^r(\omega) \rho_{\alpha}(\omega) \rho_{\alpha'}(\omega) \quad (37)$$

3.3. Discussion

$$|\mathcal{J}|^2 = T^2 T'^2 G_{ba}^a(\omega) G_{ab}^r(\omega) \quad (38)$$

$$H_{\text{IM}} = \sum_{aa} T_{aa} C_{\alpha}^{\dagger} C_a + \text{hc} \quad (39)$$



$$J = \frac{ie}{\hbar} \sum_{\alpha a} (T_{\alpha a} C_{\alpha}^{+} C_a - T_{a\alpha} C_a C_{\alpha}) \quad (40)$$

$$T(\mathbf{k}_{\parallel}) = \sum_{\alpha a} T_{\alpha a} \exp \{i\mathbf{k}_{\parallel} \cdot (\mathbf{R}_{\alpha} - \mathbf{R}_a)\} \quad (41)$$

$$G_{pq}(\mathbf{k}_{\parallel}, \omega) = \sum_{\mathbf{p}_1, \mathbf{q}_1} G_{\mathbf{R}_p, \mathbf{R}_q}(\omega) \exp \{i\mathbf{k}_{\parallel} \cdot (\mathbf{p}_{\parallel} - \mathbf{q}_{\parallel})\}$$

$$J = \int \frac{d^2 \mathbf{k}_{\parallel}}{(2\pi)^2} \frac{(2\pi) e T^2(\mathbf{k}_{\parallel}) T'^2(\mathbf{k}_{\parallel})}{\hbar} \int_{\mu}^{\mu+eV} \frac{d\omega}{2\pi} G_{ba}^a(\mathbf{k}_{\parallel}, \omega) G_{ab}^r(\mathbf{k}_{\parallel}, \omega) \rho_{\alpha}(\mathbf{k}_{\parallel}, \omega) \rho_{\alpha'}(\mathbf{k}_{\parallel}, \omega) \quad (42)$$

$$\rho_{\alpha}(\mathbf{k}_{\parallel}, \omega) = -\frac{1}{\pi} \text{sgn}(\omega - \mu) \text{Im } g_{\alpha\alpha}^c(\mathbf{k}_{\parallel}, \omega) \quad (43)$$

Appendix

$$\langle \phi_0 | T \{ A(t) B(t') \dots \} | \phi_0 \rangle$$

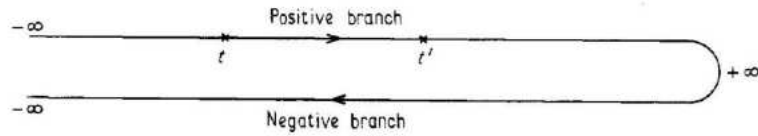
$$\langle \phi_0 | S(-\infty, +\infty) T \{ \tilde{A}(t) \tilde{B}(t') \dots S(+\infty, -\infty) \} | \phi_0 \rangle$$

$$S(t', t) = T \left[\exp \left\{ -i \int_t^{t'} H_{\text{int}}(t_1) dt_1 \right\} \right]$$

$$S(+\infty, -\infty) \phi_0 = e^{ix} \phi_0$$

$$\langle \phi_0 | T \{ A(t) B(t') \dots \} | \phi_0 \rangle = \frac{\langle \phi_0 | T \{ \tilde{A}(t) \tilde{B}(t') \dots S(+\infty, -\infty) \} | \phi_0 \rangle}{\langle \phi_0 | S(+\infty, -\infty) | \phi_0 \rangle}$$

$$\langle \phi_0 | S(-\infty, +\infty) T \{ \tilde{A}(t) \tilde{B}(t') \dots S(+\infty, -\infty) \} | \phi_0 \rangle = \langle \phi_0 | T_c \{ \tilde{A}(t) \tilde{B}(t') \dots S_c \} | \phi_0 \rangle \quad (\text{A.1})$$



$$S_c = S(-\infty, +\infty) S(+\infty, -\infty)$$

$$\begin{aligned} G_{ij}^+(t, t') &= -i \langle \phi_0 | T_c \{ c_i(t_+) c_j^+(t'_-) \} | \phi_0 \rangle \\ &= i \langle \phi_0 | c_j^+(t') c_i(t) | \phi_0 \rangle \end{aligned}$$

$$\begin{aligned} G_{ij}^-(t, t') &= -i \langle \phi_0 | T_c \{ c_i(t_-) c_j^+(t'_+) \} | \phi_0 \rangle \\ &= -i \langle \phi_0 | c_i(t) c_j^+(t') | \phi_0 \rangle \end{aligned}$$

$$\begin{aligned} G_{ij}^c(t, t') &= -i \langle \phi_0 | T_c \{ c_i(t_+) c_j^+(t'_+) \} | \phi_0 \rangle \\ &= -i \langle \phi_0 | T \{ c_i(t) c_j^+(t') \} | \phi_0 \rangle \end{aligned}$$

$$\begin{aligned} \tilde{G}_{ij}^c(t, t') &= -i \langle \phi_0 | T_c \{ c_i(t_-) c_j^+(t'_+) \} | \phi_0 \rangle \\ &= -i \langle \phi_0 | \tilde{T} \{ c_i(t) c_j^+(t') \} | \phi_0 \rangle \end{aligned}$$

$$\mathbf{G} = \begin{pmatrix} G^c & G^- \\ G^+ & \tilde{G}^c \end{pmatrix}$$

$$G_{ij}(t, t') = G_{ij}^0(t, t') + \int G_{ik}^0(t, t_1) \Sigma_{kl}(t_1, t_2) G_{ij}(t_2, t') dt_1 dt_2 \quad (\text{A.2})$$

$$\Sigma = \begin{pmatrix} \Sigma^c & \Sigma^- \\ \Sigma^+ & \tilde{\Sigma}^c \end{pmatrix}$$

$$G^c + \tilde{G}^c = G^+ + G^-$$

$$\Sigma^c + \tilde{\Sigma}^c = -(\Sigma^+ + \Sigma^-)$$

$$\mathbf{G} \rightarrow \frac{1 - i\sigma_y}{2} \begin{pmatrix} G^c & G^- \\ G^+ & \tilde{G}^c \end{pmatrix} \frac{1 + i\sigma_y}{2} = \begin{pmatrix} 0 & G^a \\ G^r & F \end{pmatrix} \quad (\text{A.3})$$

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

$$G^a = G^c - G^+ = -\tilde{G}^c + G^-$$

$$G^r = G^c - G^- = -\tilde{G}^c + G^+$$

$$F = G^c + \tilde{G}^c = G^+ + G^-$$

$$\begin{pmatrix} \Sigma^c & \Sigma^- \\ \Sigma^+ & \tilde{\Sigma}^c \end{pmatrix} \text{ becomes } \begin{pmatrix} \Omega & \Sigma^r \\ \Sigma^a & 0 \end{pmatrix}$$

$$\Sigma^a = \Sigma^c + \Sigma^+ = -(\tilde{\Sigma}^c + \Sigma^-)$$

$$\Sigma^r = \Sigma^c + \Sigma^- = -(\tilde{\Sigma}^c + \Sigma^+)$$

$$\Omega = \Sigma^c + \tilde{\Sigma}^c = -(\Sigma^+ + \Sigma^-)$$

$$H_{\text{ini}}(t) = \{ T C_\alpha^+(t) C_a(t) + T' C_{\alpha'}^+(t) C_b(t) \} + \text{hc}$$

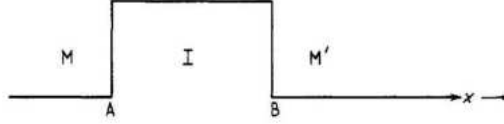
$$\sigma = \begin{pmatrix} 0 & T \\ T & 0 \end{pmatrix} (\delta_{p\alpha} \delta_{qa} + \delta_{pu} \delta_{q\alpha}) + \begin{pmatrix} 0 & T' \\ T' & 0 \end{pmatrix} (\delta_{p\alpha'} \delta_{qb} + \delta_{pb} \delta_{q\alpha'})$$

Tunneling current II: Free electron description by Caroli et. al.

1. Introduction

$$H = \sum_{ij} T_{ij} C_i^+ C_j \quad (1)$$

$$H = \int dx \psi^+(x) \left\{ -\frac{\hbar^2}{2m} \nabla^2 + V(x) \right\} \psi(x) \quad (2)$$



2. Expression of the current in the continuous representation

2.1. The Green function at zero bias

$$\int H(x, x') \Phi(x') dx' = E \Phi(x) \quad (3)$$

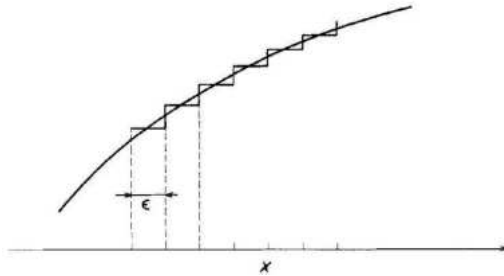
$$H(x, x') = -\frac{\hbar^2}{2m} \delta''(x - x') + V(x) \delta(x - x') \quad (4)$$

$$\int dx'' \{ \omega \delta(x - x'') - H(x, x'') \} \mathbf{G}_\omega(x'', x') = \delta(x - x') \quad (5)$$

$$H^{(\epsilon)}(x, x') = -\frac{\hbar^2}{2m\epsilon^2} \{ \delta(x - x' - \epsilon) - 2\delta(x - x') + \delta(x - x' + \epsilon) \} + V(x) \delta(x - x') \quad (6)$$

$$H(x, x') = \lim_{\epsilon \rightarrow 0} H^{(\epsilon)}(x, x') \quad (7)$$

$$-\frac{\hbar^2}{2m\epsilon^2} \{ \Phi^{(\epsilon)}(x + \epsilon) - 2\Phi^{(\epsilon)}(x) + \Phi^{(\epsilon)}(x - \epsilon) \} + V(x) \Phi^{(\epsilon)}(x) = E \Phi^{(\epsilon)}(x) \quad (8)$$

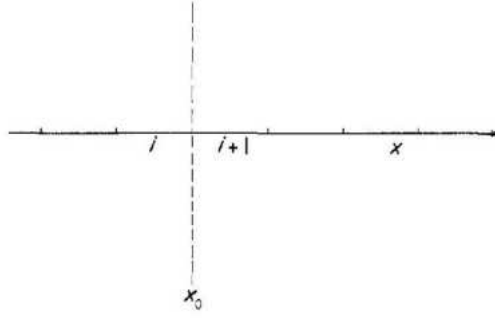


$$-\frac{\hbar^2}{2m\epsilon^2} \{ \Phi_{p+1}^{(\epsilon)} - 2\Phi_p^{(\epsilon)} + \Phi_{p-1}^{(\epsilon)} \} + V_p \Phi_p^{(\epsilon)} = E \Phi_p^{(\epsilon)} \quad (9)$$

$$\sum_{p''} (\omega \delta_{pp''} - H_{pp''}) \mathbf{G}_{p''p'}^{(\epsilon)} = \frac{\delta_{pp'}}{\epsilon} \quad (10)$$

$$H_{pp''} = \frac{\hbar^2}{2m\epsilon^2} (2\delta_{pp''} - \delta_{p,p''-1} - \delta_{p,p''+1}) + V_p \delta_{pp''} \quad (11)$$

$$G = \epsilon G^{(\epsilon)}.$$



$$\sum_{p''} (\omega \delta_{pp''} - h_{pp''}) g_{p''p'} = \delta_{pp'} \quad (12)$$

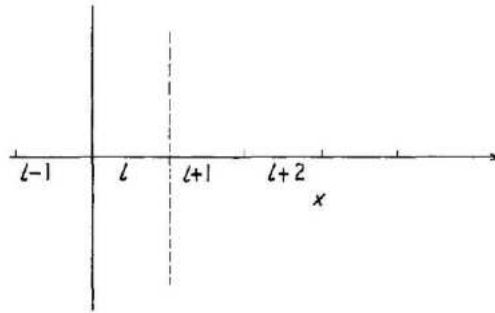
$$h_{pp''} = \begin{cases} H_{pp''} & \text{if } p \text{ and } p'' \geq i+1 \\ 0 & \text{if } p \text{ or } p'' < i+1. \end{cases}$$

$$T = -\hbar^2/2m\epsilon^2$$

$$G_{lk} = \begin{cases} \frac{\tilde{g}_{li} T g_{i+1,k}}{D} & l \leq i; k \geq i+1; \\ \tilde{g}_{lk} + \frac{\tilde{g}_{li} T g_{i+1,i+1} T \tilde{g}_{ik}}{D} & l, k \leq i \end{cases}$$

$$G_{ii} = \frac{\tilde{g}_{ii}}{D}; \quad G_{i,i+1} = \frac{\tilde{g}_{ii} T g_{i+1,i+1}}{D} \quad (13)$$

$$D = 1 - T^2 \tilde{g}_{ii} g_{i+1,i+1}.$$



$$g_l^0 = \frac{1}{\omega + 2T - V_l} \quad (14)$$

$$g_l = \frac{g_l^0}{1 - T^2 g_l^0 g_{l+1}^0} \quad (15)$$

$$\frac{1}{g_i} = \omega + 2T - V_l - T^2 g_{l+1} \quad (16)$$

$$g_l = \frac{1}{T} \{1 + \epsilon \gamma_l + \epsilon^2 \delta_l + O(\epsilon^3)\} \quad (17)$$

$$T \{-\epsilon \gamma_l + \epsilon^2 (\gamma_l^2 - \delta_l)\} = \omega - V_l - T (\epsilon \gamma_{l+1} + \epsilon^2 \delta_{l+1}) + O(\epsilon^3). \quad (18)$$

$$-\frac{\hbar^2}{2m} \left(\frac{d\gamma}{dx} + \gamma^2 \right) = \omega - V(x) \quad (19)$$

$$\tilde{g}_l = \frac{1}{T} \{1 + \epsilon \tilde{\gamma}_l + O(\epsilon^3)\} \quad (20)$$

$$\lim_{\epsilon \rightarrow 0} \tilde{\gamma}_i = \tilde{\gamma}(x),$$

$$\tilde{\gamma}(x):$$

$$-\frac{\hbar^2}{2m} \left(\tilde{\gamma}^2 - \frac{d\tilde{\gamma}}{dx} \right) = \omega V(x) \quad (21)$$

$$G(x_{0-}, x_{0+}) = \lim_{\epsilon \rightarrow 0} G_{ii}^{(\epsilon)} = + \frac{2m}{\hbar^2} \frac{1}{\gamma(x_0) + \tilde{\gamma}(x_0)} \quad (22)$$

$$G(x_{0-}, x_{0+}) = \lim_{\epsilon \rightarrow 0} G_{i,i+1}^{(\epsilon)} = G(x_{0-}, x_{0+}) \quad (23)$$

$$g_i(il) \xrightarrow{\epsilon \rightarrow 0} \frac{1}{T} \exp \int_{x_i}^{x_i} \gamma(t) dt \quad (24)$$

$$\tilde{g}_i(il) \xrightarrow{\epsilon \rightarrow 0} \frac{1}{T} \exp \int_x^{x_i} \tilde{\gamma}(t) dt \quad (25)$$

$$G(x_1, x_2) = \frac{2m}{\hbar^2 \{ \gamma(x_0) + \tilde{\gamma}(x_0) \}} \exp \left\{ \int_{x_1}^{x_0} \tilde{\gamma}(t) dt + \int_{x_0}^{x_2} \gamma(t) dt \right\} \quad (26)$$

and if $x_1, x'_1 < x_0$

$$G(x_1, x'_1) = \tilde{g}(x_1, x'_1) + \frac{2m}{\hbar^2 \{ \gamma(x_0) + \tilde{\gamma}(x_0) \}} \exp \left\{ \int_{x_0}^{x_1} \tilde{\gamma}(t) dt + \int_{x_0}^{x'_1} \tilde{\gamma}(t) dt \right\} \quad (27)$$

$$g_\omega(x, x') = \sum_{\alpha} \frac{\Psi_{\alpha}^*(x) \Psi_{\alpha}(x')}{\omega - E_{\alpha}} \quad (28)$$

$$g_i^{(\epsilon)}(ii) \simeq g(x_0 + \epsilon, x_0 + \epsilon) \simeq -\frac{2m\epsilon}{\hbar^2} \{1 + \epsilon\gamma(x_0)\}. \quad (29)$$

$$\Psi_{\alpha}(x_0) = 0,$$

$$\Psi_x(x_0 + \epsilon) \simeq \epsilon \Psi'_x(0),$$

$$\gamma(x_0) = -\frac{\hbar^2}{2m} \lim_{x \rightarrow x_0} \sum_{\alpha} \frac{\Psi_{\alpha}^*(x_0) \Psi'_{\alpha}(x)}{\omega - E_{\alpha}} = \lim_{x \rightarrow x_0} \lim_{x' \rightarrow x_0} \left(-\frac{\hbar^2}{2m} \right) \frac{\partial}{\partial x} \frac{\partial}{\partial x'} g(x, x'). \quad (30)$$

2.2 Formal expression of the current

$$\langle J \rangle = \frac{e\hbar'}{2m} \lim_{x' \rightarrow x} \left(\frac{\partial}{\partial x} - \frac{\partial}{\partial x'} \right) G^+(xt, x't) \Big|_{x=x_0} \quad (31)$$

$$G^+(xt, x't') = i \langle \Psi^+(x't') \Psi(xt) \rangle \quad (32)$$

$$\langle J \rangle = \frac{e\hbar}{2m} \int \frac{\{ G_{\omega}^{+(\epsilon)}(x_0 + \epsilon, x_0) - G_{\omega}^{+(\epsilon)}(x_0, x_0 + \epsilon) \}}{\epsilon} \frac{d\omega}{2\pi} = -\frac{eT}{\hbar} \int \frac{d\omega}{2\pi} (G_{10}^+ - G_{01}^+) \quad (33)$$

$$\hat{G}_{pq} = \begin{pmatrix} G_{pq}^c & G_{pq}^- \\ G_{pq}^+ & \tilde{G}_{pq}^c \end{pmatrix} \quad \hat{\Sigma}_{pq} = \begin{pmatrix} \Sigma_{pq}^c & \Sigma_{pq}^- \\ \Sigma_{pq}^+ & \tilde{\Sigma}_{pq}^c \end{pmatrix} \quad (34)$$

$$\begin{aligned}
G^a &= G^c - G^+ = -\tilde{G}^c + g^- \\
G^r &= G^c - G^- = -\tilde{G}^c + g^+ \\
\Sigma^a &= \Sigma^c + \Sigma^+ = -(\tilde{\Sigma}^c + \Sigma^-) \\
\Sigma^r &= \Sigma^c + \Sigma^- = -(\tilde{\Sigma}^c + \Sigma^+)
\end{aligned} \tag{35}$$

$$\hat{\Sigma}_{pq} = T \hat{1} (\delta_{p0} \delta_{q1} + \delta_{p1} \delta_{q0}). \tag{36}$$

$$\hat{G}_{pq} = \hat{g}_{pq} + \sum_{lm} \hat{g}_{pl} \Sigma_{lm} \hat{G}_{mq} \tag{37}$$

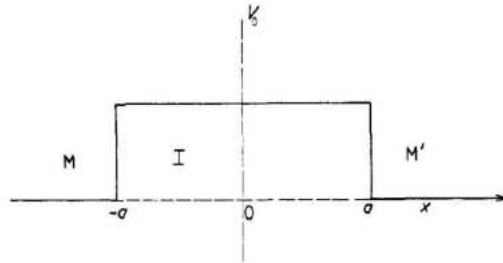
$$T [G_{10}^+ - G_{01}^+] = \frac{T}{1 - T^2 g_{11}^a g_{00}^a} [g_{00}^+ g_{11}^- - g_{00}^- g_{11}^+] \frac{T}{1 - T^2 g_{11}^r g_{00}^r} \tag{38}$$

$$\begin{aligned}
g_{00}^a &= (g_{00}^r)^* = \frac{\tilde{g}_{00}}{\tilde{g}^*} \quad \text{for } \omega > \mu_L \\
&= 2i \operatorname{Im} \tilde{g}_{00} = 0 \quad \text{for } \omega > \mu_L
\end{aligned} \tag{39}$$

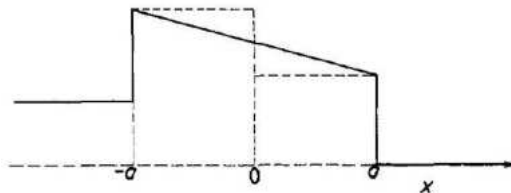
$$\langle J \rangle = \frac{8e}{\hbar} \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \frac{\operatorname{Im} \{\gamma(x_0)\} \operatorname{Im} \{\tilde{\gamma}(x_0)\}}{|\gamma(x_0) + \tilde{\gamma}(x_0)|^2} \left[\Theta(\mu_L - \omega) \Theta(\omega - \mu_R) - \Theta(\mu_R - \omega) \Theta(\omega - \mu_L) \right] \tag{40}$$

$$\begin{aligned}
\frac{d}{dx} \left\{ \frac{\operatorname{Im} \gamma(x) \operatorname{Im} \tilde{\gamma}(x)}{|\gamma(x) + \tilde{\gamma}(x)|^2} \right\} &= 0 \\
\langle J \rangle &= \frac{8e}{\hbar} \int_{\mu}^{\mu+eV} \frac{d\omega}{2\pi} \frac{\operatorname{Im} \gamma(x_0) \operatorname{Im} \tilde{\gamma}(x_0)}{|\gamma(x_0) + \tilde{\gamma}(x_0)|^2}
\end{aligned} \tag{41}$$

3. The current through a square barrier



$$\gamma_{\omega}(0) = \tilde{\gamma}_{\omega+eV}(0). \tag{42}$$



$$-\frac{\hbar^2}{2m} \left[\frac{d\gamma}{dx} + \gamma^2 \right] = \begin{cases} \omega & x > a \\ \omega - V_0 & x < a. \end{cases} \quad (43)$$

$$\gamma_\omega(x) = \begin{cases} (-\tilde{\omega})^{1/2} \text{th} \left\{ x(-\tilde{\omega})^{1/2} + b_1 \right\} & x > a \\ (\tilde{V}_0 - \tilde{\omega})^{1/2} \text{th} \left\{ x(\tilde{V}_0 - \tilde{\omega})^{1/2} + b_2 \right\} & |x| < a \end{cases}$$

$$\tilde{\omega} = \frac{2m\omega}{\hbar^2} \quad \tilde{V}_0 = \frac{2mV_0}{\hbar^2}$$

$$\gamma_\omega(x) = (-\tilde{\omega})^{1/2} \quad x > a$$

$$(V_0 - \tilde{\omega})^{1/2} \text{th} \left\{ a(\tilde{V}_0 - \tilde{\omega})^{1/2} + b_2 \right\} = (-\tilde{\omega})^{1/2} \quad (44)$$

$$\gamma(0) = (\tilde{V}_0 - \tilde{\omega})^{1/2} \text{th} b_2 = (\tilde{V}_0 - \tilde{\omega})^{1/2} \frac{(\tilde{V}_0 - \tilde{\omega})^{1/2} \text{th} \left\{ a(\tilde{V}_0 - \tilde{\omega})^{1/2} \right\} - (-\tilde{\omega})^{1/2}}{(-\tilde{\omega})^{1/2} \text{th} \left\{ a(\tilde{V}_0 - \tilde{\omega})^{1/2} \right\} - (\tilde{V}_0 - \tilde{\omega})^{1/2}} \quad (45)$$

$$\gamma_\omega(0) = \tilde{\gamma}_\omega(0),$$

$$\langle J \rangle = \frac{8e}{\hbar} \int_\mu^{\mu+eV} \frac{d\omega}{2\pi} \frac{\omega(V_0 - \omega)}{4\omega(V_0 - \omega) + V_0^2 \text{sh}^2 \left[2a \left\{ 2m(V_0 - \omega) / \hbar^2 \right\}^{1/2} \right]} \quad (46)$$

$$2a \left\{ 2m(V_0 - \omega) / \hbar^2 \right\}^{1/2} \gg 1$$

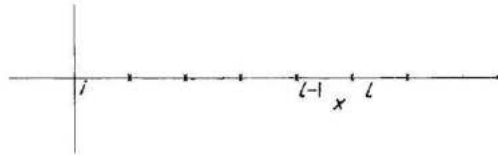
$$\langle J \rangle = \frac{8e}{\hbar} \int_\mu^{\mu+eV} \frac{d\omega}{2\pi} \frac{4\omega(V_0 - \omega)}{V_0^2} \exp \left[-4a \left\{ \frac{2m}{\hbar^2} (V_0 - \omega) \right\}^{1/2} \right] \quad (47)$$

Appendix

$$g_i(il) = g_i(i, l-1)Tg_l(ll) \quad (A1)$$

$$\gamma_\omega(x) = \sqrt{(\tilde{V}_\infty - \tilde{\omega})^{1/2}}.$$

$$g_i(il) - g_i(i, l-1) = \{Tg_l(ll) - 1\} g_i(i, l-1)_{\epsilon \rightarrow 0} \rightarrow \epsilon_i g_i(i, l-1) \quad (A2)$$



$$g_i(il) \xrightarrow{\epsilon \rightarrow 0} \frac{1}{T} \exp \int_{x_i}^{x_l} \gamma(t) dt \quad (A3)$$

$$g_i(jk) = g_{i+1}(jk) + g_{i+1}(j, i+1)Tg_i(ii)Tg_{i+1}(i+1, k) \quad (A4)$$

$$g_{i+1}(i+1, k) \simeq g_{i+1}(j, i+1) + \frac{1}{T} \quad (A5)$$

$$g_i(jk) \simeq g_{i+1}(j, k) + \frac{1}{T} \quad (\text{A6})$$

$$g_i(i + p, k) \simeq \frac{p + 1}{T}$$

$$g_i(j, i + q) \simeq \frac{q + 1}{T}$$

$$g_i(j, k) \simeq \{\min(j, k) - i + 1\} \frac{1}{T} \quad (\text{A7})$$

$$g(x, x') = \lim_{\epsilon \rightarrow 0} g_i^{(\epsilon)}(jk)$$

$$g_i(j, k) \simeq -\{\min(x - x_0, x' - x_0) + \epsilon\} \frac{2m\epsilon}{\hbar^2} \quad (\text{A8})$$

$$g(x, x') \simeq -\frac{2m}{\hbar^2} \min(x' - x_0, x' - x_0) \quad (\text{A9})$$

$$g_\omega(x, x') = \sum_\alpha \frac{\Psi_\alpha^*(x) \Psi_\alpha(x')}{\omega - E_\alpha} \quad (\text{A10})$$

$$\Psi_\alpha(x_0) = 0 \quad (\text{A11})$$

$$g_\omega(x_0 + \eta, x') \simeq \eta \sum_\alpha \frac{\Psi_\alpha^*(x) \Psi_\alpha(x')}{\omega - E_\alpha} \quad (\text{A12})$$

$$\sum_k \frac{k \sin kx'}{\omega - \hbar^2 k^2 / 2m}$$

$$\frac{\partial}{\partial x'} y(x_0 + \epsilon, x') = \gamma(x') g(x_0 + \epsilon, x') \quad (\text{A13})$$

$$\lim_{x' \rightarrow x_0} g(x_0 + \epsilon, x') = 1/T\epsilon$$

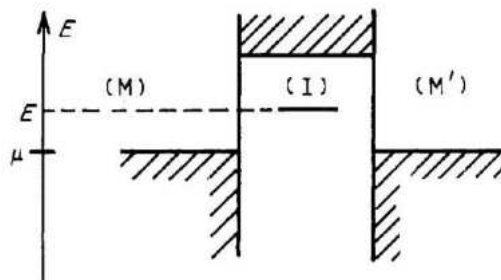
$$\lim_{x' \rightarrow x_0+} \epsilon \sum_x \frac{\Psi_\alpha'^*(x_0) \Psi_\alpha'(x')}{\omega - E_z} = \gamma(x_0) \frac{1}{\epsilon T} \quad (\text{A14})$$

$$\gamma(x_0) = \lim_{x' \rightarrow x_0+} \lim_{x \rightarrow x_0+} -\frac{\hbar^2}{2m} \frac{\partial}{\partial x'} \frac{\partial}{\partial x} g_\omega(x, x') \quad (\text{A15})$$

1.1.3 Caroli, Comberscot's et al. Tunneling III, IV: impurities, Electron-phonons

Impurities

1. Introduction



2. One dimensional model

2.1. Localized functions formalism

$$H = \sum_{ij} T_{ij} c_i^+ c_j \quad (1)$$

$$H = H_M + H_{M'} + H_I + H_c + V_{\text{imp}} \quad (2)$$

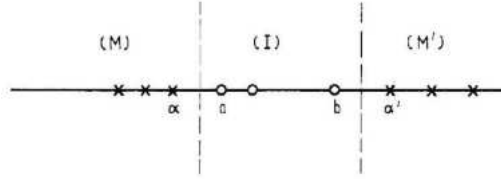
$$H_M = \sum_{\lambda, \mu} T_{\lambda\mu} c_\lambda^+ c_\mu \quad (2a)$$

$$H_{M'} = \sum_{\lambda', \mu'} T_{\lambda'\mu'} c_{\lambda'}^+ c_{\mu'} \quad (2b)$$

$$H_I = \sum_{pq} T_{pq} c_p^+ c_q \quad (2c)$$

$$H_c = T (c_\alpha^+ c_a + c_a^+ c_\alpha) + T' (c_{\alpha'}^+ c_b + c_b^+ c_{\alpha'}) \quad (3)$$

$$T = T_{\alpha a} = T_{a\alpha} \quad T' = T_{\alpha' b} = T_{b\alpha'}$$



$$V_{\text{imp}} = v c_O^+ c_O \quad (4)$$

$$J = \frac{e}{i\hbar} T (c_a^+ c_\alpha - c_\alpha^+ c_a) \quad (5)$$

$$H' = H_c + V_{\text{imp}} \quad (6)$$

$$\mathbf{G}_{ij} = \begin{pmatrix} G_{ij}^c & G_{ij}^+ \\ G_{ij}^- & \tilde{G}_{ij}^c \end{pmatrix} \quad (7)$$

$$\begin{aligned} G_{ij}^c(t, t') &= -i \langle T \{ c_i(t) c_j^+(t') \} \rangle \\ \tilde{G}_{ij}^c(t, t') &= -i \langle \tilde{T} \{ c_i(t) c_j^+(t') \} \rangle \\ G_{ij}^+(t, t') &= i \langle c_j^+(t') c_i(t) \rangle \\ G_{ij}^-(t, t') &= -i \langle c_i(t) c_j^+(t') \rangle \end{aligned} \quad (8)$$

$$\Sigma_{ij} = \begin{pmatrix} \Sigma_{ij}^c & \Sigma_{ij}^+ \\ \Sigma_{ij}^- & \tilde{\Sigma}_{ij}^c \end{pmatrix} \quad (9)$$

$$\mathbf{G}_{ij} = \mathbf{g}_{ij} + \sum_{kl} \mathbf{g}_{ik} \Sigma_{kl} \mathbf{G}_{lj} \quad (10)$$

$$\begin{aligned}
G_{ij}^r &= (G^c - G^+)_{ij} = (-\tilde{G}^c + g^-)_{ij} \\
G_{ij}^a &= (G^c - G^-)_{ij} = (-\tilde{G}^c + g^+)_{ij} \\
\Sigma_{ij}^r &= (\Sigma^c - \Sigma^+)_{ij} = (-\tilde{\Sigma}^c + \Sigma^-)_{ij} \\
\Sigma_{ij}^a &= (\Sigma^c - \Sigma^-)_{ij} = (-\tilde{\Sigma}^c + \Sigma^+)_{ij}
\end{aligned} \tag{11}$$

$$G_{ij}^\alpha = g_{ij}^\alpha + \sum_{kl} g_{ik}^\alpha \Sigma_{kl}^\alpha G_{lj}^\alpha \quad (\alpha = a \text{ or } r) \tag{12a}$$

$$G_{ij}^\pm = \sum_{kl} \left(\delta_{ik} + \sum_m G_{im}^r \Sigma_{mk}^r \right) g_{kl}^\pm \left(\delta_{lj} + \sum_m \Sigma_{lm}^a G_{mj}^a \right) + \sum_{kl} G_{ik}^r \Sigma_{kl}^\pm G_{lj}^a \tag{12b}$$

$$\begin{aligned}
\langle J \rangle &= \frac{eT}{\hbar} \{ G_{ax}^+(t, t_+) - G_{za}^+(t, t_+) \} \\
&= \frac{eT}{\hbar} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \{ G_{a\alpha}^+(\omega) - G_{\alpha a}^+(\omega) \}
\end{aligned} \tag{13}$$

2.2. Formal expression of the current

$$\Sigma^\pm = 0, \Sigma^a = \Sigma^r = \Sigma$$

$$\Sigma_{pq} = T \left(\delta_{px} \dot{\delta}_{qa} + \delta_{pa} \delta_{q\alpha} \right) + T' \left(\dot{\delta}_{p\alpha'} \delta_{bq} + \delta_{pb} \delta_{q\alpha'} \right) + v \delta_{p0} \delta_{q0} \tag{14}$$

$$G_{a\alpha}^+ - G_{\alpha a}^+ = T \left(g_{\alpha\alpha}^+ G_{aa}^- - g_{\alpha\bar{\alpha}}^- G_{aa}^+ \right) \tag{15}$$

$$g_{\frac{1}{ij}}^\pm = g_{i'j}^\pm = 0.$$

$$G_{aa}^\pm = G_a^r T g_{\alpha\alpha}^\pm T G_{aa}^a + g_{ab}^r T' g_{\alpha'x'}^\pm T' G_{ba}^a \tag{16}$$

$$\langle J \rangle = \frac{(2\pi)^2 e T^2 T'^2}{\hbar} \int_{\mu'}^{\mu' + eV} \frac{d\omega}{2\pi} |G_{ab}^r(\omega)|^2 \rho_\alpha(\omega) \rho_{\alpha'}(\omega) \tag{17}$$

$$\mu = \mu' + eV$$

$$\begin{aligned}
g_{\alpha\alpha}^+(\omega) &= 2i\pi\theta(\mu - \omega)\rho_\alpha(\omega) \\
g_{\alpha\alpha}^-(\omega) &= -2i\pi\theta(\omega - \mu)\rho_\alpha(\omega)
\end{aligned} \tag{18}$$

$$(\tilde{g}^r)_{pp'}^{-1} \tilde{g}_{p'p''}^\pm (\tilde{g}^a)_{p''q}^{-1}$$

$$g_{p'p''}^\pm(\omega) = a_{p'p''}^\pm \dot{\partial}(\omega - E) \tag{19}$$

$$(\tilde{g}^{a,r})^{-1} = \omega - H_I - V_{\text{imp}} \pm i\delta.$$

$$\sum_{p'p''} (\tilde{g}^r)_{pp'}^{-1} \tilde{g}_{p'p''}^\pm (\tilde{g}^a)_{p''q}^{-1} \propto \sum_{p'p''} (\omega - H_I - V_{\text{imp}})_{pp'} \delta(\omega - H_I - V_{\text{imp}})_{p'p''} (\omega - H_I - V_{\text{imp}})_{p''q} \equiv 0 \tag{20}$$

2.3. The transmission coefficient

$$eV - \left(E + (eV) \frac{x}{d} \right) = eV \left(1 - \frac{x}{d} \right) - E$$

$$\frac{d\langle J \rangle}{dV} = \frac{2\pi e T^2 T'^2}{\hbar} \rho_\alpha \rho_{\alpha'} |G_{ab}^r(\mu' + eV)|^2 \quad (21)$$

$$G_{ab}^r = G_{ab}^{(0)r} \Lambda_{\text{imp}}^r \quad (22)$$

$$G_{ab}^{(0)r} = \frac{g_{ab}}{(1 - T^2 g_{aa} g_{\alpha\alpha}^r) (1 - T'^2 g_{bb} g_{\alpha'\alpha'}^r) - T^2 T'^2 g_{ab} g_{ba} g_{\alpha\alpha}^r g_{\alpha'\alpha'}^r} \quad (23)$$

$$\Lambda_{\text{imp}}^r = \frac{1 + v g_{ab}^{-1} (g_{ao} g_{ob} - g_{oo} g_{ab})}{1 - v (g_{oo} + \lambda^r)} \quad (24)$$

$$\lambda^r = \frac{T^2 (g_{ao})^2 g_{\alpha\alpha}^r}{1 - T^2 g_{\alpha\alpha}^r g_{aa}} + \frac{T'^2 (g_{bo})^2 g_{\alpha'\alpha'}^r}{1 - T'^2 g_{\alpha'\alpha'}^r g_{bb}} \quad (25)$$

$$\frac{d\langle J \rangle}{dV} = \left(\frac{d\langle J \rangle}{dV} \right)^{(0)} |\Lambda_{\text{imp}}^r|^2 \quad (26)$$

$$v^{-1} - g_{oo}(\omega) = \lambda^r(E)$$

$$\omega \simeq E - \lambda^r(E) \left(\frac{dg_{oo}(\omega)}{d\omega} \Big|_{\omega=E} \right)^{-1} = \omega_0 \quad (27)$$

$$\Delta = (\text{Im } \hat{\lambda}^r) \left(\frac{dg_{oo}(\omega)}{d\omega} \Big|_{\omega=E} \right)^{-1} \quad (28)$$

$$\text{Im } \hat{\lambda}^r = \pi \rho_\alpha \frac{T^2 (g_{ao})^2}{|1 - T^2 g_{\alpha\alpha}^r g_{aa}|^2} + \pi \rho_\beta \frac{T'^2 (g_{bo})^2}{|1 - T'^2 g_{\alpha'\alpha'}^r g_{bb}|^2}$$

$$G = G_0 \frac{1}{v^2 \left(\frac{dg_{oo}}{d\omega} \Big|_{\omega=E} \right)^2 \left\{ (\mu' + eV - \tilde{E})^2 + \Delta^2 \right\}} \quad (29)$$

$$\tilde{E} = E - \left(\frac{dg_{oo}}{d\omega} \Big|_{\omega=E} \right)^{-1} \text{Re } \hat{\lambda}^r$$

3. Generalization to three dimensions

$$H_c = \sum_{\alpha\alpha} T_{\alpha\alpha} (c_\alpha^+ c_\alpha + c_\alpha c_\alpha^+) + \sum_{\alpha'\text{b}} T_{\alpha'\text{b}} (c_{\alpha'}^+ c_\text{b} + c_\text{b}^+ c_{\alpha'}) \quad (30)$$

$$T(\mathbf{k}_\parallel) = \sum_{\alpha\text{a}} T_{\alpha\text{a}} \exp \{ i \mathbf{k}_\parallel \cdot (\mathbf{R}_\alpha - \mathbf{R}_\text{a}) \}$$

$$G_{ij}^{(0)}(\mathbf{k}_\parallel, \omega) = \sum_{i_\parallel j_\parallel} G_{\mathbf{R}_i, \mathbf{R}_j}^{(0)}(\omega) \exp \{ i \mathbf{k}_\parallel \cdot (\mathbf{i}_\parallel - \mathbf{j}_\parallel) \} \quad (31)$$

$$J = \int \frac{d^2 \mathbf{k}_{\parallel}}{(2\pi)^2} \frac{(2\pi)^2 e}{\hbar} \frac{\{T(\mathbf{k}_{\parallel})\}^2 \{T'(\mathbf{k}_{\parallel})\}^2}{\hbar} \int_{\mu}^{\mu'+eV} \frac{d\omega}{2\pi} \left| G_{\text{ab}}^{(0)\text{r}}(\mathbf{k}_{\parallel}, \omega) \right|^2 \rho_{\alpha}(\mathbf{k}_{\parallel}, \omega) \cdot \rho_{\alpha'}(\mathbf{k}_{\parallel}, \omega) \quad (32)$$

$$J = \int \frac{d^2 \mathbf{k}_{\parallel}}{(2\pi)^2} \frac{d^2 \mathbf{k}'_{\parallel}}{(2\pi)^2} \frac{(2\pi)^2 e}{\hbar} \frac{\{T(\mathbf{k}_{\parallel})\}^2 \{T'(\mathbf{k}'_{\parallel})\}^2}{\hbar} \int_{\mu'}^{\mu'+eV} \frac{d\omega}{2\pi} \left| G_{\text{ab}}^{\text{r}}(\mathbf{k}_{\parallel}, \mathbf{k}'_{\parallel}, \omega) \right|^2 \cdot \rho_z(\mathbf{k}_{\parallel}, \omega) \rho_{\alpha'}(\mathbf{k}'_{\parallel}, \omega) \quad (33)$$

$$G_{ij}^{\text{r}}(\mathbf{k}_{\parallel}, \mathbf{k}'_{\parallel}, \omega) = \sum_{i_{\parallel}, j_{\parallel}} G_{\mathbf{R}_i, \mathbf{R}_j}^{\text{r}}(\omega) \exp \left\{ i(\mathbf{k}_{\parallel} \cdot \mathbf{i}_{\parallel} - \mathbf{k}'_{\parallel} \cdot \mathbf{j}_{\parallel}) \right\} \quad (34)$$

$$\left\langle \left| G_{\text{ab}}^{\text{r}}(\mathbf{k}_{\parallel}, \mathbf{k}'_{\parallel}, \omega) \right|^2 \right\rangle_{\text{av}}$$

$$G_{\mathbf{R}_a, \mathbf{R}_b}^{\text{r}}(\omega) = G_{\mathbf{R}_n, \mathbf{R}_b}^{(0)}(\omega) + \sum_{\mathbf{R}_p} G_{\mathbf{R}_a, \mathbf{R}_p}^{(0)\text{r}}(\omega) \mathcal{T}_p^{\text{r}} G_{\mathbf{R}_p, \mathbf{R}_b}^{(0)\text{r}}(\omega) \quad (35)$$

$$\mathcal{T}_{\mathbf{R}_i, \mathbf{R}_j}^{\text{r}} = \mathcal{T}_p^{\text{r}} \delta_{\mathbf{R}_i, \mathbf{R}_j} \delta_{\mathbf{R}_j, \mathbf{R}_p} \quad (36)$$

$$\mathcal{T}_p^{\text{r}} = v / \left\{ 1 - v G_{\mathbf{R}_p}^{(0)\text{r}}(\omega) \right\}$$

$$G_{\text{ab}}^{\text{r}}(\mathbf{k}_{\parallel}, \mathbf{k}'_{\parallel}, \omega) = G_{\text{ab}}^{(0)\text{r}}(\mathbf{k}_{\parallel}) \delta_{\mathbf{k}, \mathbf{k}'_{\parallel}} + \sum_{\mathbf{R}_p} \exp \left\{ i(\mathbf{k}_{\parallel} - \mathbf{k}'_{\parallel}) \cdot \mathbf{p}_{\parallel} \right\} G_{\text{ap}}^{(0)\text{r}}(\mathbf{k}_{\parallel}) \mathcal{T}_p^{\text{r}} G_{p \text{ b}}^{(0)\text{r}}(\mathbf{k}'_{\parallel}) \quad (37)$$

$$\begin{aligned} \left\langle \left| G_{\text{ab}}^{\text{r}}(\mathbf{k}_{\parallel}, \mathbf{k}'_{\parallel}, \omega) \right|^2 \right\rangle_{\text{av}} &= \left| G_{\text{ab}}^{(0)\text{r}}(\mathbf{k}_{\parallel}, \omega) \right|^2 \delta_{\mathbf{k}_1} \cdot \mathbf{k}' + \delta_{\mathbf{k}_{\parallel}, \mathbf{k}'_{\parallel}} \left\{ G_{\text{ab}}^{(\text{a})}(\mathbf{k}_{\parallel}) \sum_p c_p G_{\text{ap}}^{(0)\text{a}}(\mathbf{k}_{\parallel}, \omega) \right. \\ &\quad \times \mathcal{T}_p^{\text{a}} G_{p \text{ b}}^{(0)\text{a}}(\mathbf{k}_{\parallel}, \omega) + \text{h.c.} \left. \right\} + \sum_p c_p \left| \mathcal{T}_p^{\text{r}} G_{\text{ap}}^{(0)\text{r}}(\mathbf{k}_{\parallel}, \omega) G_{p \text{ b}}^{(0)\text{r}}(\mathbf{k}'_{\parallel}, \omega) \right|^2 \end{aligned} \quad (38)$$

$$\mathcal{T}_p^{\text{r}}(\omega) \simeq \frac{\alpha_p}{\omega - E_p + i\Delta_p} \quad (39)$$

$$\alpha_p \sim \left[\frac{dg_{pp}(\omega)}{d\omega} \Big|_{\omega=E_p} \right]^{-1} \sim \epsilon^2$$

$$G_{\text{ap}} G_{p \text{ b}} \sim G_{\text{ab}} G_{pp}$$

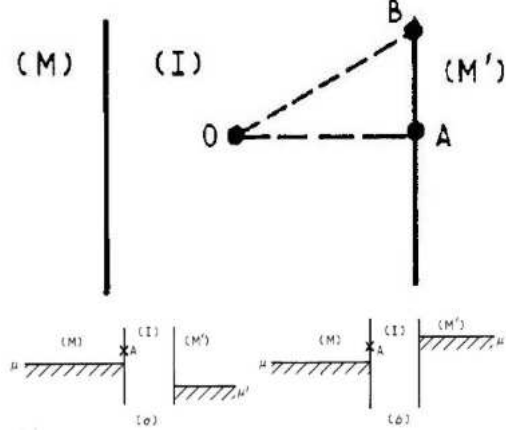
$$\mathcal{T}_p G_{pp} \sim \frac{\epsilon}{\Delta_p} \gg 1$$

$$\langle J \rangle_{\text{av}} = J^{(0)} + \frac{(2\pi)^2 e}{\hbar} \int_{\mu'}^{\mu'+eV} \frac{d\omega}{2\pi} \sum_p c_p \left| \mathcal{T}_p^{\text{T}}(\omega) \right|^2 U_{\text{ap}}(\omega) U'_{p \text{ b}}(\omega) \quad (40)$$

$$U_{\text{ap}}(\omega) = \int \frac{d^2 \mathbf{k}_{\parallel}}{(2\pi)^2} \{T(\mathbf{k}_{\parallel})\}^2 \left| G_{\text{ap}}^{(0)\text{r}}(\mathbf{k}_{\parallel}, \omega) \right|^2 \rho_a(\mathbf{k}_{\parallel}, \omega)$$

$$\frac{d\langle J \rangle_{av}}{dV} = \sum_p c_p \left| \mathcal{T}_p^r(\mu' + eV) \right|^2 A_p \quad (41)$$

$$\frac{d\langle J \rangle_{av}}{dV} = \alpha c \frac{\epsilon^2}{(\omega - \tilde{E}_0)^2 + \Delta_0^2}$$



Tunneling current IV: effects by Caroli et. al.

1. Introduction

$$H_1 = \sum_{ij} T_{ij} c_i^\dagger c_j \quad (1)$$

$$H_2 = \sum_{ijl} V_{ijl} c_i^\dagger c_j \phi_l + cc \quad (2)$$

$$H_2 = \sum_i V_i c_i^\dagger c_i \phi_i + cc$$

$$H_c = T (c_x^\dagger c_a + c_a^\dagger c_x) \quad (3)$$

$$J = \frac{ie}{\hbar} T (c_x^\dagger c_a - c_a^\dagger c_x) \quad (4)$$

2. Perturbation formalism

$$G_{ij}^+(t, t') = i \langle c_j^\dagger(t') c_i(t) \rangle$$

$$G_{ij}^-(t, t') = -i \langle c_i(t) c_j^\dagger(t') \rangle$$

$$G_{ij}^r(t, t') = -i \theta(t - t') \langle [c_i(t), c_j^\dagger(t')] \rangle_+$$

$$G_{ij}^a(t, t') = i \theta(t' - t) \langle [c_i(t), c_j^\dagger(t')] \rangle_+$$

$$D_{ij}^+(t, t') = -i \langle \phi_j(t') \phi_i(t) \rangle = D_{ji}^-(t', t)$$

$$D_{ij}^r(t, t') = -i \theta(t - t') \langle [\phi_i(t), \phi_j(t')] \rangle_-$$

$$D_{ij}^a(t, t') = i \theta(t' - t) \langle [\phi_i(t), \phi_j(t')] \rangle_-$$

$$G^r = G_0^r + g_0^r \Sigma^r G^r$$

$$G^\pm = (1 + g^r \Sigma^r) G_0^\pm (1 + \Sigma^a G^a) + g^r \Sigma^\pm G^a \quad (5a)$$

$$(AB)^\pm = A^t B^a + A^r B^\pm \quad (5b)$$

$$(AB)^r = A^r B^r$$

$$T \Sigma_{ij}^r = {}^r \Sigma_{ij}^a = T_{ij}$$

$$T \Sigma_{ij}^+ = T \Sigma_{ij}^- = 0$$

$$\begin{aligned} {}^{\text{ph}} \Sigma_{ij}^\pm(\omega) &= iV_i V_j \int \frac{d\omega'}{2\pi} D_{ij}^\pm(\omega - \omega') G_{ij}^\pm(\omega') \\ {}^{\text{ph}} \Sigma_{ij}^r(\omega) &= iV_i V_j \int \frac{d\omega'}{2\pi} \{ D_{ij}^r(\omega - \omega') G_{ij}^-(\omega') + D_{ij}^+(\omega - \omega') G_{ij}^r(\omega') \} \end{aligned} \quad (6)$$

$$\langle J \rangle = \frac{e}{\hbar} \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} T \{ G_{az}^+(\omega) - G_{aa}^{+\gamma(\omega)} \} \quad (7)$$

$$G_{za}^r = g_{\alpha a}^r T G_{aa}^r = G_z^r T g_{aa}^r \quad (8)$$

$$\begin{aligned} G_{xa}^+ &= g_{\alpha\alpha}^+ T G_{aa}^a + g_{x\alpha}^r T G_{aa}^+ \\ G_{a\alpha}^+ &= G_{aa}^r T g_{x\alpha}^+ + g_{aa}^+ T g_{x\alpha}^a \end{aligned} \quad (9)$$

$$\langle J \rangle = \frac{e}{\hbar} \int_{-\infty}^{\infty} T^2 \{ g_{\alpha\alpha}^+(\omega) G_{aa}^-(\omega) - g_{\alpha\alpha}^-(\omega) G_{aa}^+(\omega) \} \frac{d\omega}{2\pi} \quad (10)$$

$$G_{aa} = g_{aa} + g_{aa} T g_{\alpha x} T g_{aa} + \dots$$

$$G_{aa}^\pm = \Lambda^r (g_{aa}^\pm + g_{aa}^r T g_{\alpha a}^\pm T g_{aa}^a) \Lambda^a \quad (11)$$

$$\Lambda = \frac{1}{1 - T^2 g_{aa} g_{\alpha x}} \quad (12)$$

$$\langle J \rangle = \frac{e}{\hbar} \int |T \Lambda^r(\omega)|^2 [g_{x\alpha}^+(\omega) g_{aa}^-(\omega) - g_{xx}^-(\omega) g_{aa}^+(\omega)] \frac{d\omega}{2\pi} \quad (13)$$

$$\begin{aligned} \rho_x &= -\frac{1}{\pi} \text{Im } g_{\alpha\alpha}^r \\ \rho_a &= -\frac{1}{\pi} \text{Im } g_{aa}^r \end{aligned} \quad (14)$$

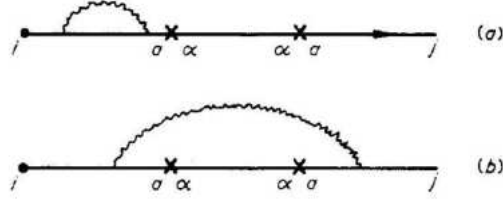
$$\begin{aligned} g_{\alpha x}^+(\omega) &= 2i\pi \rho_x(\omega) f_L(\omega) & g_{aa}^+(\omega) &= 2i\pi \rho_a(\omega) f_R(\omega) \\ g_{\alpha x}^-(\omega) &= 2i\pi \rho_x(\omega) \{f_L(\omega) - 1\} & g_{aa}^-(\omega) &= 2i\pi \rho_a(\omega) \{f_R(\omega) - 1\} \end{aligned} \quad (15)$$

$$\langle J \rangle = 2\pi \frac{e}{\hbar} \int d\omega |T\Lambda^r|^2 \rho_\alpha(\omega) \rho_a(\omega) \{f_L(\omega) - f_R(\omega)\} \quad (16)$$

$$G_{aa}^\pm = G_{aa}^r T g_{\alpha\alpha}^\pm T G_{aa}^a + g_{ab}^r T' g_{\beta\beta}^\pm T' G_{ba}^a + g_{ai}^{rph} \Sigma_{ij}^\pm G_{ja}^a \quad (17)$$

$$\langle J \rangle = \frac{e}{\hbar} \int \frac{d\omega}{2\pi} |U^r|^2 (g_{\alpha\alpha}^+ g_{\beta\beta}^- - g_{\alpha\alpha}^- g_{\beta\beta}^+) \quad (18)$$

$$U^r = T T' G_{ab}^r \quad (19)$$



$$\langle J \rangle = \frac{2\pi e}{\hbar} \int d\omega |U^r(\omega)|^2 \rho_\alpha(\omega) \rho_\beta(\omega) \{f_L(\omega) - f_R(\omega)\} \quad (20)$$

$$G_{ab}^r = \frac{g_{ab}^r}{(1 - T^2 g_{\alpha\alpha}^r g_{aa}^r) (1 - T'^2 g_{\beta\beta}^r g_{bb}^r) - T^2 T'^2 g_{ab}^r g_{ba}^r g_{\alpha\alpha}^r g_{\beta\beta}^r} \quad (21)$$

$$\langle J \rangle = \frac{e}{\hbar} \int \frac{d\omega}{2\pi} |\tilde{U}^r|^2 (\tilde{g}_{\alpha\alpha}^+ \tilde{g}_{\beta\beta}^- - \tilde{g}_{\alpha\alpha}^- \tilde{g}_{\beta\beta}^+) \quad (22)$$

$$\begin{aligned} \tilde{U}^r &= T T' g_{ab}^r \\ \tilde{g}_{x\alpha}^\pm &= \frac{g_{xz}^\pm}{|1 - T^2 g_{x\alpha}^r g_{aa}^r|^2} \end{aligned} \quad (23)$$

$$\langle J \rangle = \frac{2\pi e}{\hbar} \int d\omega |\tilde{U}^r(\omega)|^2 \tilde{\rho}_\alpha(\omega) \tilde{\rho}_\beta(\omega) \{f_L(\omega) - f_R(\omega)\} \quad (24)$$

$$\tilde{\rho}_\alpha = \frac{\rho_\alpha}{|1 - T^2 g_{xx}^r g_{aa}^r|^2} = -\frac{1}{\pi T^2 g_{aa}^r} \text{Im} \left[\frac{1}{1 - T^2 g_{\alpha\alpha}^r g_{aa}^r} \right] \quad (25)$$

$$\langle J \rangle = \frac{2\pi e}{\hbar} |U^r|^2 \int_{\mu_R}^{\mu_L} d\omega \tilde{\rho}_\alpha(\omega - \mu_L) \tilde{\rho}_\beta(\omega - \mu_R) \quad (26)$$

3. Electron-phonon interaction in the electrodes of an MIM junction

$$\begin{aligned} \frac{d}{d\omega} (\text{Im } \Sigma^r) &\sim \delta(|\omega - \mu| - \omega_0) \text{sgn}(\omega - \mu) \\ \frac{d}{d\omega} (\text{Re } \Sigma^r) &\sim \frac{2\omega_0}{\omega_0^2 - (\omega - \mu)^2} \\ g_{\alpha x}^r &= \sum_k \frac{1}{\epsilon_k - \omega - \Sigma^r(k, \omega)} \end{aligned} \quad (27)$$

$$g_{x\alpha}^r(\omega) = g_{\alpha z}^{r0} \{\omega + \Sigma^r(\omega)\} \quad (28)$$

$$\begin{aligned}
g_{xx}^r &= g_{x\alpha}^{r0} + \delta g_{\alpha x}^r \\
\delta g_{x\alpha}^r &= \sum_{\lambda, \mu} g_{x\lambda}^{r0} \Sigma_{\lambda\mu}^r g_{\mu\alpha}^{ro}
\end{aligned} \tag{29}$$

$$\begin{aligned}
\tilde{\rho}_\alpha &= \tilde{\rho}_\alpha^0 + \delta \tilde{\rho}_\alpha \\
\tilde{\rho}_\beta &= \tilde{\rho}_\beta^0 + \delta \tilde{\rho}_\beta
\end{aligned} \tag{30}$$

$$\begin{aligned}
\delta \tilde{\rho}_\alpha &= -\frac{1}{\pi} \text{Im} \frac{\delta g_{\alpha x}^r}{(1 - T^2 g_{ad} G_{x\alpha}^{r0})^2} \\
\delta \tilde{\rho}_\beta &= -\frac{1}{\pi} \text{Im} \frac{\delta g_{\beta\beta}^r}{(1 - T^2 g_{bb} g_{\beta\beta}^{ro})^2}
\end{aligned} \tag{31}$$

$$\tilde{\rho}_\alpha^0 \tilde{\rho}_\beta^0 + \tilde{\rho}_\alpha^0 \delta \tilde{\rho}_\beta + \tilde{\rho}_\beta^0 \delta \tilde{\rho}_\alpha$$

$$G = G_0 + \delta G_I + \delta G_R \tag{32}$$

$$\begin{aligned}
G^0 &= \frac{2\pi e^2}{\hbar} |U^r|^2 \tilde{\rho}_\alpha^0 \tilde{\rho}_\beta^0 \\
\delta G_I(V) &= \frac{2\pi e^2}{\hbar} |U^r|^2 \tilde{\rho}_\beta^0 \delta \tilde{\rho}_\alpha(-eV) \\
\delta G_R(V) &= \frac{2\pi e^2}{\hbar} |U^r|^2 \tilde{\rho}_\alpha^0 \delta \tilde{\rho}_\beta(+eV)
\end{aligned} \tag{33}$$

$$\delta \tilde{\rho}_\alpha = -\frac{1}{\pi A} (\cos \theta \text{Im} \delta g_{\alpha x}^r - \sin \theta \text{Re} \delta g_{\alpha\alpha}^r) \tag{34}$$

$$\delta g_{\alpha\alpha}^r = \sum_{\lambda} (g_{\alpha\lambda}^{r0}) \Sigma_{\lambda\lambda}^r \tag{35}$$

$$\begin{aligned}
\mathcal{G}_{\gamma\gamma}^r &= \mathcal{G}_{\gamma\gamma}^{r0} + \delta \mathcal{G}_{\gamma\gamma}^r \\
\delta \mathcal{G}_{\gamma\gamma}^r &= \sum_{\lambda\mu} \mathcal{G}_{\gamma\lambda}^{ro} \Sigma_{\lambda\mu}^r \mathcal{G}_{\mu\gamma}^{ro}
\end{aligned} \tag{36}$$

$$\Delta G_L = \left(\frac{e^2}{2\pi\hbar} T^2 \Lambda^2 \text{Im} \mathcal{G}_{cc}^{ro} \right) \text{Im} \delta \mathcal{G}_{\gamma\gamma}^r(-eV) \tag{37}$$

$$\mathcal{G}_{\gamma\lambda}^r = \frac{g_{\gamma a} T g_{\alpha\lambda}^r}{1 - T^2 g_{aa} g_{\alpha\lambda}^r}$$

$$\delta \mathcal{G}_{\gamma\gamma}^r = (g_{\gamma a} T)^2 \frac{g_{\alpha\lambda}^r \Sigma_{\lambda\gamma}^r g_{\mu\alpha}^r}{(1 - T^2 g_{aa} g_{\alpha\alpha}^r)^2} \tag{38}$$

$$\sum_{\lambda} \text{Im} \left\{ (\mathcal{G}_{\gamma\lambda}^{r0})^2 \Sigma_{\lambda\lambda}^r(-eV) \right\} \tag{39}$$

$$g_{\alpha\lambda}^{\text{ro}}(\omega, k_{\parallel}) \sim \exp(ik_0 z) \quad (40)$$

$$k_0 = (k_F^2 - k_{\parallel}^2)^{1/2} \quad (41)$$

$$|U^{\text{r}}|^2 \sim \exp \left\{ -2d \left(\frac{2m}{\hbar^2} V_B - k_0^2 \right)^{1/2} \right\} \quad (42)$$

$$\delta g_{\alpha\alpha}^{\pm} = \sum_{\lambda\mu} \left(g_{\alpha\lambda}^{\pm 0} \Sigma_{\lambda\mu}^{\text{a}} g_{\mu\alpha}^{\text{a}0} + g_{\alpha\lambda}^{\text{r}0} \Sigma_{\lambda\mu}^{\text{r}} g_{\mu\alpha}^{0\pm} + g_{\alpha\lambda}^{\text{r}0} \Sigma_{\lambda\mu}^{\pm} g_{\mu\alpha}^{\text{a}0} \right) \quad (43)$$

4. Electron-phonon interaction in the barrier

$$\langle J_{\text{in}} \rangle = \sum_{ij} \frac{e}{\hbar} \int \frac{d\omega}{2\pi} T^2 T'^2 G_{ai}^{\text{r}}(\omega) G_{ja}^{\text{a}}(\omega) \{ g_{\alpha\alpha}^+(\omega) \Sigma_{ij}^-(\omega) - g_{\alpha\alpha}^-(\omega) \Sigma_{ij}^+(\omega) \} \quad (44)$$

$$\begin{aligned} \langle J_{\text{in}} \rangle = \sum_{ij} \frac{ie}{\hbar} V_i V_j \iint \frac{d\omega}{2\pi} \frac{d\omega'}{2\pi} T^2 G_{ai}^{\text{r}0}(\omega) G_{ja}^{\text{a}}(\omega) \\ \times \{ g_{\alpha\alpha}^+(\omega) G_{ij}^{-0}(\omega') D_{ij}^-(\omega - \omega') - g_{\alpha\alpha}^-(\omega) G_{ij}^{+0}(\omega') D_{ij}^+(\omega - \omega') \} \end{aligned} \quad (45)$$

$$G_{ij}^{\pm 0} = G_{ia}^{\text{r}0} T g_{\alpha\alpha}^{\pm} T G_{aj}^{\text{a}0} + g_{ib}^{\text{ro}} T' g_{\beta\beta}^{\pm} T' G_{bj}^{\text{a}0} \quad (46)$$

$$\begin{aligned} G_{ai}^{\text{r}0}(\omega) G_{ja}^{\text{a}0}(\omega) G_{ia}^{\text{r}}(\omega') G_{aj}^{\text{a}0}(\omega') \{ g_{\alpha\alpha}^+(\omega) g_{\alpha\alpha}^-(\omega') D_{ij}^-(\omega - \omega') \\ - g_{\alpha\alpha}^-(\omega) g_{\alpha\alpha}^+(\omega') D_{ij}^+(\omega - \omega') \} \end{aligned} \quad (47)$$

$$D_{ij}^+(\omega) = D_{ij}^-(-\omega) \quad (48)$$

$$\langle J_{\text{tn}} \rangle = \sum_{ij} \frac{2\pi e}{\hbar} \iint d\omega d\omega' W_{ij}(\omega, \omega') F \quad (49)$$

$$\begin{aligned} W_{ij}(\omega, \omega') = T^2 T'^2 V_i V_j G_{ai}^{\text{ro}}(\omega) G_{ib}^{\text{ro}}(\omega') G_{bj}^{\text{a}}(\omega') G_{ja}^{\text{a}0}(\omega) \\ F = \frac{i}{(2\pi)^3} \{ g_{\alpha\alpha}^+(\omega) g_{\beta\beta}^-(\omega') D_{ij}^-(\omega - \omega') - g_{\alpha\alpha}^-(\omega) G_{\beta\beta}^+(\omega') D_{ij}^+(\omega - \omega') \} \end{aligned} \quad (50)$$

$$\sigma_{ij}(\omega) = \sum_n A_n^*(i) A_n(j) \delta(\omega - \omega_n) \quad (51)$$

$$\begin{aligned} D_{ij}^+(\omega) = D_{ij}^-(-\omega) = -2i\pi \sigma_{ij}(\omega) N(\omega) \\ D_{ij}^+(-\omega) = D_{ij}^-(\omega) = -2i\pi \sigma_{ij}(\omega) \{ N(\omega) + 1 \} \end{aligned} \quad (52)$$

$$N(\omega) = \frac{1}{\exp \beta \omega - 1} \quad (53)$$

$$F = \rho_{\chi}(\omega) \rho_{\beta}(\omega') \sigma_{ij}(\omega - \omega') [f_L(\omega) \{1 - f_R(\omega')\} \{1 + N(\omega - \omega')\} - \{1 - f_L(\omega)\} f_R(\omega') N(\omega - \omega')] \quad (54a)$$

$$F = \rho_g(\omega) \rho_{\beta}(\omega') \sigma_{ij}(\omega' - \omega) [f_L(\omega) \{1 - f_R(\omega')\} N(\omega' - \omega) - \{1 - f_L(\omega)\} f_R(\omega') \{1 + N(\omega' - \omega)\}] \quad (54b)$$

$$\frac{dG_{\text{in}}}{dV} = \frac{d^2 J_{\text{in}}}{dV^2} = \sum_{ij} \frac{2\pi e^3}{\hbar} W_{ij} \rho_{\alpha} \rho_{\beta} \sigma_{ij}(|eV|) \text{sgn } V \quad (56)$$

$$G_{ai}^{r0}(\omega) G_{ib}^{r0}(\omega') \quad (57)$$

$$\langle \delta J_{\text{el}} \rangle = \frac{e}{\hbar} \int \frac{d\omega}{2\pi} \delta |U^r|^2 (g_{\alpha z}^+ g_{\beta\beta}^- - g_{\alpha a}^- g_{\beta\beta}^+) \quad (58)$$

$$\delta |U^r|^2 = \sum_{ij} T^2 T^2 (G_{ai}^{r0} \Sigma_{ij}^r G_{ib}^{r0} G_{ba}^{a0} + \text{cc}) \quad (59)$$

$$\Sigma_{ij}^r(\omega) \simeq iV_i V_j \int \frac{d\omega'}{2\pi} D_{ij}^+(\omega - \omega') G_{ij}^r(\omega') \quad (60)$$

Appendix

$$H = -\frac{\hbar^2}{2m} \nabla^2 + V(x)$$

$$\left(-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) - \omega \right) \mathcal{G}^{r0}(x, y; \omega) = \delta(x - y) \quad (A.1)$$

$$\mathcal{G}^{r0}(x, y) = A_{\pm} \exp(Kx) + B_{\pm} \exp(-Kx) \quad (A.2)$$

$$\mathcal{G}^{r0}(x, y) = \alpha \exp(ikx) + \beta \exp(-ikx) \quad (A.3)$$

$$\mathcal{G}^{r0}(x, y) = \begin{cases} -\frac{2m \sinh Kx}{\hbar^2} \frac{KD}{KD} [(K + iq) \exp\{K(y - 1)\} + (K - iq) \exp\{K(a - y)\}] & x < y \\ -\frac{2m \sinh Ky}{\hbar^2} \frac{KD}{KD} [(K + iq) \exp\{K(x - a)\} + (K - iq) \exp\{K(a - x)\}] & y < x < a \\ -\frac{2m \sinh Ky}{\hbar^2} \frac{\sin}{KD} 2K \exp\{iq(x - a)\} & a < x \end{cases} \quad (A.4)$$

$$D = (K + iq) \exp(-Ka) + (K - iq) \exp(Ka) \quad (A.5)$$

$$\mathcal{G}^{r0}(x, y) = -\frac{2m}{\hbar^2} 2 \sinh Ky \frac{\exp(-Ka)}{K - iq} \exp\{iq(x - a)\} \quad (A.6)$$

$$\mathcal{G}^{r0}(x, y) \simeq -\frac{2m \sinh Ky}{\hbar^2} \frac{1}{K} \exp(-Kx) \quad (A.7)$$

1.1.4 The Keldysh method

Main formulas

2 Overview of other methods

2.0.1 BGf for spinful Majorana NW by Alvarado et. al.

Boundary Green's function by Alvarado et. al.

$$H_{\text{bulk}} = \frac{1}{2} \sum_k \hat{\Psi}_k^\dagger \hat{\mathcal{H}}(k) \hat{\Psi}_k, \quad (2.1)$$

$$\hat{G}^R(k, \omega) = [\omega + i0^+ - \hat{\mathcal{H}}(k)]^{-1}, \quad (2.2)$$

$$\hat{G}_{jj'}^R(\omega) = \frac{a}{2\pi} \int_{-\pi/a}^{\pi/a} dk e^{i(j-j')ka} \hat{G}^R(k, \omega). \quad (2.3)$$

$$\hat{G}_{jj'}^R(\omega) = \frac{1}{2\pi i} \oint_{|z|=1} \frac{dz}{z} z^{j-j'} \hat{G}^R(z, \omega). \quad (2.4)$$

$$P(z, \omega) = \det [\omega - \hat{\mathcal{H}}(z)] = \frac{1}{z^N} \prod_{n=1}^{2N} [z - z_n(\omega)], \quad (2.5)$$

$$\hat{G}_{jj'}^R(\omega) = \sum_{|z_n| < 1} \frac{z_n^{j-j'} \hat{A}(z_n, \omega)}{\prod_{m \neq n} (z_n - z_m)}, \quad (2.6)$$

$\hat{A}(z, \omega)$ is the cofactor matrix of $[\omega - \hat{\mathcal{H}}(z)] z$.

$$\hat{\mathcal{G}}_{jj}(\omega) = \hat{G}_{jj}(\omega) - \hat{G}_{j0}(\omega) [\hat{G}_{00}(\omega)]^{-1} \hat{G}_{0j}(\omega). \quad (2.7)$$

$$\hat{\mathcal{G}}_L(\omega) = \hat{\mathcal{G}}_{-1, -1}(\omega), \quad \hat{\mathcal{G}}_R(\omega) = \hat{\mathcal{G}}_{11}(\omega). \quad (2.8)$$

$$\hat{r}_{L/R} = \lim_{\omega \rightarrow 0} \frac{\mathbb{1} - i\hat{\mathcal{V}}_{\pm 1}^\dagger \hat{\mathcal{G}}_{L/R}(\omega) \hat{\mathcal{V}}_{\pm 1}}{\mathbb{1} + i\hat{\mathcal{V}}_{\pm 1}^\dagger \hat{\mathcal{G}}_{L/R}(\omega) \hat{\mathcal{V}}_{\pm 1}}. \quad (2.9)$$

- (i) Hermiticity of the BdG Hamiltonian implies that every root $z_n(\omega)$ is accompanied by a root $1/z_n^*(\omega)$, where ‘ $*$ ’ denotes complex conjugation.
- (ii) Electron-hole symmetry of the BdG Hamiltonian implies that $z_n(\omega) = z_n^*(-\omega)$. In the presence of an additional symmetry $\hat{\mathcal{H}}(k) = \hat{U} \hat{\mathcal{H}}(-k) \hat{U}^\dagger$ with a unitary matrix $\hat{U}$, for every root $z_n(\omega)$, also $z_n^*(\omega)$ must be a root.
- (iii) As a consequence of (i) and (ii), $\prod_{n=1}^{2N} z_n(\omega) = 1$.
- (iv) Topological phase transitions can occur once a pair of zero-energy roots hits the unit circle, $|z_n(0)| = 1$, which corresponds to the closing and reopening of a gap in the bulk spectrum.
- (v) Equations (15.6) and (15.7) imply that subgap bound states (with energy E) localized near the boundary of a semi-infinite wire decay into the bulk in a manner controlled by $\max(|z_n(E)| < 1)$.

Spinful single-channel hybrid nanowires by Alvarado et. al.

$$\hat{\Psi}_k^T = (c_{k\uparrow}, c_{k\downarrow}, c_{-k\downarrow}^\dagger, -c_{-k\uparrow}^\dagger),$$

$$\mathcal{H}(k) = \epsilon_k \sigma_0 \tau_z + V_x \sigma_x \tau_0 + \alpha_k \sigma_z \tau_z + \Delta \sigma_0 \tau_x, \quad (2.10)$$

$$\epsilon_k = 2t[1 - \cos(ka)] - \mu$$

$$E_{k,\pm}^2 = \Delta^2 + \alpha_k^2 + V_x^2 + \epsilon_k^2 \pm 2\sqrt{\Delta^2 V_x^2 + (\alpha_k^2 + V_x^2) \epsilon_k^2}. \quad (2.11)$$

$$V_x = V_c = \sqrt{\Delta^2 + \mu^2},$$

$$2\Delta^2 [\tilde{\alpha}^2(z) + \epsilon^2(z) - V_x^2 - \omega^2] + 2\tilde{\alpha}^2(z) [V_x^2 - \omega^2 - \epsilon^2(z)] + \tilde{\alpha}^4(z) + V_x^4 + \Delta^4 + [\omega^2 - \epsilon^2(z)]^2 - 2V_x^2 [\omega^2 + \epsilon^2(z)] = 0,$$

$$\tilde{\alpha}(z) = -i\alpha(z - z^{-1})/2, \quad \epsilon(z) = -t(z + z^{-1} - 2) - \mu. \quad (2.12)$$

$$\sum_{n=1}^4 C_n(\omega) \left(z^n + \frac{1}{z^n} \right) + C_0(\omega) = 0, \quad (2.13)$$

$$\sum_{m=0}^8 a_m(\omega) z^m = 0, \quad (2.14)$$

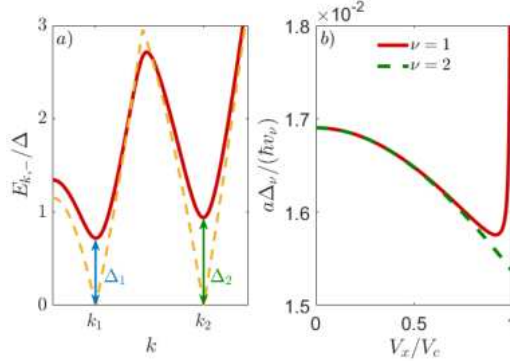


Figure 9: Spinful single-channel Majorana wire from Lutchyn and Oreg. (a) $E_{k,\pm}^2 = \Delta^2 + \alpha_k^2 + V_x^2 + \epsilon_k^2 \pm 2\sqrt{\Delta^2 V_x^2 + (\alpha_k^2 + V_x^2) \epsilon_k^2}$, topologically trivial regime $V_x < V_c$ (red), $k_{1,2} \simeq \cos^{-1} \left(\frac{2t(2t-\mu)}{\alpha^2 + 4t^2} \pm \frac{\sqrt{V_x^2(\alpha^2 + 4t^2) + \alpha^4 + 4t\mu\alpha^2 - \alpha^2\mu^2}}{\alpha^2 + 4t^2} \right)$. Here $\mu = 5$ meV, $\Delta = 2$ meV, and $V_x = 0.5V_c$. The dashed yellow curve is for $\Delta = 0$.

(b) Evolution of the two gaps (normalized to the velocities $v_{1,2}$) vs Zeeman parameter V_x for $\Delta = 0.2$ meV and $\mu = 2$ meV. We note that Δ_1 and v_1 simultaneously vanish as $V_x \rightarrow V_c$.

$$k_{1,2} \simeq \cos^{-1} \left(\frac{2t(2t-\mu)}{\alpha^2 + 4t^2} \pm \frac{\sqrt{V_x^2(\alpha^2 + 4t^2) + \alpha^4 + 4t\mu\alpha^2 - \alpha^2\mu^2}}{\alpha^2 + 4t^2} \right). \quad (2.15)$$

$$\mathcal{H}_{\text{eff}, \nu=1,2}(k) \simeq \begin{pmatrix} v_\nu(k - k_\nu) & \Delta_\nu \\ \Delta_\nu & -v_\nu(k - k_\nu) \end{pmatrix}, \quad (2.16)$$

$$\det[\omega - \mathcal{H}_{\text{eff},\nu}(z)] = 0$$

$$z_\nu(\omega) \simeq \left(1 \pm \frac{a}{v_\nu} \sqrt{\Delta_\nu^2 - \omega^2}\right) e^{ik_\nu a}, \quad (2.17)$$

$$z_\nu(\omega) = \left(1 - \tau_\nu \sqrt{\Delta_\nu^2 - \omega^2}\right) e^{i\delta_\nu}, \quad (2.18)$$

$$\tau_1 \Delta_1 = \tau_2 \Delta_2 = \eta \ll 1, \quad (2.19)$$

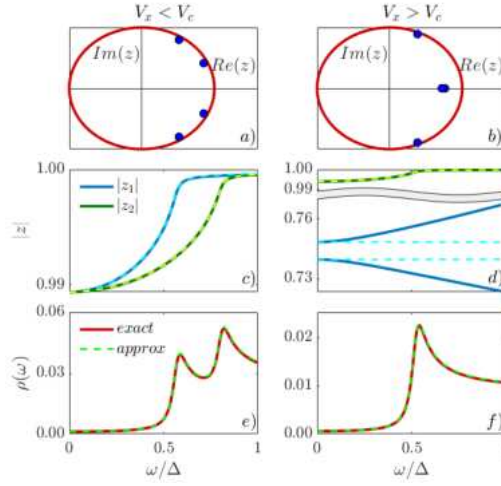


Figure 10: Behavior of the roots $z_n(\omega)$ for the spinful single-channel Majorana wire model [68, 69]. We use $V_x = 0.5V_c$ and $V_x = 1.2V_c$ as representatives for topologically trivial and nontrivial cases, respectively, with $\mu = 5$ meV and other parameters as specified in the main text. Upper panels: Roots $z_n(\omega = 0)$ (black dots) inside the unit circle (red) for (a) $V_x < V_c$ and (b) $V_x > V_c$. For illustrative purposes, we use $\Delta = 1$ meV in panels (a) and (b). For additional information, see Supplemental Material [73]. Middle panels: Modulus of the roots inside the unit circle vs ω/Δ for (c) $V_x < V_c$ and (d) $V_x > V_c$. Solid curves represent numerically exact results and dashed curves follow from Eqs. (15.19) and (15.21), respectively. Bottom panels: Energy dependence of the local bulk DoS, $\rho(\omega)$ (in meV^{-1}), for (e) $V_x < V_c$ and (f) $V_x > V_c$. The solid red curves depict numerically exact results using Eq. (15.3) and the dashed green curves show approximate results obtained from Eq. (15.25).

$$\begin{aligned} z_{1,\pm}(\omega) &= \left(1 \pm \tau_1 \sqrt{\Delta_1^2 - \omega^2}\right) e^{-\delta_1}, \\ z_{2,\pm}(\omega) &= \left(1 - \tau_2 \sqrt{\Delta_2^2 - \omega^2}\right) e^{\pm i\delta_2}, \end{aligned} \quad (2.20)$$

$$Q = \frac{\text{sgn Pf } \hat{\mathcal{H}}(k=0)}{\text{sgn Pf } \hat{\mathcal{H}}(k=\pi/a)} = \pm 1. \quad (2.21)$$

$$\rho(\omega) = -\frac{1}{\pi} \text{Im Tr} [\hat{G}_{00}(\omega)]. \quad (2.22)$$

$$\hat{G}_{00}(\omega) \simeq \sum_{|z_n| < 1} \frac{\hat{A}(z_n, \omega) + \omega \hat{A}'(z_n, \omega)}{\prod_{m \neq n} (z_n - z_m)}, \quad (2.23)$$

$$\hat{A}'(z_n, \omega) = \frac{d}{d\omega} \hat{A}(z_n(\omega), \omega).$$

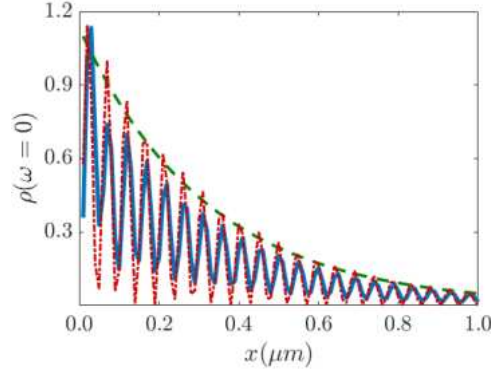


Figure 11: Spatial variation of the local DoS, $\rho_j(0)$ (in meV^{-1}), vs distance from the boundary, $x = ja$ (in μm), for the $\omega = 0$ Majorana state in a semi-infinite TS wire with $\mu = 1$ meV and $V_x = 2V_c$. The solid blue curve gives numerically exact results obtained from Eq. (15.7). Red-dotted and green-dashed curves show Eq. (15.26) with and without $2k_F$ oscillations, respectively.

$$\hat{G}_{00}(\omega) \approx \sum_{\nu=1,2} \frac{\hat{A}_\nu + \omega \hat{A}'_\nu}{\sqrt{\Delta_\nu^2 - \omega^2}}, \quad (2.24)$$

$$\rho_j(\omega = 0) \propto |z_2(0)|^{2j} \cos^2(jk_F a + \chi_0), \quad (2.25)$$

Two-channel class-D nanowire

$$\hat{\mathcal{H}}_{2\text{ch}}(k) = \begin{pmatrix} \hat{\mathcal{H}}(k) & \hat{T} \\ \hat{T}^\dagger & \hat{\mathcal{H}}(k) \end{pmatrix}, \quad (2.26)$$

$$\hat{T} = -t_y \sigma_0 \tau_z + i\alpha_y \sigma_x \tau_z + \Delta_y \sigma_0 \tau_x, \quad (2.27)$$

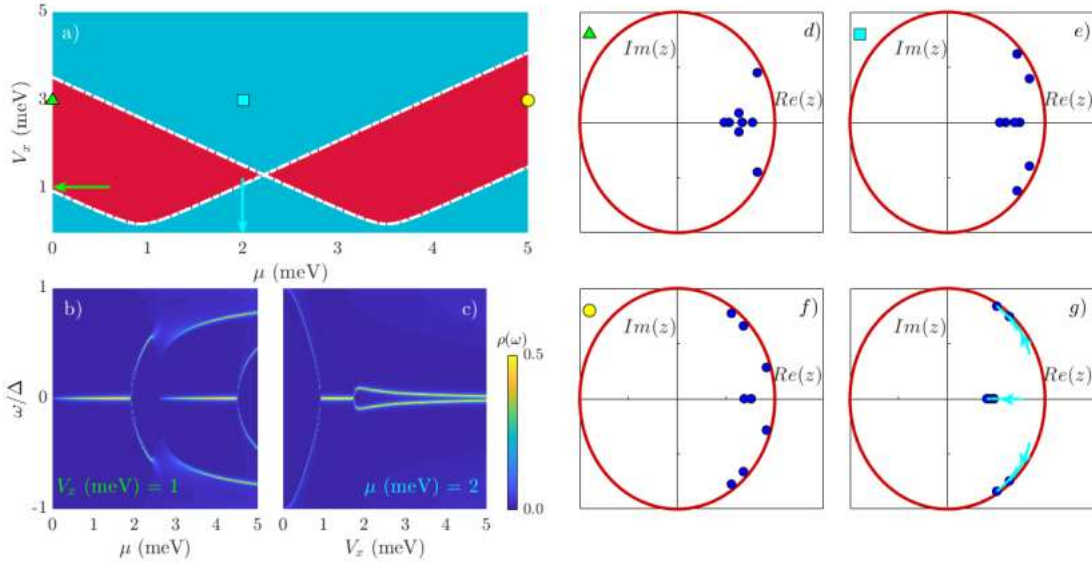


Figure 12: Two-channel spinful Majorana wire model of class D [see Eq. (15.27)] with parameters as explained in the main text. Panel (a) shows the bulk phase diagram in the μ - V_x plane. Topological nontrivial (trivial) phases are shown in red (blue). Panels (b) and (c) show the energy dependence of the local DoS, $\rho_{j=1}(\omega)$ (in meV^{-1}), at the boundary of a semi-infinite two-channel wire along the trajectories marked by arrows in panel (a). Panels (d)–(f) illustrate the roots $z_n(0)$ inside the unit circle at the three points indicated in panel (a) by a triangle (d), a square (e), and a circle (f), respectively. For additional insights, see Supplemental Material [73]. Panel (g) shows the evolution of the roots within the topologically trivial regime as V_x increases from 3 to 8 meV at constant chemical potential $\mu = 2$ meV. In panels (d)–(g), we use $\Delta = 1$ meV.

$$\text{Pf} \hat{\mathcal{H}}_{2\text{ch}}(0) = \alpha_y^4 + [(\mu - 3t_y)^2 - V_x^2 + \Delta^2] [(\mu - t_y)^2 - V_x^2 + \Delta^2] + 2\alpha_y^2 [-(\mu - 3t_y)(\mu - t_y) - V_x^2 + \Delta^2].$$

$$V_{c,\pm} = \left(\alpha_y^2 + \mu^2 - 4\mu t_y + 5t_y^2 + \Delta^2 \pm 2|\mu - 2t_y|\sqrt{t_y^2 + \alpha_y^2} \right)^{1/2}.$$

TRITOPS nanowires by Alvarado et. al.

Single-channel case

$$\hat{\mathcal{H}}_{\text{DIII}}(k) = \epsilon_k \sigma_0 \tau_z + \alpha_k \sigma_z \tau_z + \Delta_k \sigma_0 \tau_x, \quad (2.28)$$

$$\begin{aligned} \epsilon_k &= -2t \cos(ka) - \mu, \quad \alpha_k = 2\alpha \sin(ka), \\ \Delta_k &= 2\Delta \cos(ka). \end{aligned} \quad (2.29)$$

$$\hat{\Psi}_k^T \rightarrow (c_{k\uparrow}, c_{-k\downarrow}^\dagger, c_{k\downarrow}, -c_{-k\uparrow}^\dagger),$$

$$\hat{\mathcal{H}}_{\pm}(k) = (\epsilon_k \mp \alpha_k) \tilde{\sigma}_z + \Delta_k \tilde{\sigma}_x = \beta_{\pm}(k) \cdot \tilde{\sigma}, \quad (2.30)$$

$$\beta_{\pm}(k) = \begin{pmatrix} 2\Delta \cos(ka) \\ 0 \\ -\mu - 2t \cos(ka) \pm 2\alpha \sin(ka) \end{pmatrix} \quad (2.31)$$

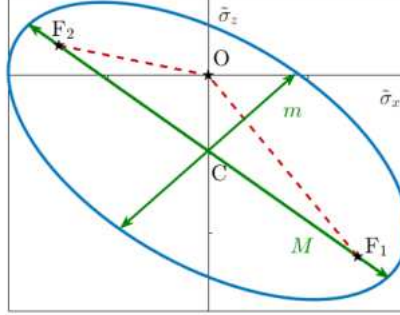


Figure 13: Curve traced out by $\beta_{-}(k)$ in the $\tilde{\sigma}_x$ - $\tilde{\sigma}_z$ plane for a single-channel TRITOPS wire in a topologically nontrivial phase [see Eqs. (15.33) and (15.34)] with $t = 0.5, \alpha = 0.8, \Delta = 1$, and $\mu = 1.04$ (all in meV). The evolution of the bulk Hamiltonian $\hat{\mathcal{H}}_{-}(k)$ upon traversal of the Brillouin zone is described by an ellipse containing the origin (O). For details, see main text and Appendix 15.9.

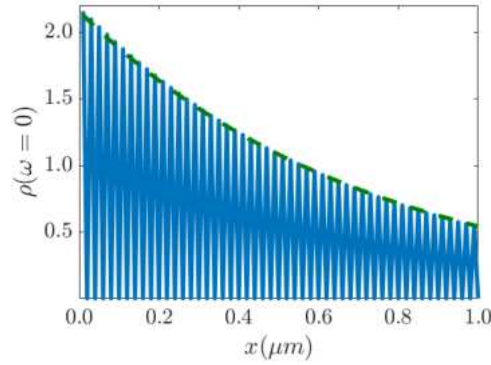


Figure 14: Spatial variation of the local DoS at zero energy (in meV^{-1}), corresponding to Majorana end states of a semi-infinite TRITOPS wire in its topological phase [see Eq. (15.30)] for $\mu = 0$ (blue solid curve). The green dashed curve shows an exponential decay on the length scale $\lambda_e = -\frac{a}{2} \ln |z_{\max}|$ [see Eq. (15.83)].

Two-channel case

$$\hat{\mathcal{H}}_{\text{DIII}, 2\text{ch}}(k) = \begin{pmatrix} \hat{\mathcal{H}}_{\text{DIII}}(k) & \hat{T}_{\text{DIII}} \\ \hat{T}_{\text{DIII}}^{\dagger} & \hat{\mathcal{H}}_{\text{DIII}}(k) \end{pmatrix}, \quad (2.32)$$

$$\hat{T}_{\text{DIII}} = -t_y \sigma_0 \tau_z + i\alpha_y \sigma_y \tau_z + \Delta_y \sigma_0 \tau_z. \quad (2.33)$$

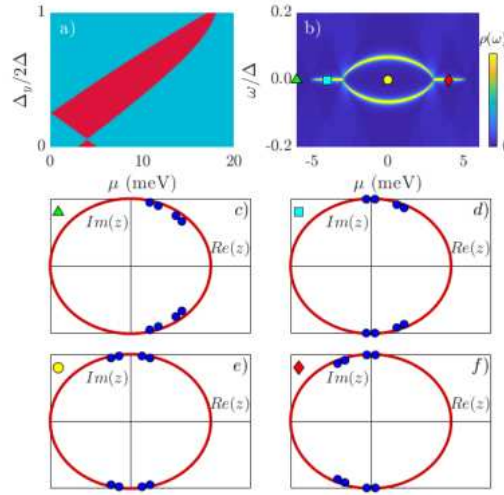


Figure 15: Two-channel TRITOPS nanowire [see Eq. (15.35)], with parameters as explained in the main text. Panel (a) shows the phase diagram in the μ - Δ_y plane, with the topologically nontrivial (trivial) phase in red (blue). (b) Local DoS, $\rho_{j=1}(\omega)$ (in meV^{-1}), at the boundary of a semi-infinite wire in the μ - ω plane for $\Delta_y = 0$. Panels (c) to (f) depict the roots $z_n(\omega = 0)$ inside the unit circle for different μ as indicated by the respective symbol in panel (b). We use $\Delta = 1 \text{ meV}$ in panels (c)–(f).

$$|\mu_{\pm}| = \sqrt{\alpha_y^2 + (t_y \pm 2\alpha)^2}. \quad (2.34)$$

Phase-biased topological Josephson junctions by Alvarado et. al.

Three-terminal TS junctions

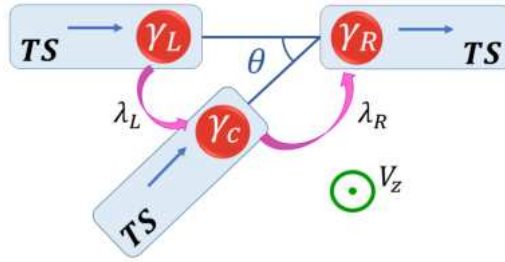


Figure 16: Three-terminal junction of spinful TS nanowires (see Sec. 15.4), with two parallel wires (L, R) and a central (C) wire at angle θ . The red dots indicate MBSs with Majorana operators $\gamma_{L,R,C}$ near the junction, with tunnel couplings $\lambda_{L,R}$ connecting the L, R wires to the C wire. We assume that no direct tunnel coupling between the L and R wires is present. A Zeeman field V_z is applied perpendicular to the plane containing the three wires. Blue arrows show the positive momentum direction in each wire.

$$R(\vartheta) = [\sigma_0 \cos(\vartheta/2) - i\sigma_y \sin(\vartheta/2)] \tau_0, \quad (2.35)$$

$$\hat{\mathcal{G}}'_{L/R} = R_y \hat{\mathcal{G}}_{L,R} R_y^{-1}, \quad (2.36)$$

$$\hat{\mathcal{G}}_C = R_z(\theta) R_y \hat{\mathcal{G}}_{L,R} R_y^{-1} R_z^{-1}(\theta). \quad (2.37)$$

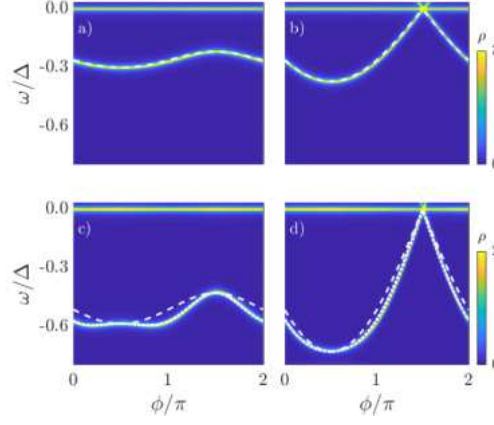


Figure 17: , $\phi_L = -\phi_R = \phi$ and $\phi_C = 0$. The TS wires are modeled as spinful nanowires with $\mu = 2$ meV, $V_z = 3V_c$, and symmetric couplings, $\lambda_L = \lambda_R = \lambda$. For other parameters, see Sec. 15.4. Panel (a) [(b)] is for $\lambda = 2$ meV and $\theta = \pi/2$ [$\theta = \pi/10$]. Panel (c) [(d)] is for $\lambda = 5$ meV and $\theta = \pi/2$ [$\theta = \pi/10$]. From blue to yellow, $\rho_{3\text{TS}}(\omega)$ (in meV^{-1}) gradually increases, where Eq. (15.43) has been evaluated in a numerically exact manner. White dotted [dashed] curves show the approximate Andreev bound state dispersion relation in Eq. (15.54) [Eq. (15.51)].

$$H_T = \frac{1}{2} \sum_{\nu=L,R} \hat{\Psi}_\nu^\dagger \hat{\lambda}_\nu \hat{\Psi}_C + \text{H.c.}, \quad \hat{\lambda}_\nu = \lambda_\nu \sigma_0 \tau_z e^{i\tau_z \phi_\nu/2}, \quad (2.38)$$

$$\hat{G}_{3\text{TS}} = \begin{pmatrix} \hat{\mathcal{G}}_L^{-1} & \hat{\lambda}_L & 0 \\ \hat{\lambda}_L^\dagger & \hat{\mathcal{G}}_C^{-1} & \hat{\lambda}_R \\ 0 & \hat{\lambda}_R^\dagger & \hat{\mathcal{G}}_R^{-1} \end{pmatrix}^{-1}, \quad (2.39)$$

$$\rho_{3\text{TS}}(\omega) = -\frac{1}{\pi} \text{Im Tr} [\hat{G}_{3\text{TS}}(\omega)]. \quad (2.40)$$

$$H_{\text{eff},\nu} = \lim_{\omega \rightarrow 0} \hat{\mathcal{G}}_\nu^{-1}(\omega). \quad (2.41)$$

$$\begin{aligned} \hat{\Psi}_L &\simeq \sqrt{\frac{\Delta_2}{t}} \begin{pmatrix} 0 \\ 1 \\ -i \\ 0 \end{pmatrix} \gamma_L, \quad \hat{\Psi}_R \simeq \sqrt{\frac{\Delta_2}{t}} \begin{pmatrix} 0 \\ -i \\ 1 \\ 0 \end{pmatrix} \gamma_R, \\ \hat{\Psi}_C &\simeq \sqrt{\frac{\Delta_2}{t}} R_z(\theta) \begin{pmatrix} 0 \\ 1 \\ -i \\ 0 \end{pmatrix} \gamma_C, \end{aligned} \quad (2.42)$$

$$H_{\text{mm}} = -i\Omega_L(\phi) \gamma_L \gamma_C - i\Omega_R(\phi) \gamma_R \gamma_C, \quad (2.43)$$

$$\begin{aligned} \Omega_L(\phi) &= \frac{2\Delta_2 \lambda_L}{t} \sin\left(\frac{\phi + \theta}{2}\right), \\ \Omega_R(\phi) &= -\frac{2\Delta_2 \lambda_R}{t} \cos\left(\frac{\phi - \theta}{2}\right). \end{aligned} \quad (2.44)$$

$$\begin{pmatrix} \gamma_L \\ \gamma_R \end{pmatrix} = \begin{pmatrix} \sin \kappa & -\cos \kappa \\ \cos \kappa & \sin \kappa \end{pmatrix} \begin{pmatrix} \tilde{\gamma}_L \\ \tilde{\gamma}_R \end{pmatrix}, \quad (2.45)$$

$$\sin \kappa = \Omega_L / \Omega$$

$$\Omega(\phi) = \sqrt{\Omega_L^2(\phi) + \Omega_R^2(\phi)}. \quad (2.46)$$

$$H_{\text{mm}} = -i\Omega(\phi)\tilde{\gamma}_L\gamma_C, \quad (2.47)$$

$$E_{\pm}(\phi) = \pm \frac{1}{2} \sqrt{\Omega_L^2(\phi) + \Omega_R^2(\phi)}. \quad (2.48)$$

$$\lambda_{LR} \simeq \lambda_L \lambda_R / \Delta.$$

$$H_{T,LR} = \frac{1}{2} \lambda_{LR} \hat{\Psi}_L^\dagger \sigma_0 e^{i\tau_z \phi} \tau_z \hat{\Psi}_R + \text{H.c.} \quad (2.49)$$

$$\Omega_{LR}(\phi) = \frac{2\Delta_2 \lambda_{LR}}{t} \cos \phi. \quad (2.50)$$

$$E_{\pm}(\phi) = \pm \frac{1}{2} \sqrt{\Omega_L^2(\phi) + \Omega_R^2(\phi) + \Omega_{LR}^2(\phi)}. \quad (2.51)$$

TRITOPS-TS junction

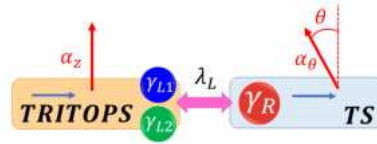


Figure 18: Blue arrows - the positive momentum; θ - relative angle of SO axes on both sides.

$$H_T = \frac{1}{2} \lambda_L \hat{\Psi}_L^\dagger \sigma_0 e^{i\tau_z \phi/2} \tau_z \hat{\Psi}_R + \text{H.c.}, \quad (2.52)$$

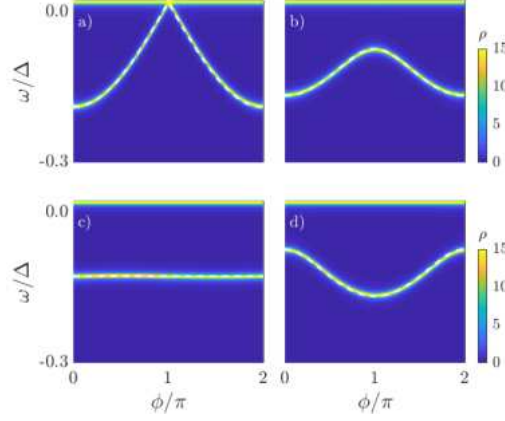


Figure 19: Phase-dependent subgap spectrum of a TRITOPS-TS Josephson junction for different values of the tilt angle θ in Fig. 79. The spinful single-channel model parameters are as described in Secs. 15.4 and 15.6, with $\mu = 1$ meV, $\lambda_L = 2$ meV, and $V_x = 1.5V_c$ on the TS side. The tilt angle is $\theta = 0$ in panel (a), $\theta = 0.3\pi$ in panel (b), $\theta = \pi/2$ in panel (c), and $\theta = 0.7\pi$ in panel (d). From blue to yellow, the color code indicates increasing DoS values at the junction, $\rho(\omega)$ (in meV^{-1}). White dashed curves show the Andreev bound states (15.59).

$$\begin{aligned}
 \hat{\Psi}_L &\simeq \sqrt{\frac{\Delta}{t}} \begin{pmatrix} 1 \\ 0 \\ i \\ 0 \end{pmatrix} \gamma_{L1} + \sqrt{\frac{\Delta}{t}} \begin{pmatrix} 0 \\ i \\ 0 \\ 1 \end{pmatrix} \gamma_{L2}, \\
 \hat{\Psi}_R &\simeq \sqrt{\frac{\Delta_2}{t}} R_y(\theta) \begin{pmatrix} i \\ -i \\ 1 \\ 1 \end{pmatrix} \gamma_R.
 \end{aligned} \tag{2.53}$$

$$H_{\min} = -i [w_1(\phi) \gamma_{L1} + w_2(\phi) \gamma_{L2}] \gamma_R \tag{2.54}$$

$$\begin{aligned}
 w_1(\phi) &= \frac{2\lambda_L \sqrt{\Delta \Delta_2}}{t} \cos \frac{\phi}{2} \cos \frac{\theta}{2}, \\
 w_2(\phi) &= -\frac{2\lambda_L \sqrt{\Delta \Delta_2}}{t} \sin \frac{\phi}{2} \sin \frac{\theta}{2}.
 \end{aligned} \tag{2.55}$$

$$E_{\pm}(\phi) = \pm \frac{1}{2} \sqrt{w_1^2(\phi) + w_2^2(\phi)}. \tag{2.56}$$

Part II

Very Useful and Famous Methods

3 PPS BGf Theory

3.1 Overview

We develop Green's function formalism to describe continuous multi-layered quasi-one-dimensional setups described by piece-wise constant single-particle Hamiltonians. The Hamiltonians of the individual layers are assumed to be quadratic polynomials in the momentum operator with matrix-valued (multichannel) coefficients. This, in particular, allows one to study transport in heterostructures consisting of multichannel conducting, superconducting, or insulating components with band structures of arbitrary complexity. We find a general expression for the single-particle Green's function of the combined setup in terms of the bulk (translationally invariant) Green's functions of its constituents. Furthermore, we provide the expression for the global density of states of the combined system and establish the bound state equation in terms of bulk Green's functions. We apply our formalism to investigate the spectrum and current-phase relations in ordinary and topological Josephson junctions, additionally showing how to account for the effects of static disorder and local Coulomb interaction.

3.1.1 Conclusions and Outlook

In this paper, we provided a detailed account of the interface Green's function technique regarding a large class of quasi-one-dimensional models of heterostructures. In our analysis, we assumed the Hamiltonians of the individual layers to be $N_c \times N_c$ matrix-valued quadratic polynomials in momentum operator conjugate to the composition axis of the structure. Such Hamiltonians may be seen as arising from the expansion of microscopic ones around the respective Fermi surfaces, allowing one to model a broad range of realistic physical systems (multi-band systems, systems with superconducting correlations, etc). For the considered class of models, we established the representation of the position space Green's function of the entire system in terms of bulk position space Green's functions of its sub-parts. By demonstrating that the calculation of latter objects is uninvolved, we opened up a simple pathway to study the multi-barrier scattering phenomena and the interface-localized bound states in systems of interest. Furthermore, our work lays one of the first stones into the analysis of many-body and disorder effects in layered systems, requiring knowledge of Green's functions as input. By further extracting the global DOS of the system from its Green's function, we reveal that the spectral information may, largely, be drawn from a single matrix d . For a system with M tunneling barriers, we show that d is an $N_c M \times N_c M$ block matrix, comprised of $N_c \times N_c$ -sized blocks admitting for a simple representation in terms of bulk Green's functions of the layers. As the spectral density relates to the logarithmic derivative of $\det d$, we also establish that the bound state energies of the system lie at its zeroes.

We shall point out that our formalism explicitly deals with ballistic contact problems (beyond the tunneling regime), commonly studied in modern-day experiments [52]. As nowadays the veil of theoretical secrecy over the ballistic regime is just being lifted, we hope our method finds its extensive use in this regard.

The method is further exemplified on models of normal and topological Josephson junctions. As a prototypical example, we considered a junction between two conventional superconductors with an induced phase difference across the interface. We derived the exact energy- and current-phase relations for the considered model, showing how they reduce to the commonly known

formulas in the Andreev approximation, for which the chemical potential in superconducting leads is assumed to be the largest energy scale in the model. Further, we studied whether the universal Josephson relations, arising in the Andreev limit, are robust against static disorder. By disorder-dressing the bulk Green's functions of superconductors in the T -matrix approximation and using those to assemble the Green's function of the composite system, we demonstrated the stability of the Josephson relations in the Andreev approximation, revealing only small corrections to the ABS energy beyond it. Next, we studied the effect of the local Coulomb interaction at the contact point between superconductors, showing how it leads to the $0 - \pi$ phase transition, detectable through the change in the current-phase relation of the system.

As for the topological Josephson systems, we considered models of short and long junctions between two Majorana wires. We demonstrated how to calculate energy-phase and current-phase relations for the considered models beyond the low-energy approximation, reducing the numerical computation to root-searching problems.

In addition, we indicate that our method may serve as a convenient starting point to develop low-energy approximations for considered models. When projecting the Hamiltonians of different subsystems onto the low-energy window of interest, one is typically challenged to identify and further implement the correct matching conditions on the distinct low-energy degrees of freedom. Our method allows us to elegantly circumvent this problem by developing the low-energy approximations of bulk single-layer Green's functions, as in Ref. [41]. In this paper, we demonstrated how this strategy is realized in the Majorana junction model (see Appendix 3.4.4), reproducing the low-energy results of Ref. [33].

3.1.2 BGf glossary

(напишу четко, какое обозначение что значит)

3.1.3 Introduction

Literature overview

Description of equilibrium and transport properties of layered quantum systems is a common problem in the domains of quantum electronics and solid state theory[1, 2]. Albeit typically a single-particle quantum mechanical problem, its solution is rarely simple due to the intricate band structures of the materials forming the layers. Most commonly, nowadays, these systems are analyzed numerically within the tight-binding approximation (like e.g. in Ref. [3]), allowing one to get a good grip on the low-lying excitations, as well as to assess the effects of the static disorder. The main idea behind this approach consists[4] in numerical studies of the ballistic conductance in mesoscopic structures within a lattice model, being expressed in terms of the lattice Green's function. It can be equally well applied to the study of equilibrium properties like Josephson current (JC) in Josephson junctions (JJ) of arbitrary width, see e.g. Refs. [5],[6].

An alternative to the tight-binding numerics is the scattering matrix approach [7, 8] applied to continuum ballistic models. One of its key early-day achievements was the calculation of the two-terminal conductance in terms of the transmission probabilities. The microscopic justification of this method relies either on taking the continuum limit of the wavefunction matching (WFM) of Ref. [4] or on studying the continuum limit of the atomistic Green's function (AGF) as in Ref. [9]. The equivalence of both WFM and AGF approaches has been fully substantiated in Ref. [10]. The relation between transmission and reflection coefficients of the scattering matrix on the one hand and Green's function, on the other hand, is widely known as the Fisher-Lee relation[11]. Its mode-resolved generalization has been recently proposed in Ref. [12].

In application to superconducting systems, the scattering matrix approach has been extended by Blonder, Tinkham, and Klapwijk[13] (BTK) on the basis of solutions of the Bogoliubov-de-Gennes equations. The BTK theory has been further generalized by C. Beenakker to the multichannel case in Ref. [14]. In this work the compact equation for the sub-gap Andreev bound states[15, 16] (ABS) and the expression for the continuous excitation spectrum of the JJ have been established in terms of normal and Andreev scattering matrices. In practice, however, these matrices are often treated in the so-called Andreev limit $\Delta_0 \ll \mu$, where Δ_0 is the absolute value of the superconducting order parameter and μ is the chemical potential, essentially neglecting the normal reflection at the superconducting interface.

An alternative theoretical description of the superconducting tunneling and proximity effects has been developed by G. Arnold in Refs. [17],[18] using standard nonequilibrium Green's functions. His approach is based on the theory of Feuchtwang[19, 20, 21], which does not make use of the tunneling Hamiltonian. This theory shares many common features with the AGF method of Caroli *et al* [9] mentioned above, and these techniques will serve as a starting point for our present consideration.

Discussing the approaches to interface physics, one cannot but mention that quasiclassical Green's functions treated in terms of the Eilenberger and the Usadel equations are also traditionally used for describing superconducting proximity effects and the JJs, see Refs. [22] and [23] for reviews.

Description of the article

In the first part of this section, we provide multichannel and multi-interface generalizations of results in Refs. [9], [17],[18], [19], [20], [21], essentially building upon the method of Ref. [9]. In particular, we derive a compact closed-form expression for the composite Green's function $G(x, x')$ of a quasi-one-dimensional heterostructure () given merely in terms of the bulk Green's functions $G^{(0,m)}(x, x')$ of its constituting layers (labeled by $m = 0, \dots, M$).

The knowledge of $G(x, x')$ is beneficial for several reasons. It gives direct access to the system's spectrum, allows for a calculation of various physical observables (like e.g. the density of states (DOS) and the current), may serve as a starting point for low-energy approximations (like e.g. the Andreev limit), as well as provides essential input for a perturbative diagrammatic treatment of disorder and many-body interaction effects. The bulk single-layer Green's functions $G^{(0,m)}(x, x')$ are obtained by means of the standard Fourier transformation and are computationally inexpensive. For few-channel models $G^{(0,m)}(x, x')$ may often be calculated analytically, as demonstrated in our examples below.

This idea (the extraction of the properties of inhomogeneous quantum systems from their bulk counterpart), is traditional in solid-state theory. This subject has a long history, and it often bears different names: the quantum theory of surface states[24, 25, 26, 27, 28] or the method of embedding[29, 30, 31], for example.

In the second part of this section, we apply the derived expression for $G(x, x')$ to the two paradigmatic examples: (1) Josephson model of a tunneling barrier between two *s*-wave superconductors, and (2) the model of a JJ between two semiconducting wires with strong spin-orbit interaction and proximity induced superconductivity [32, 33], submersed into the parallel magnetic field (each of which is capable of hosting a Majorana zero mode[34, 35]). In particular, we demonstrate that both the excitation spectrum (i.e. the global DOS) as well as the JC can be obtained in terms of a single matrix

$$d = \begin{pmatrix} G(0, 0) & G(0, W) \\ G(W, 0) & G(W, W) \end{pmatrix}^{-1}$$

for the finite-width junction (with the two interfaces at $x = 0$ and $x = W$), degenerating to $d = [G(0,0)]^{-1}$ in the short junction limit (single interface at $x = 0$). In particular, the correction to the global DOS due to the tunneling between the layers is given by

$$\delta\rho = -\frac{1}{\pi} \frac{\partial}{\partial\omega} \text{Im} \ln \det d,$$

which resembles analogous expressions for $\delta\rho$ in terms of the scattering matrix[36, 37]. A general equation for the bound states acquires a particularly simple form $\det d = 0$, which is equally well applicable to all types of heterostructures. The relation of our equations for the bound states and for $\delta\rho$ to their scattering matrix analogs becomes transparent on the basis of the Fisher-Lee relation[11] (????). The matrix d^{-1} appears as a common factor in expressions for all components of the scattering matrix, and it can be generally interpreted as a core part of the T -matrix in the coordinate representation. The advantage of computing d directly from the Green's function $G(x, x')$ consists, though, in its additive form, combining contributions from adjacent layers (and possibly from a local contact potential at their interface). This property of d is thus analogous to that of self-energies.

Using our formulas, we study the excitation spectrum and the JC in various parametric regimes of the two models discussed above, relaxing the Andreev approximation (which might be essential for one-dimensional wires) and allowing for arbitrary values of the junction's width W . The knowledge of the explicit form of $G(x, x')$ allows us to account for the effects of random potential disorder and local Coulomb interaction at the interface of the JJ. We find that the effect of static disorder arises only beyond the Andreev limit. Taking into account the local Coulomb interaction, we predict a crossover between the 0- and the π -junctions.

3.1.4 Algorithm of usage of interface Green's functions

(в 1й части это же, но короче)

3.1.5 Checking if all is fine at middle stages of computation

(пропишу, что должно получаться, чтобы сверять свойства.)

3.1.6 How to figure out this bGf method (???)

(потом отвечу, пока не знаю.)

Guessing about bGF formalism

How bGF formalism connects with Shrodinger's equation, Hewtonian laws or something other simple?

(a very stupid and a very cool question!!! I'll think about it later)

How we could solve same problems without bGf method (???)

(потом отвечу, пока не знаю, но вопрос важный, казалось бы, квантмех много решали, там же тоже такие были задачи??? что, из него методов на них не найти? заготовлю штук 4 задач элементарных для иллюстрации, как их решать без этого метода и в рамках него!)

3.2 Interface Green's functions for discrete models

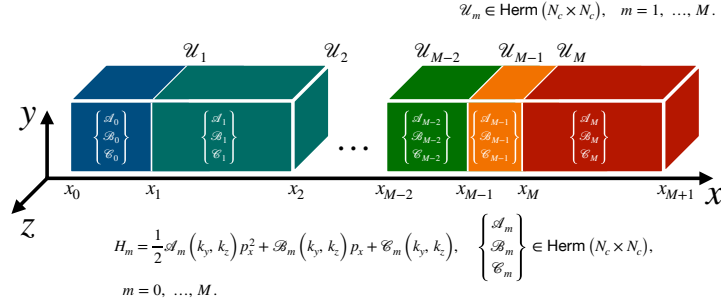


Figure 20: Sketch of the model: A heterostructure consists of $M + 1$ layers described by individual matrix-valued Hamiltonians H_m , which are consequently coupled with each other across matrix-valued tunneling barriers $\mathcal{U}_m$.

3.2.1 Model and problem formulation

Let us consider a quasi-one-dimensional system that consists of $M + 1$ layers (labeled by $m = 0, \dots, M$) subsequently connected with each other across tunneling barriers (also labeled by $m = 1, \dots, M$) along the spatial x -axis, see Fig. 31. We further assume that m -th layer extends over the spatial range (x_m, x_{m+1}) , with x_1 and x_M being the coordinates of the leftmost and the rightmost interfaces, respectively. The leftmost boundary of the whole system is x_0 and the rightmost one is x_{M+1} . An important class of models with semi-infinite leads is also included in the present consideration: They are realized by setting $x_0 = -\infty$ and $x_{M+1} = +\infty$. For the m -th subsystem we assume the Hamiltonian of the form

$$H_m = \frac{1}{2} \mathcal{A}_m p^2 + \mathcal{B}_m p + \mathcal{C}_m, \quad (3.1)$$

where $\mathcal{A}_m$, $\mathcal{B}_m$, $\mathcal{C}_m$ are constant $N_c \times N_c$ hermitian matrices, parametrically dependent on the conserved transverse (quasi)momenta k_y and k_z , with N_c being the number of channels (e.g. spin and/or orbital quantum numbers, etc.), and $p \equiv p_x = -i\partial_x$ is the momentum operator in the direction along the system.

The tunneling barriers between the subsystems are modeled by delta-like potentials at the corresponding interfaces:

$$\mathcal{U}(x) = \sum_{m=1}^M \mathcal{U}_m \delta(x - x_m), \quad (3.2)$$

where $\mathcal{U}_m$ are generally considered as $N_c \times N_c$ hermitian matrices (which, in particular, allows us to treat magnetically active contacts as well).

The Hamiltonian of the combined system is written as

$$H = \frac{1}{2} p \mathcal{A}(x) p + \frac{1}{2} \{ \mathcal{B}(x), p \} + \mathcal{C}(x) + \mathcal{U}(x), \quad (3.3)$$

where

$$\left\{ \begin{matrix} \mathcal{A}(x) \\ \mathcal{B}(x) \\ \mathcal{C}(x) \end{matrix} \right\} = \sum_{m=0}^M \Theta(x_m < x < x_{m+1}) \left\{ \begin{matrix} \mathcal{A}_m \\ \mathcal{B}_m \\ \mathcal{C}_m \end{matrix} \right\}. \quad (3.4)$$

The eigenvalue problem for the Hamiltonian H is complemented by matching conditions for the wave function and its derivative at each interface. While the wavefunction is continuous,

$$\Psi(x_m^-) = \Psi(x_m^+) \equiv \Psi(x_m), \quad x_m^\pm = x_m \pm 0^+, \quad (3.5)$$

its derivative satisfies the following matching condition at the m -th interface:

$$\frac{\mathcal{A}_m \Psi'(x_m^+) - \mathcal{A}_{m-1} \Psi'(x_m^-)}{2} = \left[\mathcal{U}_m - i \frac{\mathcal{B}_m - \mathcal{B}_{m-1}}{2} \right] \Psi(x_m). \quad (3.6)$$

This condition is traditionally derived by integrating the Schrödinger equation $H\Psi(x) = E\Psi(x)$ over the infinitesimally small region (x_m^-, x_m^+) enclosing the contact's coordinate x_m . Physically it expresses the conservation of the current density across the interface.

The main goal of our present consideration is to find an expression for the (retarded) Green's function of the whole system¹, which satisfies the equation

$$[z - H]G(x, x'; z) = \delta(x - x') \quad (3.7)$$

for arbitrary complex-valued spectral parameter z (with $\text{Im } z > 0$, in particular for $z = \omega + i0^+$, in the case of the retarded function). The differential equation $[z - H]G(x, x'; z) = \delta(x - x')$ is complemented by the vanishing boundary conditions $G(x_0, x'; z) = G(x_{M+1}, x'; z) = 0$ at the system's ends. By the virtue of the Lehmann representation, the Green's function $G(x, x'; z)$ must also obey the matching conditions $\Psi(x_m^-) = \Psi(x_m^+) \equiv \Psi(x_m)$ and $\frac{\mathcal{A}_m \Psi'(x_m^+) - \mathcal{A}_{m-1} \Psi'(x_m^-)}{2} = \left[\mathcal{U}_m - i \frac{\mathcal{B}_m - \mathcal{B}_{m-1}}{2} \right] \Psi(x_m)$ in the variable x . We note that $G(x, x'; z)$ can be alternatively defined as a solution of a reciprocal differential equation with respect to the variable x' equipped with the boundary conditions $G(x, x_0; z) = G(x, x_{M+1}; z) = 0$, while the corresponding interface matching conditions are obtained by hermitian conjugating $\Psi(x_m^-) = \Psi(x_m^+) \equiv \Psi(x_m)$ and $\frac{\mathcal{A}_m \Psi'(x_m^+) - \mathcal{A}_{m-1} \Psi'(x_m^-)}{2} = \left[\mathcal{U}_m - i \frac{\mathcal{B}_m - \mathcal{B}_{m-1}}{2} \right] \Psi(x_m)$.

As an input we use the *translationally invariant* (aka bulk) Green's functions $G^{(0,m)}(x, x'; z)$ satisfying the equation

$$[z - H_m]G^{(0,m)}(x, x'; z) = \delta(x - x') \quad (3.8)$$

on the whole spatial axis. They are evaluated by means of the Fourier transformation

$$G^{(0,m)}(x, x'; z) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \frac{e^{ik(x-x')}}{z - h_m(k)}, \quad (3.9)$$

where $h_m(k)$ is the Hamiltonian obtained from H_m in $H_m = \frac{1}{2}\mathcal{A}_m p^2 + \mathcal{B}_m p + \mathcal{C}_m$ by the substitution $p \rightarrow k$.

Often the integral in $G^{(0,m)}(x, x'; z)$ may be evaluated analytically by means of the integration in the complex k -plane: To this end, one needs to establish the complex-valued roots $k(z)$ of the secular equation $\det[z - h_m(k)] = 0$ with positive imaginary parts, $\text{Im } k(z) > 0$. Then the result of integration in $G^{(0,m)}(x, x'; z)$ is represented as a sum of residues at the corresponding poles $k(z)$ (see Appendix 3.2.1 for details). For few-channel models, the equation $\det[z - h_m(k)] = 0$ admits analytical solutions (some of them will be demonstrated in the following examples), while in general the roots $k(z)$ have to be determined numerically. As it will be demonstrated later, the root-searching routine is the only numerical part in solving the

¹Note that the heterostructural Green's function $G(x, x')$ is a highly non-trivial object, containing information about all possible multi-barrier scattering phenomena and interface-localized bound states.

problem of establishing $G(x, x'; z)$, and for this reason, the Green's function approach developed below should provide a considerable speed up in the study of arbitrary heterostructures.

An opening move in establishing $G(x, x'; z)$ is a determination of the set of Green's functions $G^m(x, x'; z)$ describing every isolated layer m on the corresponding spatial interval (x_m, x_{m+1}) . The Green's function $G^m(x, x'; z)$ also satisfies $[z - H_m]G^{(0,m)}(x, x'; z) = \delta(x - x')$, but – in contrast to $G^{(0,m)}(x, x'; z)$ – it vanishes at the interval's ends, that is $G^m(x_m, x'; z) = G^m(x_{m+1}, x'; z) = 0$.

Thankfully, the determination of $G^m(x, x')$ (the argument z is omitted for brevity) is conveniently solved by the so-called boundary Green's function technique, allowing one to write a simple relation between the propagators of infinite $[G^{(0,m)}(x, x')]$ and bounded $[G^m(x, x')]$ systems (see Appendix 3.3.1, and references therein, for the summary of the key results).

Next, we find out how the isolated layers are coupled with each other.

As such, this coupling is dictated by the matching conditions $\Psi(x_m^-) = \Psi(x_m^+) \equiv \Psi(x_m)$ and $\frac{\mathcal{A}_m \Psi'(x_m^+) - \mathcal{A}_{m-1} \Psi'(x_m^-)}{2} = \left[\mathcal{U}_m - i \frac{\mathcal{B}_m - \mathcal{B}_{m-1}}{2} \right] \Psi(x_m)$. The key question is then how does one implement these requirements in terms of a local potential? Once this is accomplished, a set of Dyson's equations may be set up, relating the full Green's function $G(x, x')$ to those $G^m(x, x')$ of the individual layers. Furthermore, it is expected that the set of these equations admits a closed-form analytical solution, as by the locality of the coupling potential the integral equations are expected to reduce to algebraic ones for a finite number of unknowns.

To give an answer to the key question we underscore the two approaches which we find especially useful in practice.

The first approach, inspired by the earlier ideas of the seminal paper Ref. [9], is based on discretizing the equation $[z - H]G(x, x'; z) = \delta(x - x')$ on a lattice with the spacing a , solving the obtained tight-binding counterpart, and taking carefully the continuum limit $a \rightarrow 0$. As in the tight-binding description the space consists of discrete points, the tunneling between the disjoint parts of the system may be defined unambiguously, and hence the above-described program with a formal solution of the corresponding Dyson equations can be successfully executed. Further evaluation of the limit $a \rightarrow 0$ allows us to recover the complete Green's function of the corresponding continuum theory, expressed in terms of boundary Green's functions of the individual layers as well as their spatial derivatives up to the second order. In this paper, we describe the major steps of implementing this approach for the models of our present interest.

The second approach is based on (the multichannel generalization of) the Sturm-Liouville theory for second-order differential operators. It constructively exploits the matching conditions $\Psi(x_m^-) = \Psi(x_m^+) \equiv \Psi(x_m)$ and $\frac{\mathcal{A}_m \Psi'(x_m^+) - \mathcal{A}_{m-1} \Psi'(x_m^-)}{2} = \left[\mathcal{U}_m - i \frac{\mathcal{B}_m - \mathcal{B}_{m-1}}{2} \right] \Psi(x_m)$ to recover $G(x, x')$ staying within the paradigm of continuum models. The results of its application naturally reproduce those described below in the present paper.

Green's functions of translationally invariant systems

Let us consider the infinite-space model described by the Hamiltonian $H_m = \frac{1}{2}\mathcal{A}_m p^2 + \mathcal{B}_m p + \mathcal{C}_m$ and evaluate its Green's function

$$G^{(0,m)}(x, x'; z) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \frac{e^{ik(x-x')}}{z - h_m(k)}, \quad (3.10)$$

in the position representation. It can be conveniently rewritten as

$$G^{(0,m)}(x, x'; z) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \frac{e^{ik(x-x')}}{Q(k; z)} P(k; z), \quad (3.11)$$

where $P(k; z) = \text{adj}[z - h_m(k)]$ is the adjugate matrix, and $Q(k; z) = \det[z - h_m(k)]$ is a $2N_c$ -order polynomial of k , admitting the representation

$$Q(k; z) = \det(\tfrac{1}{2}\mathcal{A}_m) \prod_{s=1}^{2N_c} (k - k_s(z)) \quad (3.12)$$

in terms of the roots $k_s(z)$.

By the hermiticity of $h_m(k)$ the determinant $Q(k; \omega)$ on the real frequency axis may only have either a pair of complex conjugated roots or purely real roots. Upon the shift $\omega \rightarrow \omega + i0^+$ the latter acquire infinitesimal imaginary parts, whose signs are given by the signs of $\frac{dk_s(\omega)}{d\omega}$, or equivalently by the signs of the dispersion slope (group velocity) $\left. \frac{d\omega_k}{dk} \right|_{k=k_s}$. For the Hamiltonian $H_m = \frac{1}{2}\mathcal{A}_m p^2 + \mathcal{B}_m p + \mathcal{C}_m$, a number of dispersion branches pointing up (down) at negative k is the same as a number of dispersion branches pointing up (down) at positive k . This implies that for every ω the number of real roots with the positive dispersion slope equals the number of real roots with the negative dispersion slope. Therefore we can generally classify all roots into the two equally sized sets $\{k_{s,+}(\omega + i0^+)\}_{s=1}^{N_c}$ and $\{k_{s,-}(\omega + i0^+)\}_{s=1}^{N_c}$ according to

$$\text{Im}[k_{s,+}(\omega + i0^+)] > 0, \quad \text{Im}[k_{s,-}(\omega + i0^+)] < 0. \quad (3.13)$$

Evaluating $G^{(0,m)}(x, x'; z)$ at $z = \omega + i0^+$ we close the integration contour in the upper and lower halves of the complex momentum plane for $x > x'$ and $x < x'$, respectively, to correspondingly pick up the residues at the isolated singularities $k_{s,+}$ or $k_{s,-}$. This results in

$$\begin{aligned} G^{(0,m)}(x, x'; \omega + i0^+) &= \sum_{s=1}^{N_c} \frac{i\Theta(x - x')e^{ik_{s,+}|x-x'|}}{\det(\tfrac{1}{2}\mathcal{A}_m) \prod_{s' \neq s} (k_{s,+} - k_{s',+}) \prod_{s'} (k_{s,+} - k_{s',-})} P(k_{s,+}; \omega) \\ &\quad - \sum_{s=1}^{N_c} \frac{i\Theta(x' - x)e^{-ik_{s,-}|x-x'|}}{\det(\tfrac{1}{2}\mathcal{A}_m) \prod_{s'} (k_{s,-} - k_{s',+}) \prod_{s' \neq s} (k_{s,-} - k_{s',-})} P(k_{s,-}; \omega). \end{aligned} \quad (3.14)$$

Thus the evaluation of

$$G^{(0,m)}(x, x'; z) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \frac{e^{ik(x-x')}}{z - h_m(k)}, \quad (3.15)$$

reduces to a problem of the polynomial factorization $Q(k; z) = \det(\tfrac{1}{2}\mathcal{A}_m) \prod_{s=1}^{2N_c} (k - k_s(z))$ with a root post-selection criteria

$$\text{Im}[k_{s,+}(\omega + i0^+)] > 0, \quad \text{Im}[k_{s,-}(\omega + i0^+)] < 0. \quad (3.16)$$

In many practical applications, these tasks may be executed analytically, otherwise one can resort to numerical methods, such as the NumPy's built-in routine `numpy.roots`, for example [53].

3.2.2 Single-barrier setup

In this subsection, we extrapolate the construction technique of the composite Green's function of a double-layer system of Ref. [9] to the multichannel case.

Let us start by considering the simplest case of a single barrier separating left L ($m = 0$, $x \in (x_0, 0)$, $x_0 < 0$) and right R ($m = 1$, $x \in (0, x_2)$, $x_2 > 0$) subsystems at $x = x_1 = 0$.

We introduce the following lattice counterpart of the single-barrier continuum model

$$H = H_L + H_R + |0\rangle W_0 \langle 0| - |0\rangle t_L \langle -1| - |-1\rangle t_L^\dagger \langle 0| - |1\rangle t_R \langle 0| - |0\rangle t_R^\dagger \langle 1|, \quad (3.17)$$

where the left and right disjoint subsystems are defined on the lattice sites $n_0 \leq n \leq -1$ and $1 \leq n \leq n_2$, respectively. They are described by the Hamiltonians

$$H_L = - \sum_{n=n_0+1}^{-1} \left(|n\rangle t_L \langle n-1| + |n-1\rangle t_L^\dagger \langle n| \right) + \sum_{n=n_0}^{-1} |n\rangle W_L \langle n|, \quad (3.18)$$

$$H_R = - \sum_{n=1}^{n_2-1} \left(|n+1\rangle t_R \langle n| + |n\rangle t_R^\dagger \langle n+1| \right) + \sum_{n=1}^{n_2} |n\rangle W_R \langle n|, \quad (3.19)$$

with constant, matrix-valued (in the channel space) hopping amplitudes t_L, t_R and onsite potentials W_L, W_R . The central site $n = 0$ is characterized by the onsite potential W_0 . It is coupled to both subsystems by the same nearest-neighbor hopping amplitudes t_L and t_R as occur in the Hamiltonians H_L and H_R , respectively.

Treating the coupling term $|0\rangle t_L \langle -1| - |-1\rangle t_L^\dagger \langle 0| - |1\rangle t_R \langle 0| - |0\rangle t_R^\dagger \langle 1|$ as a perturbation, we set up the following Dyson equation in the coordinate representation for the Green's function $G = \frac{1}{z-H}$ of the full system

$$G_{n,n'} = G_{n,n'}^L + G_{n,n'}^C + G_{n,n'}^R - \left(G_{n,-1}^L t_L^\dagger + G_{n,1}^R t_R \right) G_{0,n'} - \delta_{n,0} G_{0,0}^C \left(t_L G_{-1,n'} + t_R^\dagger G_{1,n'} \right) \quad (3.20)$$

in terms of the Green's functions

$$G^L = \frac{1}{z - H_L}, \quad G^C = |0\rangle \frac{1}{z - W_0} \langle 0|, \quad G^R = \frac{1}{z - H_R}$$

of the three disjoint subsystems. The Green's function $G_{n,n'}^L$ is nonzero only for $n, n' \leq -1$, while $G_{n,n'}^R$ is nonzero only for $n, n' \geq 1$. In the corresponding domains, they are expressed according to $G_{n,n'}^m = G_{n,n'}^{(0,m)} - G_{n,0}^{(0,m)} [G_{0,0}^{(0,m)}]^{-1} G_{0,n'}^{(0,m)}$ in terms of the Green's functions $G_{n,n'}^{(L,0)}$ and $G_{n,n'}^{(R,0)}$ of the two auxiliary models defined on the whole lattice and using the constant parameters from the left and the right subsystems, respectively.

Due to the locality of the perturbation, the Dyson equation admits the explicit solution

$$G_{n,n'} = G_{n,n'}^L + G_{n,n'}^R + F_n D^{-1} \bar{F}_{n'}, \quad (3.21)$$

$$D := z - W_0 - t_L G_{-1,-1}^L t_L^\dagger - t_R^\dagger G_{1,1}^R t_R \quad (3.22)$$

$$F_n := -\delta_{n,0} + G_{n,-1}^L t_L^\dagger + G_{n,1}^R t_R, \quad (3.23)$$

$$\bar{F}_{n'} := -\delta_{0,n'} + t_L G_{-1,n'}^L + t_R^\dagger G_{1,n'}^R. \quad (3.24)$$

We particularly note that

$$F_0 = \bar{F}_0 = -1, \quad (3.25)$$

$$G_{0,0} = D^{-1}. \quad (3.26)$$

On the basis of the solution $G_{n,n'} = G_{n,n'}^L + G_{n,n'}^R + F_n D^{-1} \bar{F}_{n'}$ it is straightforward to compute the global DOS. Setting $z = \omega + i0^+$, we evaluate

$$\sum_{n=n_0}^{n_2} \text{tr} [G_{n,n} - G_{n,n}^L - G_{n,n}^R] = \text{tr} \left[\frac{1}{D} \right] + \sum_{n=n_0}^{-1} \text{tr} \left[\frac{1}{D} t_L G_{-1,n}^L G_{n,-1}^L t_L^\dagger \right] + \sum_{n=1}^{n_2} \text{tr} \left[t_R^\dagger G_{1,n}^R G_{n,1}^R t_R \frac{1}{D} \right], \quad (3.27)$$

where the trace operation is performed in the channel space. Using the identities

$$\sum_{n=n_0}^{-1} G_{-1,n}^L G_{n,-1}^L = -\frac{\partial G_{-1,-1}^L}{\partial \omega}, \quad (3.28)$$

$$\sum_{n=1}^{n_2} G_{1,n}^R G_{n,1}^R = -\frac{\partial G_{1,1}^R}{\partial \omega}, \quad (3.29)$$

which are most easily proven in the Lehmann representation, and the Jacobi's formula $\text{tr} [D^{-1} \frac{\partial D}{\partial \omega}] = \frac{\partial}{\partial \omega} \ln \det D$, we establish that

$$\rho(\omega) = \rho^L(\omega) + \rho^R(\omega) - \frac{1}{\pi} \text{Im} \frac{\partial}{\partial \omega} \ln \det D(\omega + i0^+). \quad (3.30)$$

The last term in this expression represents a correction to the global DOS due to the tunneling between the subsystems. It has a form analogous to that of the familiar expression in terms of the scattering matrix[37]. The terms $\rho^L(\omega)$ and $\rho^R(\omega)$ represent the global DOS of the left and right disjoint subsystems, respectively.

In order to derive the continuum limit of $G_{n,n'} = G_{n,n'}^L + G_{n,n'}^R + F_n D^{-1} \bar{F}_{n'}$ we make certain assumptions about the scaling of the Hamiltonian parameters with the lattice constant a . In particular, we define for $m = L, R$

$$t_m + t_m^\dagger = \frac{\mathcal{A}_m}{a^2}, \quad (3.31)$$

$$i(t_m - t_m^\dagger) = \frac{\mathcal{B}_m}{a}, \quad (3.32)$$

$$W_m = \frac{\mathcal{A}_m}{a^2} + \mathcal{C}_m, \quad (3.33)$$

where $\mathcal{A}_m, \mathcal{B}_m, \mathcal{C}_m = O(a^0)$ are constant matrices.

In addition, we choose

$$W_0 = \frac{\mathcal{A}_L + \mathcal{A}_R}{2a^2} + \frac{\mathcal{U}_1}{a}. \quad (3.34)$$

This choice of the leading $O(\frac{1}{a^2})$ term in W_0 is important [40] (?? черт его знает, какое отношение эта статья имеет к формуле ниже или выше?) for ensuring the fulfillment of the matching conditions

$$\Psi(x_m^-) = \Psi(x_m^+) \equiv \Psi(x_m); \quad \frac{\mathcal{A}_m \Psi'(x_m^+) - \mathcal{A}_{m-1} \Psi'(x_m^-)}{2} = \left[\mathcal{U}_m - i \frac{\mathcal{B}_m - \mathcal{B}_{m-1}}{2} \right] \Psi(x_m)$$

for the Green's function in the continuum limit [see also in the end of the subsection]. In turn, the subleading $O(\frac{1}{a})$ term induces the impurity delta-potential $\mathcal{U}(x) = \sum_{m=1}^M \mathcal{U}_m \delta(x - x_m)$, which is generally matrix-valued.

It is important to emphasize that our limiting procedure essentially differs from the one frequently used in the wide-band limit treatment of the tunneling regime $t'_{L,R} \ll t_{L,R}$, where the hopping amplitudes $t_{L,R}$ in the leads considerably dominate over the hopping amplitudes $t'_{L,R}$ from the leads onto the quantum dot at the site $n = 0$, and the characteristic tunneling rates $\Gamma_{L,R} = \pi(t'_{L,R})^2/t_{L,R}$ giving rise to (the imaginary part of) the dot's self-energy are then much smaller than the bandwidths of the leads $\sim t_{L,R}$. Recall that in our treatment we set $t'_{L,R} = t_{L,R}$.

Defining the continuous variable $x = na$ in the limit $a \rightarrow 0$, we recover the continuum analog $H_m = \frac{1}{2}\mathcal{A}_m p^2 + \mathcal{B}_m p + \mathcal{C}_m$ of the lattice Hamiltonians

$$H_L = - \sum_{n=n_0+1}^{-1} \left(|n\rangle t_L \langle n-1| + |n-1\rangle t_L^\dagger \langle n| \right) + \sum_{n=n_0}^{-1} |n\rangle W_L \langle n|, \quad (3.35)$$

$$H_R = - \sum_{n=1}^{n_2-1} \left(|n+1\rangle t_R \langle n| + |n\rangle t_R^\dagger \langle n+1| \right) + \sum_{n=1}^{n_2} |n\rangle W_R \langle n|, \quad (3.36)$$

Relating the continuum and the lattice Green's functions via

$$G(x, x') = \lim_{a \rightarrow 0} \frac{1}{a} G_{n, n'}, \quad (3.37)$$

we derive in Appendix 3.3.1 the continuum analog of $G_{n, n'} = G_{n, n'}^L + G_{n, n'}^R + F_n D^{-1} \bar{F}_{n'}$. It reads

$$G(x, x') = G^L(x, x') + G^R(x, x') + F(x) d^{-1} \bar{F}(x'), \quad (3.38)$$

where $G^L(x, x')$ and $G^R(x, x')$ are the boundary Green's functions of the corresponding disjoint regions $(x_0, 0)$ and $(0, x_2)$ which identically vanish outside of them, respectively. The last term $F(x) d^{-1} \bar{F}(x')$ contains non-diagonal contributions in the subsystem's basis (i.e. it is generically non-zero for all $x, x' \in (x_0, x_2)$), and thereby it mediates the coupling between the subsystems. It is defined through the following objects

$$F(x) = \lim_{a \rightarrow 0} F_n = -\Theta(-x) \left[G_2^{(0,L)}(x, 0^+) + G_2^L(x, 0^-) - G_2^{(0,L)}(x, 0^-) \right] \frac{\mathcal{A}_L}{2} \\ + \Theta(x) \left[G_2^{(0,R)}(x, 0^-) + G_2^R(x, 0^+) - G_2^{(0,R)}(x, 0^+) \right] \frac{\mathcal{A}_R}{2}, \quad (3.39)$$

$$\bar{F}(x') = \lim_{a \rightarrow 0} \bar{F}_{n'} = -\Theta(-x') \frac{\mathcal{A}_L}{2} \left[G_1^{(0,L)}(0^+, x') + G_1^L(0^-, x') - G_1^{(0,L)}(0^-, x') \right] \\ + \Theta(x') \frac{\mathcal{A}_R}{2} \left[G_1^{(0,R)}(0^-, x') + G_1^R(0^+, x') - G_1^{(0,R)}(0^+, x') \right], \quad (3.40)$$

$$d = \lim_{a \rightarrow 0} a D = -\mathcal{U}_1 - \frac{1}{8} \mathcal{A}_L \lim_{x \rightarrow 0^-} \frac{d^2}{dx^2} G^L(x, x) \mathcal{A}_L - \frac{1}{8} \mathcal{A}_R \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G^R(x, x) \mathcal{A}_R, \quad (3.41)$$

where we employed a set of compact notations for the coordinate derivatives of Green's functions

$$G_1^*(x, x') := \frac{\partial G^*(x, x')}{\partial x}, \quad G_2^m(x, x') := \frac{\partial G^*(x, x')}{\partial x'}.$$

The obtained expressions represent a multichannel generalization of Eqs.

$$G(x_{0-}, x_{0+}) \lim_{\epsilon \rightarrow 0} G_{ii}^{(\epsilon)} = + \frac{2m}{\hbar^2} \frac{1}{\gamma(x_0) + \bar{\gamma}(x_0)},$$

$$\gamma(x_0) = - \frac{\hbar^2}{2m} \lim_{x \rightarrow x_0} \sum_{\alpha} \frac{\Psi_{\alpha}^*(x_0) \Psi'_{\alpha}(x)}{\omega - E_{\alpha}} = \lim_{x \rightarrow x_0} \lim_{x' \rightarrow x_0} \left(- \frac{\hbar^2}{2m} \right) \frac{\partial}{\partial x} \frac{\partial}{\partial x'} g(x, x')$$

in Ref. [9]. The similar expressions for the two-channel case were previously derived in Refs. [17], [18].

The expression $G(x, x') = G^L(x, x') + G^R(x, x') + F(x) d^{-1} \bar{F}(x')$ can be interpreted in terms of the T -matrix as follows. Formally rewriting it as

$$G(x, x') = G^{L+R}(x, x') + \langle x | G^{L+R} l^\dagger d^{-1} l G^{L+R} | x' \rangle \quad (3.42)$$

$$= G^{L+R}(x, x') + \int dy \int dy' G^{L+R}(x, y) T(y, y') G^{L+R}(y', x'), \quad (3.43)$$

where $G^{L+R} = G^L + G^R$ and l is the formally introduced boundary hermitian operator in terms of

$$\langle x | G^{L+R} l^\dagger | y \rangle = i\delta(y)F(x), \quad (3.44)$$

$$\langle y' | l G^{L+R} | x' \rangle = -i\delta(y')\bar{F}(x'), \quad (3.45)$$

we identify $T(y, y') = \langle y | l^\dagger d^{-1} l | y' \rangle$ with the T -operator in the coordinate representation.

The functions

$$F(x) = \lim_{a \rightarrow 0} F_n = -\Theta(-x) \left[G_2^{(0,L)}(x, 0^+) + G_2^L(x, 0^-) - G_2^{(0,L)}(x, 0^-) \right] \frac{\mathcal{A}_L}{2} \\ + \Theta(x) \left[G_2^{(0,R)}(x, 0^-) + G_2^R(x, 0^+) - G_2^{(0,R)}(x, 0^+) \right] \frac{\mathcal{A}_R}{2}, \quad (3.46)$$

$$\bar{F}(x') = \lim_{a \rightarrow 0} \bar{F}_{n'} = -\Theta(-x') \frac{\mathcal{A}_L}{2} \left[G_1^{(0,L)}(0^+, x') + G_1^L(0^-, x') - G_1^{(0,L)}(0^-, x') \right] \\ + \Theta(x') \frac{\mathcal{A}_R}{2} \left[G_1^{(0,R)}(0^-, x') + G_1^R(0^+, x') - G_1^{(0,R)}(0^+, x') \right] \quad (3.47)$$

have the remarkable properties

$$F(0^+) = F(0^-) \equiv F(0) = -1, \quad (3.48)$$

$$\bar{F}(0^+) = \bar{F}(0^-) \equiv \bar{F}(0) = -1, \quad (3.49)$$

which follow from the standard jump conditions

$$[G_2^{(0,m)}(0, 0^+) - G_2^{(0,m)}(0, 0^-)] \frac{\mathcal{A}_m}{2} = 1, \quad \frac{\mathcal{A}_m}{2} [G_1^{(0,m)}(0^+, 0) - G_1^{(0,m)}(0^-, 0)] = 1, \quad (3.50)$$

$$\left(G_1^*(x, x') := \frac{\partial G^*(x, x')}{\partial x}, \quad G_2^*(x, x') := \frac{\partial G^*(x, x')}{\partial x'} \right).$$

The expressions $F(0^+) = F(0^-) \equiv F(0) = -1$ and $F(0^+) = F(0^-) \equiv F(0) = -1$ thus appear to be consistent with their lattice analogues $F_n := -\delta_{n,0} + G_{n,-1}^L t_L^\dagger + G_{n,1}^R t_R$ and $\bar{F}_{n'} := -\delta_{0,n'} + t_L G_{-1,n'}^L + t_R^\dagger G_{1,n'}^R$, respectively.

By the virtue of $G^L(0, 0) = G^R(0, 0)$ and $F(0^+) = F(0^-) \equiv F(0) = -1$, $\bar{F}(0^+) = \bar{F}(0^-) \equiv \bar{F}(0) = -1$ we recover

$$G(0, 0) = d^{-1}. \quad (3.51)$$

This result is also consistent with its lattice analog $G_{0,0} = D^{-1}$. Replacing $D \rightarrow d/a$ in the lattice version of the global DOS $\rho(\omega) = \rho^L(\omega) + \rho^R(\omega) - \frac{1}{\pi} \text{Im} \frac{\partial}{\partial \omega} \ln \det D(\omega + i0^+)$, we immediately obtain its continuum counterpart

$$\rho(\omega) = \rho^L(\omega) + \rho^R(\omega) - \frac{1}{\pi} \text{Im} \frac{\partial}{\partial \omega} \ln \det d(\omega + i0^+). \quad (3.52)$$

The global DOS $\rho^m(\omega)$ of the disjoint subsystem with L length is generally given by

$$\rho^{L/R}(\omega) = -\frac{1}{\pi} \text{Im} \int_L dx \text{tr} G^{L/R}(x, x; \omega + i0^+). \quad (3.53)$$

Since $d(\omega)$ is hermitian on the real frequency axis, its determinant $\det d(\omega)$ is real-valued. The real-valued roots of the equation

$$\det d(\omega) = 0 \quad (3.54)$$

determine energies E_B of bound states induced by the tunneling between the two subsystems. Their contribution to the global DOS $\rho(\omega) = \rho^L(\omega) + \rho^R(\omega) - \frac{1}{\pi} \text{Im} \frac{\partial}{\partial \omega} \ln \det d(\omega + i0^+)$ naturally appears in the form $\sum_B \delta(\omega - E_B)$.

In the infinite-space model (that is, when both the left and the right subsystems are semi-infinite), the bound states found from the equation $\det d(\omega) = 0$ are localized near the interface of the two subsystems, and their energies E_B reside in the bandgaps of the whole system. It is also remarkable that in this case the matrix d can be alternatively written as

$$d = [G^{(0,L)}(0, 0)]^{-1} - p_0, \quad (3.55)$$

and hence the bound state equation $\det d(\omega) = 0$ acquires the form

$$\det [1 - G^{(0,L)}(0, 0; \omega) p_0(\omega)] = 0. \quad (3.56)$$

Here

$$p_0 = \mathcal{U}_1 + \frac{1}{8} \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} [\mathcal{A}_R G^R(x, x) \mathcal{A}_R - \mathcal{A}_L G_0^L(x, x) \mathcal{A}_L]$$

is the (generally energy-dependent) term breaking the translational invariance of the auxiliary infinite-space model using the parameters of the left subsystem, i.e. described by $G^{(0,L)}(x, x')$. In addition, we employed the auxiliary Green's function $G_0^L(x, x')$ which describes the model to the right from the hard-wall potential at $x = 0$ but uses the parameters of the left subsystem (see Appendix 3.3.1 for its explicit expression as well as for the adopted conventions regarding notations). To establish $d = [G^{(0,L)}(0, 0)]^{-1} - p_0$ we used the identity

$$[G^{(0,L)}(0)]^{-1} = -\frac{1}{8} \mathcal{A}_L \left[\lim_{x \rightarrow 0^-} \frac{d^2}{dx^2} G^L(x, x) + \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G_0^L(x, x) \right] \mathcal{A}_L, \quad (3.57)$$

which naturally arises in the translationally invariant case $\mathcal{U}_1 = 0$ and $(\mathcal{A}_L, \mathcal{B}_L, \mathcal{C}_L) = (\mathcal{A}_R, \mathcal{B}_R, \mathcal{C}_R)$.

Reminding ourselves that the boundary Green's functions $G^L(x, x')$ and $G^R(x, x')$, entering

$$G(x, x') = G^L(x, x') + G^R(x, x') + F(x) d^{-1} \bar{F}(x'),$$

$$d = \lim_{a \rightarrow 0} a D = -\mathcal{U}_1 - \frac{1}{8} \mathcal{A}_L \lim_{x \rightarrow 0^-} \frac{d^2}{dx^2} G^L(x, x) \mathcal{A}_L - \frac{1}{8} \mathcal{A}_R \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G^R(x, x) \mathcal{A}_R,$$

admit a simple representation (see Appendix 3.3.1) in terms of $G^{(0,L)}(x, x')$ and $G^{(0,R)}(x, x')$, respectively, we assert that the above construction completes our program of establishing $G(x, x')$ in the special $M = 1$ case.

To verify that $G(x, x') = G^L(x, x') + G^R(x, x') + F(x) d^{-1} \bar{F}(x')$ does provide the solution of the equation $[z - H]G(x, x'; z) = \delta(x - x')$ with the matching conditions $\Psi(x_m^-) = \Psi(x_m^+) \equiv \Psi(x_m)$ and $\frac{\mathcal{A}_m \Psi'(x_m^+) - \mathcal{A}_{m-1} \Psi'(x_m^-)}{2} = [\mathcal{U}_m - i \frac{\mathcal{B}_m - \mathcal{B}_{m-1}}{2}] \Psi(x_m)$, we do the following steps:

1) We note that both $G^L(x, x') + G^R(x, x')$ and $F(x)$ satisfy this differential equation.

2) On the basis of the relations $G^L(0, x') = G^R(0, x') = 0$ and $F(0^+) = F(0^-) \equiv F(0) = -1$ we establish the continuity of $G(x, x')$ at $x = 0$, i.e. the condition $\Psi(x_m^-) = \Psi(x_m^+) \equiv \Psi(x_m)$ is fulfilled.

3) We express the condition $\frac{\mathcal{A}_m \Psi'(x_m^+) - \mathcal{A}_{m-1} \Psi'(x_m^-)}{2} = [\mathcal{U}_m - i \frac{\mathcal{B}_m - \mathcal{B}_{m-1}}{2}] \Psi(x_m)$ for $G(x, x')$ in the form

$$\frac{\mathcal{A}_R}{2} [G_1^R(0^+, x') + F'(0^+) d^{-1} \bar{F}(x')] - \frac{\mathcal{A}_L}{2} [G_1^L(0^-, x') + F'(0^-) d^{-1} \bar{F}(x')] = - [\mathcal{U}_1 - i \frac{\mathcal{B}_R - \mathcal{B}_L}{2}] d^{-1} \bar{F}(x'), \quad (3.58)$$

and check whether it is fulfilled for $x' \neq 0$ (i.e. for $x' > 0^+$ and for $x' < 0^-$). Under this condition we identify $\bar{F}(x') = \frac{1}{2}\mathcal{A}_R G_1^R(0^+, x') - \frac{1}{2}\mathcal{A}_L G_1^L(0^-, x')$, and it remains to prove that

$$d = -\mathcal{U}_1 - \frac{\mathcal{A}_R}{2}F'(0^+) + \frac{\mathcal{A}_L}{2}F'(0^-) + i\frac{\mathcal{B}_R - \mathcal{B}_L}{2}. \quad (3.59)$$

The fulfillment of this condition is shown in Appendix 3.3.1. Thus it is finally justified that the derived $G(x, x')$ satisfies the both matching conditions.

It is remarkable that the formulas

$$d = -\mathcal{U}_1 - \frac{\mathcal{A}_R}{2}F'(0^+) + \frac{\mathcal{A}_L}{2}F'(0^-) + i\frac{\mathcal{B}_R - \mathcal{B}_L}{2}$$

$$F'(0^+) = -G_1^{(0,R)}(0^+, 0) [G^{(0,R)}(0, 0)]^{-1}$$

$$F'(0^-) = -G_1^{(0,L)}(0^-, 0) [G^{(0,L)}(0, 0)]^{-1}$$

represent the simplest way of determining the matrix d in the single-barrier case:

$$d = -\mathcal{U}_1 + \mathcal{L}_R - \mathcal{L}_L, \quad (3.60)$$

$$\mathcal{L}_R = \frac{\mathcal{A}_R}{2}G_1^{(0,R)}(0^+, 0)[G^{(0,R)}(0, 0)]^{-1} + \frac{i}{2}\mathcal{B}_R, \quad (3.61)$$

$$\mathcal{L}_L = \frac{\mathcal{A}_L}{2}G_1^{(0,L)}(0^-, 0)[G^{(0,L)}(0, 0)]^{-1} + \frac{i}{2}\mathcal{B}_L. \quad (3.62)$$

In this setting, the relation of d to the matching condition $\frac{\mathcal{A}_m \Psi'(x_m^+) - \mathcal{A}_{m-1} \Psi'(x_m^-)}{2} = \left[\mathcal{U}_m - i\frac{\mathcal{B}_m - \mathcal{B}_{m-1}}{2} \right] \Psi(x_m)$, expressing the conservation of the current density, becomes especially transparent. The importance of the objects $\mathcal{L}_{R/L}$ for the properties of the boundary charge in the half-space models has been recently elucidated in Ref. [41].

3.2.3 Double-barrier setup

So far we know that

$$G_X^m(x, x') = G^{(0,m)}(x, x') - G^{(0,m)}(x, X) \left[G^{(0,m)}(X, X) \right]^{-1} G^{(0,m)}(X, x'), \quad (3.63)$$

and

$$G_{X,Y}^m(x, x') = G^{(0,m)}(x, x') - G^{(0,m)}(x, X) [G_Y^m(X, X)]^{-1} G_Y^m(X, x') - G^{(0,m)}(x, Y) [G_X^m(Y, Y)]^{-1} G_X^m(Y, x').$$

(??? add words, what are they?)

From them we obtain

$$G_{0,W}^C(x, x') = g(x - x') - g(x) \left\{ g(0) - g(-W)[g(0)]^{-1}g(W) \right\}^{-1} \left\{ g(-x') - g(-W)[g(0)]^{-1}g(W - x') \right\} - g(x - W) \left\{ g(0) - g(W)[g(0)]^{-1}g(-W) \right\}^{-1} \left\{ g(W - x') - g(W)[g(0)]^{-1}g(-x') \right\}, \quad (3.64)$$

in terms of the translationally invariant counterpart $G^{(C,0)}(x, 0) \equiv g(x)$. According to $G^m(x, x') = \Theta(x_{m+1} > x > x_m) G_{x_m, x_{m+1}}^m(x, x')$, the boundary Green's function of the central region is then given by $G^C(x, x') = \Theta(W > x > 0) G_{0,W}^C(x, x')$.

Taking in account that $G_{0,W}^C(x, x')$ identically vanishes when x and x' belong to different spatial regions [either $(-\infty, 0)$ or $(0, W)$ or (W, ∞)], we get the identities

$$\frac{\partial^2}{\partial x \partial x'} G_{0,W}^C(0^-, W^+) = \frac{\partial^2}{\partial x \partial x'} G_{0,W}^C(0^-, W^-) = \frac{\partial^2}{\partial x \partial x'} G_{0,W}^C(0^+, W^+) = 0. \quad (3.65)$$

They allow us to replace

$$\begin{aligned} & \frac{1}{4} \mathcal{A}_C G_{12}^C(0^+, W^-) \mathcal{A}_C = \\ & = \frac{1}{4} \mathcal{A}_C \left[\frac{\partial^2 G_{0,W}^C(0^+, W^-)}{\partial x \partial x'} - \frac{\partial^2 G_{0,W}^C(0^-, W^-)}{\partial x \partial x'} + \frac{\partial^2 G_{0,W}^C(0^-, W^+)}{\partial x \partial x'} - \frac{\partial^2 G_{0,W}^C(0^+, W^+)}{\partial x \partial x'} \right] \mathcal{A}_C. \end{aligned} \quad (3.66)$$

$$(3.67)$$

After the lengthy calculation using the formula for $G_{0,W}^C(x, x')$ and $[G_2^{(0,m)}(0, 0^+) - G_2^{(0,m)}(0, 0^-)] \frac{\mathcal{A}_m}{2} = 1$, $\frac{\mathcal{A}_m}{2} [G_1^{(0,m)}(0^+, 0) - G_1^{(0,m)}(0^-, 0)] = 1$ we obtain

$$\frac{1}{4} \mathcal{A}_C G_{12}^C(0^+, W^-) \mathcal{A}_C = \left\{ g(W) - g(0)[g(-W)]^{-1}g(0) \right\}^{-1}. \quad (3.68)$$

Analogously we deduce

$$\frac{1}{4} \mathcal{A}_C G_{12}^C(W^-, 0^+) \mathcal{A}_C = \left\{ g(-W) - g(0)[g(W)]^{-1}g(0) \right\}^{-1}. \quad (3.69)$$

To derive

$$-\frac{1}{8} \mathcal{A}_C \left[\lim_{x \rightarrow 0^-} \frac{d^2}{dx^2} G_0^C(x, x) + \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G^C(x, x) \right] \mathcal{A}_C = \left\{ g(0) - g(-W)[g(0)]^{-1}g(W) \right\}^{-1}, \quad (3.70)$$

we use the identities

$$\frac{\partial^2}{\partial x \partial x'} G_{0,W}^C(0^+, 0^-) = \frac{\partial^2}{\partial x \partial x'} G_{0,W}^C(0^-, 0^+) = 0 \quad (3.71)$$

to replace

$$\begin{aligned} & -\frac{1}{8} \mathcal{A}_C \left[\lim_{x \rightarrow 0^-} \frac{d^2}{dx^2} G_0^C(x, x) + \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G^C(x, x) \right] \mathcal{A}_C \\ & = -\frac{1}{8} \mathcal{A}_C \left[\lim_{x \rightarrow 0^-} \frac{d^2 G_{0,W}^C(x, x)}{dx^2} + \lim_{x \rightarrow 0^+} \frac{d^2 G_{0,W}^C(x, x)}{dx^2} - 2 \frac{\partial^2 G_{0,W}^C(0^+, 0^-)}{\partial x \partial x'} - 2 \frac{\partial^2 G_{0,W}^C(0^-, 0^+)}{\partial x \partial x'} \right] \mathcal{A}_C. \end{aligned} \quad (3.72)$$

After the lengthy calculation (????) using formula for $G_{0,W}^C(x, x')$ and $[G_2^{(0,m)}(0, 0^+) - G_2^{(0,m)}(0, 0^-)] \frac{\mathcal{A}_m}{2} = 1$, $\frac{\mathcal{A}_m}{2} [G_1^{(0,m)}(0^+, 0) - G_1^{(0,m)}(0^-, 0)] = 1$ we got what we stated.

Analogously we deduce

$$-\frac{1}{8} \mathcal{A}_C \left[\lim_{x \rightarrow W^+} \frac{d^2}{dx^2} G_W^C(x, x) + \lim_{x \rightarrow W^-} \frac{d^2}{dx^2} G^C(x, x) \right] \mathcal{A}_C = \left\{ g(0) - g(W)[g(0)]^{-1}g(-W) \right\}^{-1}. \quad (3.73)$$

3.2.4 Multiple barriers

USELESS AND TOO COMPLICATED TOPIC, there is no need to look at it! FIRST, MASTER 2-BARRIER CASE

(below is a collection of results, I'll add real derivation if need it)

Overview of the results

A direct generalization of the single-barrier result $G(x, x') = G^L(x, x') + G^R(x, x') + F(x)d^{-1}\bar{F}(x')$ to the case of multiple barriers is given by the expression (see Appendix 3.2.4 for its derivation)

$$G(x, x') = \sum_{m=0}^M G^m(x, x') + \sum_{m,m'=1}^M F_m(x)(d^{-1})_{m,m'}\bar{F}_{m'}(x'), \quad (3.74)$$

where d is a block tridiagonal matrix, with the additional barrier indices m, m' labeling the blocks. The diagonal blocks

$$d_{m,m} = -\mathcal{U}_m - \frac{1}{8}\mathcal{A}_{m-1} \lim_{x \rightarrow x_m^-} \frac{d^2}{dx^2} G^{m-1}(x, x) \mathcal{A}_{m-1} - \frac{1}{8}\mathcal{A}_m \lim_{x \rightarrow x_m^+} \frac{d^2}{dx^2} G^m(x, x) \mathcal{A}_m \quad (3.75)$$

is a generalization of

$$d = \lim_{a \rightarrow 0} a D = -\mathcal{U}_1 - \frac{1}{8}\mathcal{A}_L \lim_{x \rightarrow 0^-} \frac{d^2}{dx^2} G^L(x, x) \mathcal{A}_L - \frac{1}{8}\mathcal{A}_R \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G^R(x, x) \mathcal{A}_R$$

They are defined locally at the interface positions x_m . In turn, the off-diagonal blocks

$$d_{m,m+1} = \frac{1}{4}\mathcal{A}_m G_{12}^m(x_m^+, x_{m+1}^-) \mathcal{A}_m, \quad (3.76)$$

$$d_{m+1,m} = \frac{1}{4}\mathcal{A}_m G_{12}^m(x_{m+1}^-, x_m^+) \mathcal{A}_m \quad (3.77)$$

describe the propagation between the two adjacent barriers x_m and x_{m+1} across the m -th subsystem. These blocks thereby account for the quantum interference effects.

The functions

$$F_m(x) = -\Theta(x_{m-1} < x < x_m) \left[G_2^{(0,m-1)}(x, x_m^+) + G_2^{m-1}(x, x_m^-) - G_2^{(0,m-1)}(x, x_m^-) \right] \frac{\mathcal{A}_{m-1}}{2} \\ + \Theta(x_m < x < x_{m+1}) \left[G_2^{(0,m)}(x, x_m^-) + G_2^m(x, x_m^+) - G_2^{(0,m)}(x, x_m^+) \right] \frac{\mathcal{A}_m}{2}, \quad (3.78)$$

$$\bar{F}_{m'}(x') = -\Theta(x_{m'-1} < x' < x_{m'}) \frac{\mathcal{A}_{m'-1}}{2} \left[G_1^{(0,m'-1)}(x_{m'}^+, x') + G_1^{m'-1}(x_{m'}^-, x') - G_1^{(0,m'-1)}(x_{m'}^-, x') \right] \\ + \Theta(x_{m'} < x' < x_{m'+1}) \frac{\mathcal{A}_{m'}}{2} \left[G_1^{(0,m')}(x_{m'}^-, x') + G_1^{m'}(x_{m'}^+, x') - G_1^{(0,m')}(x_{m'}^+, x') \right] \quad (3.79)$$

generalize the expressions

$$F(x) = \lim_{a \rightarrow 0} F_n = -\Theta(-x) \left[G_2^{(0,L)}(x, 0^+) + G_2^L(x, 0^-) - G_2^{(0,L)}(x, 0^-) \right] \frac{\mathcal{A}_L}{2} \\ + \Theta(x) \left[G_2^{(0,R)}(x, 0^-) + G_2^R(x, 0^+) - G_2^{(0,R)}(x, 0^+) \right] \frac{\mathcal{A}_R}{2}, \quad (3.80)$$

$$\bar{F}(x') = \lim_{a \rightarrow 0} \bar{F}_{n'} = -\Theta(-x') \frac{\mathcal{A}_L}{2} \left[G_1^{(0,L)}(0^+, x') + G_1^L(0^-, x') - G_1^{(0,L)}(0^-, x') \right] \\ + \Theta(x') \frac{\mathcal{A}_R}{2} \left[G_1^{(0,R)}(0^-, x') + G_1^R(0^+, x') - G_1^{(0,R)}(0^+, x') \right] \quad (3.81)$$

respectively. They also possess the properties

$$F_m(x_{m'}) = -\delta_{m,m'}, \quad (3.82)$$

$$\bar{F}_{m'}(x_m) = -\delta_{m',m}, \quad (3.83)$$

analogous to $F_0 = \bar{F}_0 = -1$ in the single-barrier case. On their basis we establish that

$$G(x_m, x_{m'}) = (d^{-1})_{m,m'}, \quad (3.84)$$

and thus reveal the physical meaning of the matrix d : It is the inverse of the propagator between the contacts x_m and $x_{m'}$.

The global single-barrier DOS $\rho(\omega) = \rho^L(\omega) + \rho^R(\omega) - \frac{1}{\pi} \text{Im} \frac{\partial}{\partial \omega} \ln \det d(\omega + i0^+)$ has a straightforward multi-barrier generalization (see Appendix 3.2.4)

$$\rho(\omega) = \sum_{m=0}^M \rho^m(\omega) - \frac{1}{\pi} \text{Im} \frac{\partial}{\partial \omega} \ln \det d(\omega + i0^+), \quad (3.85)$$

where the determinant of $d(\omega + i0^+)$ is now evaluated for the $N_c M \times N_c M$ matrix.

The bound state equation $\det d(\omega) = 0$ retains its form in the multi-barrier case.

In the double-barrier case $M = 2$ we can further simplify the matrix d .

We assign the values $x_1 = 0$ and $x_2 = W$ to the contact coordinates, such that W is the width of the central (C) region. We also assume that both the left (L) and the right (R) subsystems are semi-infinite, and in the following we use the labeling of the regions $m = L, C, R$ instead of $m = 0, 1, 2$. Owing to the expressions

$$\begin{aligned} \frac{1}{4} \mathcal{A}_C G_{12}^C(0^+, W^-) \mathcal{A}_C &= \{g(W) - g(0)[g(-W)]^{-1}g(0)\}^{-1}, \\ \frac{1}{4} \mathcal{A}_C G_{12}^C(W^-, 0^+) \mathcal{A}_C &= \{g(-W) - g(0)[g(W)]^{-1}g(0)\}^{-1}, \\ -\frac{1}{8} \mathcal{A}_C \left[\lim_{x \rightarrow 0^-} \frac{d^2}{dx^2} G_0^C(x, x) + \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G^C(x, x) \right] \mathcal{A}_C &= \{g(0) - g(-W)[g(0)]^{-1}g(W)\}^{-1}, \\ -\frac{1}{8} \mathcal{A}_C \left[\lim_{x \rightarrow W^+} \frac{d^2}{dx^2} G_W^C(x, x) + \lim_{x \rightarrow W^-} \frac{d^2}{dx^2} G^C(x, x) \right] \mathcal{A}_C &= \{g(0) - g(W)[g(0)]^{-1}g(-W)\}^{-1}. \end{aligned} \quad (3.86)$$

derived in Appendix 3.2.3 we state

$$d_{\text{tw bar}} \equiv \begin{pmatrix} G(0, 0) & G(0, W) \\ G(W, 0) & G(W, W) \end{pmatrix}^{-1} = \begin{pmatrix} G^{(C,0)}(0, 0) & G^{(C,0)}(0, W) \\ G^{(C,0)}(W, 0) & G^{(C,0)}(W, W) \end{pmatrix}^{-1} - \begin{pmatrix} p_0 & 0 \\ 0 & p_W \end{pmatrix}. \quad (3.87)$$

The last term can be interpreted as the self-energy term breaking the translational invariance of the auxiliary infinite-space model, which uses the parameters of the central subsystem. It is expressed in terms of the energy-dependent matrices

$$p_0 = \mathcal{U}_1 + \frac{1}{8} \lim_{x \rightarrow 0^-} \left[\mathcal{A}_L \frac{d^2}{dx^2} G^L(x, x) \mathcal{A}_L - \mathcal{A}_C \frac{d^2}{dx^2} G_0^C(x, x) \mathcal{A}_C \right], \quad (3.88)$$

$$p_W = \mathcal{U}_2 + \frac{1}{8} \lim_{x \rightarrow W^+} \left[\mathcal{A}_R \frac{d^2}{dx^2} G^R(x, x) \mathcal{A}_R - \mathcal{A}_C \frac{d^2}{dx^2} G_W^C(x, x) \mathcal{A}_C \right]. \quad (3.89)$$

Hereby we introduced the auxiliary Green's functions $G_0^C(x, x')$ and $G_W^C(x, x')$ describing the models to the left from the hard-wall potential at $x = 0$ and to the right from the hard-wall potential at $x = W$, respectively, and both using the parameters of the central subsystem (see Appendix 3.3.1 for their explicit expressions).

Applying the formulas

$$\mathcal{A}_R F'(0^+) - iB_R = \frac{1}{4} \mathcal{A}_R \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G^R(x, x) \mathcal{A}_R, \quad (3.90)$$

$$\mathcal{A}_L F'(0^-) - iB_L = -\frac{1}{4} \mathcal{A}_L \lim_{x \rightarrow 0^-} \frac{d^2}{dx^2} G^L(x, x) \mathcal{A}_L. \quad (3.91)$$

we further simplify p_0, p_W to

$$p_0 = \mathcal{U}_1 + \mathcal{L}_L - \mathcal{L}_{L \rightarrow C}, \quad (3.92)$$

$$p_W = \mathcal{U}_2 - \mathcal{L}_R + \mathcal{L}_{R \rightarrow C}, \quad (3.93)$$

where

$$\begin{aligned} \mathcal{L}_R &= \frac{\mathcal{A}_R}{2} G_1^{(0,R)}(0^+, 0) [G^{(0,R)}(0, 0)]^{-1} + \frac{i}{2} \mathcal{B}_R \\ \mathcal{L}_L &= \frac{\mathcal{A}_L}{2} G_1^{(0,L)}(0^-, 0) [G^{(0,L)}(0, 0)]^{-1} + \frac{i}{2} \mathcal{B}_L, \end{aligned}$$

$$\mathcal{L}_{L \rightarrow C} = \frac{\mathcal{A}_C}{2} G_1^{(0,C)}(0^-, 0) [G^{(0,C)}(0, 0)]^{-1} + \frac{i}{2} \mathcal{B}_C, \quad (3.94)$$

$$\begin{aligned} \mathcal{L}_{R \rightarrow C} &= \frac{\mathcal{A}_C}{2} G_1^{(0,C)}(0^+, 0) [G^{(0,C)}(0, 0)]^{-1} + \frac{i}{2} \mathcal{B}_C \\ &= \mathcal{L}_{L \rightarrow C} + [G^{(0,C)}(0, 0)]^{-1}. \end{aligned} \quad (3.95)$$

These formulas allow one to express the matrix d in the double-barrier case given by $d_{\text{tw bar}}$ in terms of the objects $\mathcal{L}$ containing only first derivatives of the translationally invariant Green's functions $G^{(0,m)}$.

Green's function derivation for the multi-barrier model

Let us define positions of arbitrary M barriers as $n = n_1, \dots, n_M$ and assigning the local onsite potentials W_{n_m} to each of them, we also introduce $M + 1$ disjoint subsystems defined on the intervals $n_m + 1 \leq n \leq n_{m+1} - 1$ and describe them by the Green's functions G^m .

Switching on the coupling $\sum_{m=1}^M v_m$ between the disjoint subsystems via the hoppings through the barrier sites n_m ,

$$v_m = -|n_m\rangle t_{n_m-1} \langle n_m-1| - |n_m-1\rangle t_{n_m-1}^\dagger \langle n_m| - |n_m+1\rangle t_{n_m} \langle n_m| - |n_m\rangle t_{n_m}^\dagger \langle n_m+1|, \quad (3.96)$$

we set up the Dyson equation treating v_m 's as a perturbation. It has the form

$$\begin{aligned} G_{n,n'} &= G_{n,n'}^{\{m\}} - \sum_{m'=1}^M \left(G_{n,n_{m'}-1}^{m'-1} t_{n_{m'}-1}^\dagger + G_{n,n_{m'}+1}^{m'} t_{n_{m'}} \right) G_{n_{m'},n'} \\ &\quad - \sum_{m'=1}^M \delta_{n,n_{m'}} g_{n_{m'},n_{m'}}^{m'} \left(t_{n_{m'}-1} G_{n_{m'}-1,n'} + t_{n_{m'}}^\dagger G_{n_{m'}+1,n'} \right) \end{aligned} \quad (3.97)$$

generalizing $G_{n,n'} = G_{n,n'}^L + G_{n,n'}^C + G_{n,n'}^R - (G_{n,-1}^L t_L^\dagger + G_{n,1}^R t_R) G_{0,n'} - \delta_{n,0} G_{0,0}^C (t_L G_{-1,n'} + t_R^\dagger G_{1,n'})$.

Hereby $G_{n,n'}^{\{m\}}$ contains not only all Green's functions G^m of the finite-ranged disjoint subsystems, but also those $g_{n,n'}^m = \frac{\delta_{n,n_m}\delta_{n',n_m}}{z-W_{n_m}}$ of the barrier sites n_m .

Below we find the solution of the Dyson equation for $G_{n,n'}$ following the same routine as in Section 3.2.2.

Introducing the M -component vectors

$$\bar{\mathcal{G}}_{m'}^{(\pm)} = G_{n_{m'} \pm 1, n'}, \quad (3.98)$$

$$\bar{\mathcal{G}}_{m'}^{(0)} = G_{n_{m'}, n'}, \quad (3.99)$$

we first derive the equation

$$\begin{pmatrix} \bar{\mathcal{G}}^{(+)} \\ \bar{\mathcal{G}}^{(-)} \\ \bar{\mathcal{G}}^{(0)} \end{pmatrix} = \begin{pmatrix} 1 & 0 & -T^{+,0} \\ 0 & 1 & -T^{-,0} \\ -T^{0,+} & -T^{0,-} & 1 \end{pmatrix}^{-1} \begin{pmatrix} \bar{\mathcal{G}}_{\text{disj}}^{(+)} \\ \bar{\mathcal{G}}_{\text{disj}}^{(-)} \\ \bar{\mathcal{G}}_{\text{disj}}^{(0)} \end{pmatrix}, \quad (3.100)$$

where

$$\begin{aligned} T_{m,m'}^{0,+} &:= -\delta_{m,m'} g_{n_{m'}, n_{m'}}^{m'} t_{n_{m'}}^\dagger, \\ \bar{\mathcal{G}}_{\text{disj}, m'}^{(+)} &:= G_{n_{m'}+1, n'}^{m'}, \\ T_{m,m'}^{0,-} &:= -\delta_{m,m'} g_{n_{m'}, n_{m'}}^{m'} t_{n_{m'}-1}, \\ \bar{\mathcal{G}}_{\text{disj}, m'}^{(-)} &:= G_{n_{m'}-1, n'}^{m'-1}, \\ T_{m,m'}^{+,0} &:= -\delta_{m,m'-1} G_{n_{m'}-1+1, n_{m'}-1}^{m'-1} t_{n_{m'}-1}^\dagger - \delta_{m,m'} G_{n_{m'}+1, n_{m'}+1}^{m'} t_{n_{m'}}, \\ \bar{\mathcal{G}}_{\text{disj}, m'}^{(0)} &:= g_{n_{m'}, n_{m'}}^{m'} \delta_{n_{m'}, n'}, \\ T_{m,m'}^{-,0} &:= -\delta_{m,m'} G_{n_{m'}-1, n_{m'}-1}^{m'-1} t_{n_{m'}-1}^\dagger - \delta_{m,m'+1} G_{n_{m'}+1-1, n_{m'}+1}^{m'} t_{n_{m'}}. \end{aligned}$$

Finding explicitly the solutions for $\bar{\mathcal{G}}_{m'}^{(\pm)} = G_{n_{m'} \pm 1, n'}$, $\bar{\mathcal{G}}_{m'}^{(0)} = G_{n_{m'}, n'}$, as follows:

$$G_{n_m, n'} = - \sum_{m'=1}^M (D^{-1})_{m, m'} \bar{F}_{m', n'}, \quad (3.101)$$

$$g_{n_m, n_m}^m \left(t_{n_m-1} G_{n_m-1, n'} + t_{n_m}^\dagger G_{n_m+1, n'} \right) = g_{n_m, n_m}^m \delta_{n_m, n'} + \sum_{m'=1}^M (D^{-1})_{m, m'} \bar{F}_{m', n'}, \quad (3.102)$$

we insert them into

$$\begin{aligned} G_{n, n'} &= G_{n, n'}^{\{m\}} - \sum_{m'=1}^M \left(G_{n, n_{m'}-1}^{m'-1} t_{n_{m'}-1}^\dagger + G_{n, n_{m'}+1}^{m'} t_{n_{m'}} \right) G_{n_{m'}, n'} \\ &\quad - \sum_{m'=1}^M \delta_{n, n_{m'}} g_{n_{m'}, n_{m'}}^{m'} \left(t_{n_{m'}-1} G_{n_{m'}-1, n'} + t_{n_{m'}}^\dagger G_{n_{m'}+1, n'} \right) \end{aligned} \quad (3.103)$$

Thereby we establish the composite Green's function of the whole system

$$G_{n, n'}^{\text{multi}} = \sum_{m=0}^M G_{n, n'}^m + \sum_{m, m'=1}^M F_{n, m} (D^{-1})_{m, m'} \bar{F}_{m', n'}, \quad (3.104)$$

where

$$D_{m,m'} = \delta_{m,m'} \left(z - W_{n_m} - t_{n_m-1} G_{n_m-1,n_m-1}^{m-1} t_{n_m-1}^\dagger - t_{n_m}^\dagger G_{n_m+1,n_m+1}^m t_{n_m} \right) - \delta_{m,m'+1} t_{n_m-1} G_{n_m-1,n_m-1+1}^{m-1} t_{n_m-1} - \delta_{m,m'-1} t_{n_m}^\dagger G_{n_m+1,n_m+1-1}^m t_{n_m+1-1}^\dagger, \quad (3.105)$$

$$F_{n,m} = -\delta_{n,n_m} + G_{n,n_m-1}^{m-1} t_{n_m-1}^\dagger + G_{n,n_m+1}^m t_{n_m}, \quad (3.106)$$

$$\bar{F}_{m',n'} = -\delta_{n_{m'},n'} + t_{n_{m'}-1} G_{n_{m'}-1,n'}^{m'-1} + t_{n_{m'}}^\dagger G_{n_{m'}+1,n'}^{m'}. \quad (3.107)$$

$G_{n,n'}^{\text{multi}}$ is a multi-barrier generalization of the single-barrier $G_{n,n'} = G_{n,n'}^L + G_{n,n'}^R + F_n D^{-1} \bar{F}_{n'}$. Observing that

$$\sum_n \bar{F}_{m',n} F_{n,m} = \frac{\partial}{\partial \omega} D_{m',m}, \quad (3.108)$$

(???? why???) we also express the global DOS in the multi-barrier setup by the analogy with the single-barrier formula $\rho(\omega) = \rho^L(\omega) + \rho^R(\omega) - \frac{1}{\pi} \text{Im} \frac{\partial}{\partial \omega} \ln \det D(\omega + i0^+)$ as

$$\rho(\omega) = \sum_{m=0}^M \rho^m(\omega) - \frac{1}{\pi} \text{Im} \frac{\partial}{\partial \omega} \ln \det D(\omega + i0^+). \quad (3.109)$$

To take the continuum limit of $D_{m,m'}$ we approximate in $D_{m,m'}$ [cf. Eq. $W_0 = \frac{\mathcal{A}_L + \mathcal{A}_R}{2a^2} + \frac{\mathcal{U}_1}{a}$]

$$W_{n_m} \approx \frac{\mathcal{A}_{m-1} + \mathcal{A}_m}{2a^2} + \frac{\mathcal{U}_m}{a} \quad (3.110)$$

and

$$t_{n_m-1} G_{n_m-1,n_m-1}^{m-1} t_{n_m-1}^\dagger \approx -\frac{\mathcal{A}_{m-1}}{2a^2} + \frac{1}{8a} \mathcal{A}_{m-1} \lim_{x \rightarrow x_m^-} \frac{d^2}{dx^2} G^{m-1}(x, x) \mathcal{A}_{m-1}, \quad (3.111)$$

$$(3.112)$$

$$t_{n_m}^\dagger G_{n_m+1,n_m+1}^m t_{n_m} \approx -\frac{\mathcal{A}_m}{2a^2} + \frac{1}{8a} \mathcal{A}_m \lim_{x \rightarrow x_m^+} \frac{d^2}{dx^2} G^m(x, x) \mathcal{A}_m, \quad (3.113)$$

analogously to the single-barrier case discussed in Appendix 3.3.1. In addition, there are two new objects in $D_{m,m'}$, which are approximated by

$$t_{n_m-1} G_{n_m-1,n_m-1+1}^{m-1} t_{n_m-1} \approx -\frac{1}{4a} \mathcal{A}_{m-1} G_{12}^{m-1}(x_m^-, x_{m-1}^+) \mathcal{A}_{m-1}, \quad (3.114)$$

$$(3.115)$$

$$t_{n_m}^\dagger G_{n_m+1,n_m+1-1}^m t_{n_m+1-1}^\dagger \approx -\frac{1}{4a} \mathcal{A}_m G_{12}^m(x_m^+, x_{m+1}^-) \mathcal{A}_m. \quad (3.116)$$

Altogether these terms provide the continuum limit of $D_{m,m'}$ giving rise to the expressions

$$d_{m,m} = -\mathcal{U}_m - \frac{1}{8} \mathcal{A}_{m-1} \lim_{x \rightarrow x_m^-} \frac{d^2}{dx^2} G^{m-1}(x, x) \mathcal{A}_{m-1} - \frac{1}{8} \mathcal{A}_m \lim_{x \rightarrow x_m^+} \frac{d^2}{dx^2} G^m(x, x) \mathcal{A}_m \quad (3.117)$$

is a generalization of

$$d = \lim_{a \rightarrow 0} a D = -\mathcal{U}_1 - \frac{1}{8} \mathcal{A}_L \lim_{x \rightarrow 0^-} \frac{d^2}{dx^2} G^L(x, x) \mathcal{A}_L - \frac{1}{8} \mathcal{A}_R \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G^R(x, x) \mathcal{A}_R$$

They are defined locally at the interface positions x_m . In turn, the off-diagonal blocks

$$d_{m,m+1} = \frac{1}{4} \mathcal{A}_m G_{12}^m(x_m^+, x_{m+1}^-) \mathcal{A}_m, \quad (3.118)$$

$$d_{m+1,m} = \frac{1}{4} \mathcal{A}_m G_{12}^m(x_{m+1}^-, x_m^+) \mathcal{A}_m \quad (3.119)$$

defining the matrix $d = \lim_{a \rightarrow 0} a D$ in the multi-barrier case.

A derivation of the continuum limit of

$$F_{n,m} = -\delta_{n,n_m} + G_{n,n_m-1}^{m-1} t_{n_m-1}^\dagger + G_{n,n_m+1}^m t_{n_m},$$

$$\bar{F}_{m',n'} = -\delta_{n_m',n'} + t_{n_m'-1} G_{n_m'-1,n'}^{m'-1} + t_{n_m'}^\dagger G_{n_m'+1,n'}^{m'}.$$

employs the same type of the a -expansion as in

$$G_{n,-1}^L = G_{n,-1}^L - G_{n,0}^L \approx \quad (3.120)$$

$$\approx -a^2 G_2^{(0,L)}(x, 0^+) + a^2 G^{(0,L)}(x, 0) [G^{(0,L)}(0, 0)]^{-1} G_2^{(0,L)}(0, 0^-) = \quad (3.121)$$

$$= -a^2 [G_2^{(0,L)}(x, 0^+) + G_2^L(x, 0^-) - G_2^{(L,0)}(x, 0^-)], \quad (3.122)$$

$$(3.123)$$

$$G_{n,1}^R = G_{n,1}^R - G_{n,0}^R \approx \quad (3.124)$$

$$\approx a^2 G_2^{(0,R)}(x, 0^-) - a^2 G^{(0,R)}(x, 0) [G^{(0,R)}(0, 0)]^{-1} G_2^{(0,R)}(0, 0^+)$$

$$= a^2 [G_2^{(0,R)}(x, 0^-) + G_2^R(x, 0^+) - G_2^{(0,R)}(x, 0^+)], \quad (3.125)$$

It leads us to the final result

$$F_m(x) = -\Theta(x_{m-1} < x < x_m) \left[G_2^{(0,m-1)}(x, x_m^+) + G_2^{m-1}(x, x_m^-) - G_2^{(0,m-1)}(x, x_m^-) \right] \frac{\mathcal{A}_{m-1}}{2} \\ + \Theta(x_m < x < x_{m+1}) \left[G_2^{(0,m)}(x, x_m^-) + G_2^m(x, x_m^+) - G_2^{(0,m)}(x, x_m^+) \right] \frac{\mathcal{A}_m}{2}, \quad (3.126)$$

$$\bar{F}_{m'}(x') = -\Theta(x_{m'-1} < x' < x_{m'}) \frac{\mathcal{A}_{m'-1}}{2} \left[G_1^{(0,m'-1)}(x_{m'}^+, x') + G_1^{m'-1}(x_{m'}^-, x') - G_1^{(0,m'-1)}(x_{m'}^-, x') \right] \\ + \Theta(x_{m'} < x' < x_{m'+1}) \frac{\mathcal{A}_{m'}}{2} \left[G_1^{(0,m')} (x_{m'}^-, x') + G_1^{m'}(x_{m'}^+, x') - G_1^{(0,m')} (x_{m'}^+, x') \right] \quad (3.127)$$

analogous to the result

$$F(x) = \lim_{a \rightarrow 0} F_n = -\Theta(-x) \left[G_2^{(0,L)}(x, 0^+) + G_2^L(x, 0^-) - G_2^{(0,L)}(x, 0^-) \right] \frac{\mathcal{A}_L}{2} \\ + \Theta(x) \left[G_2^{(0,R)}(x, 0^-) + G_2^R(x, 0^+) - G_2^{(0,R)}(x, 0^+) \right] \frac{\mathcal{A}_R}{2}, \quad (3.128)$$

$$\bar{F}(x') = \lim_{a \rightarrow 0} \bar{F}_{n'} = -\Theta(-x') \frac{\mathcal{A}_L}{2} \left[G_1^{(0,L)}(0^+, x') + G_1^L(0^-, x') - G_1^{(0,L)}(0^-, x') \right] \\ + \Theta(x') \frac{\mathcal{A}_R}{2} \left[G_1^{(0,R)}(0^-, x') + G_1^R(0^+, x') - G_1^{(0,R)}(0^+, x') \right] \quad (3.129)$$

of the single-barrier case.

3.2.5 Easy illustrative examples

Example 1

Example 2

3.3 Interface Green's functions for continuum models

(тут то, что на самом деле нужно)

3.3.1 Boundary Green's functions

For completeness of the exposition, we shall quote some basic results on the propagators of bounded systems. For a detailed account of the subject, we refer an interested reader to Refs. [54] (тут много развивающей возни с функциями Грина, не понятно, что от туда нужно конкретно для этой статьи), [55] (??? тут вские штуки, которые не нужно знать, ну и общие формулы, лучше не открывать вообще без потребности), while for various interesting applications to transport and solid state phenomena see Refs. [41, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65]. (? откуда вообще потребность в таком параде цитат? итак можно в интернет зайти и самим найти шутк 20 статей разных)

Derivation of the discrete limit in the single-barrier model

Consider a model described by Green's function $G^{(0,m)}(x, x')$. Inserting an ultra-local infinite-height potential at $x = X$, which creates a hard-wall boundary, we separate our initial system into two disjoint subsystems $x > X$ and $x < X$. The Green's functions of the both of them are given by the expression

$$G_X^m(x, x') = G^{(0,m)}(x, x') - G^{(0,m)}(x, X) [G^{(0,m)}(X, X)]^{-1} G^{(0,m)}(X, x'), \quad (3.130)$$

satisfying the boundary conditions $G_X^m(X, x') = G_X^m(x, X) = 0$. It also holds that $G_X^m(x, x')$ identically vanishes when either $x > X, x' < X$ or $x < X, x' > X$.

The expression for $G_X^m(x, x')$ also has the lattice analog (see e.g. in Ref. [65])

$$G_{n,n'}^m = G_{n,n'}^{(0,m)} - G_{n,0}^{(0,m)} [G_{0,0}^{(0,m)}]^{-1} G_{0,n'}^{(0,m)} \quad (3.131)$$

for a nearest-neighbor tight-binding model with the infinite-height potential at the site $n = 0$.

Adding the second hard-wall boundary at $x = Y$, we obtain the Green's function

$$G_{X,Y}^m(x, x') = G_X^m(x, x') - G_X^m(x, Y) [G_X^m(Y, Y)]^{-1} G_X^m(Y, x') = \quad (3.132)$$

$$= G_Y^m(x, x') - G_Y^m(x, X) [G_Y^m(X, X)]^{-1} G_Y^m(X, x'), \quad (3.133)$$

satisfying the boundary conditions $G_{X,Y}^m(X, x') = G_{X,Y}^m(x, X) = G_{X,Y}^m(Y, x') = G_{X,Y}^m(x, Y) = 0$. It can be also represented in the symmetrized form

$$G_{X,Y}^m(x, x') = G^{(0,m)}(x, x') - G^{(0,m)}(x, X) [G_Y^m(X, X)]^{-1} G_Y^m(X, x') - G^{(0,m)}(x, Y) [G_X^m(Y, Y)]^{-1} G_X^m(Y, x').$$

Assuming that $X < Y$, we observe that the Green's function $G_{X,Y}^m(x, x')$ identically vanishes when x and x' belong to different intervals (smaller than X , between X and Y , larger than Y). In addition, the following identities hold:

$$G_{X,Y}^m(x, x') = G_X^m(x, x') \quad x, x' < X, \quad (3.134)$$

$$G_{X,Y}^m(x, x') = G_Y^m(x, x') \quad x, x' > Y. \quad (3.135)$$

With these definitions, we obtain the following expression for the boundary Green's functions introduced in the main text:

$$G^m(x, x') = \Theta(x_{m+1} > x > x_m) G_{x_m, x_{m+1}}^m(x, x'), \quad (3.136)$$

where we assumed that neither $x_0 = -\infty$ nor $x_{M+1} = \infty$. In these special cases, we get simpler expressions on the basis of

$$G_X^m(x, x') = G^{(0,m)}(x, x') - G^{(0,m)}(x, X) \left[G^{(0,m)}(X, X) \right]^{-1} G^{(0,m)}(X, x'), \quad (3.137)$$

as follows

$$G^0(x, x') = \Theta(x_1 - x) G_{x_1}^0(x, x'), \quad x_0 = -\infty, \quad (3.138)$$

$$G^M(x, x') = \Theta(x - x_M) G_{x_M}^M(x, x'), \quad x_{M+1} = \infty. \quad (3.139)$$

Derivation of the continuum limit in the single-barrier model

To derive the continuum limit $a \rightarrow 0$ of the lattice Green's functions $G(x, x') = G^L(x, x') + G^R(x, x') + F(x)d^{-1}\bar{F}(x')$ we expand $G_{-1,-1}^L$ and $G_{1,1}^R$ by two orders higher than the leading one: This is necessary to cancel the leading $O(\frac{1}{a^2})$ term in $D := z - W_0 - t_L G_{-1,-1}^L t_L^\dagger - t_R^\dagger G_{1,1}^R t_R$ and to correctly extract the subleading $O(\frac{1}{a})$ contribution.

Now we will do an expansion, for which we need the formulas below:

$$G_{-1,0}^L = G_{0,-1}^L = G_{0,0}^L = 0; \quad G_{1,0}^R = G_{0,1}^R = G_{0,0}^R = 0,$$

which are inherent to the boundary Green's functions. We also can use the formula $G_{n,n'}^m = G_{n,n'}^{(0,m)} - G_{n,0}^{(0,m)} [G_{0,0}^{(0,m)}]^{-1} G_{0,n'}^{(0,m)}$ for our case:

$$\begin{aligned} G_{-1,0}^L &= G_{-1,0}^{(0,L)} - G_{-1,0}^{(0,L)} [G_{0,0}^{(0,L)}]^{-1} G_{0,0}^{(0,L)} \\ G_{0,-1}^L &= G_{0,-1}^{(0,L)} - G_{0,0}^{(0,L)} [G_{0,0}^{(0,L)}]^{-1} G_{0,-1}^{(0,L)} \\ G_{0,0}^L &= G_{0,0}^{(0,L)} - G_{0,0}^{(0,L)} [G_{0,0}^{(0,L)}]^{-1} G_{0,0}^{(0,L)} \end{aligned}$$

Finally,

$$G_{-1,-1}^{(0,L)} = G_{0,0}^{(0,L)}.$$

(???) And we do the transformation:

$$G_{-1,-1}^L = G_{0,0}^{(0,L)} - G_{-1,0}^{(0,L)} + G_{0,0}^{(0,L)} - G_{0,-1}^{(0,L)} - \left[G_{-1,0}^{(0,L)} - G_{0,0}^{(0,L)} \right] [G_{0,0}^{(0,L)}]^{-1} \left[G_{0,-1}^{(0,L)} - G_{0,0}^{(0,L)} \right] \quad (3.140)$$

Now we need to go to continuum limit, for this we identify $G_{n,n'}^{(0,L)}$ with $G^{(0,L)}(x, x')$ via

$$G(x, x') = \lim_{a \rightarrow 0} \frac{1}{a} G_{n,n'},$$

and use the Taylor approximation:

$$G_{-1,0}^{(0,L)} = a G^{(0,L)}(0 - a, 0) \approx a \left(G^{(0,L)}(0, 0) + G^{(0,L)'}(0, 0) \cdot (-a) + \frac{G^{(0,L)''}(0, 0)}{2} \cdot (-a)^2 \right)$$

and, as before, use the subindices 1 and 2 to indicate partial derivatives of the latter function with respect to the corresponding arguments. We get:

$$G_{-1,-1}^L \approx a^2 G_1^{(0,L)}(0^-, 0) - \frac{a^3}{2} G_{11}^{(0,L)}(0^-, 0) + a^2 G_2^{(0,L)}(0, 0^-) - \frac{a^3}{2} G_{22}^{(0,L)}(0, 0^-) - a^3 G_1^{(0,L)}(0^-, 0) [G^{(0,L)}(0, 0)]^{-1} G_2^{(0,L)}(0, 0^-). \quad (3.141)$$

Observing that $t_m = \frac{\mathcal{A}_m - ia\mathcal{B}_m}{2a^2}$ and (???? понятно, что тут мы взяли выражение выше, убрали $\sim a^2$, но почему? и 2 откуда-то появилась?)

$$\lim_{x \rightarrow 0^-} \frac{d^2}{dx^2} G^L(x, x) = -G_{11}^{(0,L)}(0^-, 0) - G_{22}^{(0,L)}(0, 0^-) - 2G_1^{(0,L)}(0^-, 0) [G^{(0,L)}(0, 0)]^{-1} G_2^{(0,L)}(0, 0^-), \quad (3.142)$$

and using the jump conditions

$$[G_2^{(0,m)}(0, 0^+) - G_2^{(0,m)}(0, 0^-)] \frac{\mathcal{A}_m}{2} = 1, \\ \frac{\mathcal{A}_m}{2} [G_1^{(0,m)}(0^+, 0) - G_1^{(0,m)}(0^-, 0)] = 1,$$

we obtain

$$t_L G_{-1,-1}^L t_L^\dagger = \frac{\mathcal{A}_L - ia\mathcal{B}_L}{2a^2} G_{-1,-1}^L \frac{\mathcal{A}_L + ia\mathcal{B}_L}{2a^2} = -\frac{\mathcal{A}_L}{2a^2} + \frac{1}{8a} \mathcal{A}_L \lim_{x \rightarrow 0^-} \frac{d^2}{dx^2} G^L(x, x) \mathcal{A}_L + O(a^0). \quad (3.143)$$

(???)

Analogously we find that

$$t_R^\dagger G_{1,1}^R t_R = -\frac{\mathcal{A}_R}{2a^2} + \frac{1}{8a} \mathcal{A}_R \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G^R(x, x) \mathcal{A}_R + O(a^0). \quad (3.144)$$

Collecting together the terms defining D in $D := z - W_0 - t_L G_{-1,-1}^L t_L^\dagger - t_R^\dagger G_{1,1}^R t_R$ we observe the cancellation of the $O(\frac{1}{a^2})$ terms thanks to the choice of the a -scaling in $W_0 = \frac{\mathcal{A}_L + \mathcal{A}_R}{2a^2} + \frac{\mathcal{U}_1}{a}$. The remaining $O(\frac{1}{a})$ terms give rise to

$$d := \lim_{a \rightarrow 0} a D = -\mathcal{U}_1 - \frac{1}{8} \mathcal{A}_L \lim_{x \rightarrow 0^-} \frac{d^2}{dx^2} G^L(x, x) \mathcal{A}_L - \frac{1}{8} \mathcal{A}_R \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G^R(x, x) \mathcal{A}_R$$

Note that the linear z -dependence is suppressed in this limit since it appears in the next-to-subleading order of the a -expansion.

To approximate all the remaining terms in $G_{n,n'} = G_{n,n'}^L + G_{n,n'}^R + F_n D^{-1} \bar{F}_{n'}$, it suffices to expand them up to the leading order which turns out to be $O(a^2)$:

(???? there is a typo below, since now $G_{n,-1}^L$ cancels with $G_{n,-1}^L$ etc.)

$$G_{n,-1}^L = G_{n,-1}^L - G_{n,0}^L \approx \quad (3.145)$$

$$\approx -a^2 G_2^{(0,L)}(x, 0^+) + a^2 G^{(0,L)}(x, 0) [G^{(0,L)}(0, 0)]^{-1} G_2^{(0,L)}(0, 0^-) = \quad (3.146)$$

$$= -a^2 [G_2^{(0,L)}(x, 0^+) + G_2^L(x, 0^-) - G_2^{(L,0)}(x, 0^-)], \quad (3.147)$$

$$(3.148)$$

$$G_{n,1}^R = G_{n,1}^R - G_{n,0}^R \approx \quad (3.149)$$

$$\approx a^2 G_2^{(0,R)}(x, 0^-) - a^2 G^{(0,R)}(x, 0) [G^{(0,R)}(0, 0)]^{-1} G_2^{(0,R)}(0, 0^+) =$$

$$= a^2 [G_2^{(0,R)}(x, 0^-) + G_2^R(x, 0^+) - G_2^{(0,R)}(x, 0^+)], \quad (3.150)$$

$$G_{-1,n'}^L = G_{-1,n'}^L - G_{0,n'}^L \approx \quad (3.151)$$

$$\begin{aligned} &\approx -a^2 G_1^{(0,L)}(0^+, x') + a^2 G_1^{(0,L)}(0^-, 0) [G^{(0,L)}(0, 0)]^{-1} G^{(0,L)}(0, x') \\ &= -a^2 \left[G_1^{(0,L)}(0^+, x') + G_1^L(0^-, x') - G_1^{(0,L)}(0^-, x') \right], \end{aligned} \quad (3.152)$$

$$(3.153)$$

$$G_{1,n'}^R = G_{1,n'}^R - G_{0,n'}^R \approx \quad (3.154)$$

$$\begin{aligned} &\approx a^2 G_1^{(0,R)}(0^-, x') - a^2 G_1^{(0,R)}(0^+, 0) [G^{(0,R)}(0, 0)]^{-1} G^{(0,R)}(0, x') \\ &= a^2 \left[G_1^{(0,R)}(0^-, x') + G_1^R(0^+, x') - G_1^{(0,R)}(0^+, x') \right]. \end{aligned} \quad (3.155)$$

Multiplying them with $t_L^\dagger$, t_R , t_L , $t_R^\dagger$, respectively, cancels the factor a^2 , and we eventually obtain

$$\begin{aligned} F(x) = \lim_{a \rightarrow 0} F_n = & -\Theta(-x) \left[G_2^{(0,L)}(x, 0^+) + G_2^L(x, 0^-) - G_2^{(0,L)}(x, 0^-) \right] \frac{\mathcal{A}_L}{2} \\ & + \Theta(x) \left[G_2^{(0,R)}(x, 0^-) + G_2^R(x, 0^+) - G_2^{(0,R)}(x, 0^+) \right] \frac{\mathcal{A}_R}{2}, \end{aligned} \quad (3.156)$$

$$\begin{aligned} \bar{F}(x') = \lim_{a \rightarrow 0} \bar{F}_{n'} = & -\Theta(-x') \frac{\mathcal{A}_L}{2} \left[G_1^{(0,L)}(0^+, x') + G_1^L(0^-, x') - G_1^{(0,L)}(0^-, x') \right] \\ & + \Theta(x') \frac{\mathcal{A}_R}{2} \left[G_1^{(0,R)}(0^-, x') + G_1^R(0^+, x') - G_1^{(0,R)}(0^+, x') \right]. \end{aligned} \quad (3.157)$$

Proof that $d = -\mathcal{U}_1 - \frac{\mathcal{A}_R}{2} F'(0^+) + \frac{\mathcal{A}_L}{2} F'(0^-) + i \frac{\mathcal{B}_R - \mathcal{B}_L}{2}$

To prove the formula for d it suffices to show that

$$\mathcal{A}_R F'(0^+) - iB_R = \frac{1}{4} \mathcal{A}_R \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G^R(x, x) \mathcal{A}_R, \quad (3.158)$$

$$\mathcal{A}_L F'(0^-) - iB_L = -\frac{1}{4} \mathcal{A}_L \lim_{x \rightarrow 0^-} \frac{d^2}{dx^2} G^L(x, x) \mathcal{A}_L. \quad (3.159)$$

First we evaluate

$$F'(0^+) = \left[G_{12}^{(0,R)}(0^+, 0^-) - G_1^{(0,R)}(0^+, 0) [G^{(0,R)}(0, 0)]^{-1} G_2^{(0,R)}(0, 0^+) \right] \frac{\mathcal{A}_R}{2}, \quad (3.160)$$

$$F'(0^-) = -\left[G_{12}^{(0,L)}(0^-, 0^+) - G_1^{(0,L)}(0^-, 0) [G^{(0,L)}(0, 0)]^{-1} G_2^{(0,L)}(0, 0^-) \right] \frac{\mathcal{A}_L}{2}. \quad (3.161)$$

Using the identities

$$G_{12}^{(0,R)}(0^+, 0^-) = G_1^{(0,R)}(0^+, 0) [G^{(0,R)}(0, 0)]^{-1} G_2^{(0,R)}(0, 0^-), \quad (3.162)$$

$$G_{12}^{(0,L)}(0^-, 0^+) = G_1^{(0,L)}(0^-, 0) [G^{(0,L)}(0, 0)]^{-1} G_2^{(0,L)}(0, 0^+), \quad (3.163)$$

previously derived in Ref. [41] (?? somehow) [see Eqs. (B2) and (B3) therein], as well as the jump conditions $[G_2^{(0,m)}(0, 0^+) - G_2^{(0,m)}(0, 0^-)] \frac{\mathcal{A}_m}{2} = 1$ and $\frac{\mathcal{A}_m}{2} [G_1^{(0,m)}(0^+, 0) - G_1^{(0,m)}(0^-, 0)] = 1$, we simplify

$$F'(0^+) = -G_1^{(0,R)}(0^+, 0) [G^{(0,R)}(0, 0)]^{-1}, \quad (3.164)$$

$$F'(0^-) = -G_1^{(0,L)}(0^-, 0) [G^{(0,L)}(0, 0)]^{-1}. \quad (3.165)$$

Next, we evaluate

$$-\frac{1}{8}\mathcal{A}_R \left[\lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G^R(x, x) - \lim_{x \rightarrow 0^-} \frac{d^2}{dx^2} G_0^R(x, x) \right] \mathcal{A}_R \quad (3.166)$$

$$= \frac{1}{4}\mathcal{A}_R G_1^{(0,R)}(0^+, 0)[G^{(0,R)}(0, 0)]^{-1} G_2^{(0,R)}(0, 0^+)\mathcal{A}_R - \frac{1}{4}\mathcal{A}_R G_1^{(0,R)}(0^-, 0)[G^{(0,R)}(0, 0)]^{-1} G_2^{(0,R)}(0, 0^+)\mathcal{A}_R \quad (3.167)$$

$$= \frac{1}{4}\mathcal{A}_R [G_{12}^{(0,R)}(0^+, 0^-) - G_{12}^{(0,R)}(0^-, 0^+)] \mathcal{A}_R - \mathcal{A}_R F'(0^+) - [G^{(0,R)}(0, 0)]^{-1} \quad (3.168)$$

and

$$-\frac{1}{8}\mathcal{A}_L \left[\lim_{x \rightarrow 0^-} \frac{d^2}{dx^2} G^L(x, x) - \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G_0^L(x, x) \right] \mathcal{A}_L \quad (3.169)$$

$$= \frac{1}{4}\mathcal{A}_L G_1^{(0,L)}(0^-, 0)[G^{(0,L)}(0, 0)]^{-1} G_2^{(0,L)}(0, 0^+)\mathcal{A}_L - \frac{1}{4}\mathcal{A}_L G_1^{(0,L)}(0^+, 0)[G^{(0,L)}(0, 0)]^{-1} G_2^{(0,L)}(0, 0^+)\mathcal{A}_L \quad (3.170)$$

$$= \frac{1}{4}\mathcal{A}_L [G_{12}^{(0,L)}(0^-, 0^+) - G_{12}^{(0,L)}(0^+, 0^-)] \mathcal{A}_L + \mathcal{A}_L F'(0^-) - [G^{(0,L)}(0, 0)]^{-1}. \quad (3.171)$$

Further using $[G^{(0,L)}(0)]^{-1} = -\frac{1}{8}\mathcal{A}_L \left[\lim_{x \rightarrow 0^-} \frac{d^2}{dx^2} G^L(x, x) + \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G_0^L(x, x) \right] \mathcal{A}_L$ we establish that

$$\mathcal{A}_R F'(0^+) - \frac{1}{4}\mathcal{A}_R \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G^R(x, x) \mathcal{A}_R = \frac{1}{4}\mathcal{A}_R [G_{12}^{(0,R)}(0^+, 0^-) - G_{12}^{(0,R)}(0^-, 0^+)] \mathcal{A}_R, \quad (3.172)$$

$$\mathcal{A}_L F'(0^-) + \frac{1}{4}\mathcal{A}_L \lim_{x \rightarrow 0^-} \frac{d^2}{dx^2} G^L(x, x) \mathcal{A}_L = \frac{1}{4}\mathcal{A}_L [G_{12}^{(0,L)}(0^+, 0^-) - G_{12}^{(0,L)}(0^-, 0^+)] \mathcal{A}_L. \quad (3.173)$$

The relations

$$\mathcal{A}_R F'(0^+) - iB_R = \frac{1}{4}\mathcal{A}_R \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G^R(x, x) \mathcal{A}_R, \quad (3.174)$$

$$\mathcal{A}_L F'(0^-) - iB_L = -\frac{1}{4}\mathcal{A}_L \lim_{x \rightarrow 0^-} \frac{d^2}{dx^2} G^L(x, x) \mathcal{A}_L. \quad (3.175)$$

follow from these by virtue of the jump condition in the mixed derivatives

$$\frac{1}{4}\mathcal{A}_m [G_{12}^{(0,m)}(0^+, 0^-) - G_{12}^{(0,m)}(0^-, 0^+)] \mathcal{A}_m = i\mathcal{B}_m, \quad (3.176)$$

which is a generalization of the equation

$$G_{12}^{(0)}(x^+, x^-) - G_{12}^{(0)}(x^-, x^+) = i4mA(x). \quad (3.177)$$

in Ref. [41].

3.3.2 Easy illustrative examples

Example 1

Example 2

3.4 Main examples

3.4.1 Basic Josephson model by PPS BGf

By a JJ, one commonly understands a weak link between a pair of superconductors. When two BCS condensates (labeled by 1 and 2) are brought together and the tunneling of Cooper pairs between them is then switched on, the combined system finds a new ground state, in which the difference $\varphi = \varphi_1 - \varphi_2$ of the phases φ_1, φ_2 of the corresponding superconducting order parameters $\Delta_{1,2} = |\Delta_{1,2}|e^{i\varphi_{1,2}}$ adjusts itself to a particular value $\varphi = \varphi_{\min}$. In particular, the most common cases are $\varphi_{\min} = 0$ and $\varphi_{\min} = \pi$, defining the so-called 0- and π -junctions. This result implies that the ground state energy of the combined system is a function of the aforementioned phase difference, with a minimum at $\varphi = \varphi_{\min}$. In this regard, it is important to study the spectral flow of Josephson systems with the externally varied phase difference, and this defines the first type of problems for showcasing our formalism. Note that in real experiments, the phase difference is typically varied by closing the system into a circular geometry far away from the tunneling region, and varying the magnetic flux threading the resulting ring. Specifically, in one of the following examples, we shall see how local Coulomb interaction at the contact between the condensates may lead to a crossover between the above-mentioned $\varphi_{\min} = 0$ and $\varphi_{\min} = \pi$ ground states.

As Josephson systems feature, by construction, superconducting components, local charge conservation is violated resulting in the non-zero persistent current (known as the Josephson current (JC)) between the condensates comprising the junction. It turns out[43] that such a current is also a φ -dependent quantity that, quite generally, may be shown to be the φ -derivative of the aforesaid ground state energy. The study of the experimentally measurable JC defines the second problem, which we consider for demonstrating the potential of Green's function formalism.

Hamiltonian for JJ

In what follows we restrict our consideration to spin- $\frac{1}{2}$ s -wave superconductors, although this formalism allows including arbitrary matrix structure and momentum dependence of the order parameters up to $O(p^2)$.

We consider the following second quantized Hamiltonian

$$\mathcal{H} = \int_{-\infty}^{\infty} dx \hat{\psi}^\dagger(x) \left[p \frac{1}{2m(x)} p + \frac{1}{2} \{A(x), p\} + V(x) \right] \hat{\psi}(x) + \frac{1}{2} \int_{-\infty}^{\infty} dx \left[\hat{\psi}^\dagger(x) \hat{\Delta}(x) (\hat{\psi}^\dagger(x))^T + \text{h.c.} \right], \quad (3.178)$$

where $\hat{\psi}(x) = (\hat{\psi}_\uparrow(x), \hat{\psi}_\downarrow(x))^T$ is a two-component spinor, whose spin components $\sigma = \uparrow, \downarrow$ are the field operators obeying the standard fermionic anticommutation relations

$$\{\hat{\psi}_\sigma(x), \hat{\psi}_{\sigma'}(x')\} = 0, \quad (3.179)$$

$$\{\hat{\psi}_\sigma(x), \hat{\psi}_{\sigma'}^\dagger(x')\} = \delta_{\sigma,\sigma'} \delta(x - x'). \quad (3.180)$$

The effective mass $m(x)$ is a piece-wise constant scalar function of x ; $A(x) = A^\dagger(x)$, $V(x) = V^\dagger(x)$ are piece-wise constant 2×2 hermitian matrices, and $\hat{\Delta}(x) = -\hat{\Delta}^T(x) = \Delta(x)i\sigma_y$ is the

antisymmetric 2×2 s -wave paring matrix expressed in terms of the spatially dependent scalar order parameter $\Delta(x)$. We assign the following spatial dependence to these objects:

$$\begin{pmatrix} m(x) \\ A(x) \\ V(x) \\ \Delta(x) \end{pmatrix} = \Theta(-x) \begin{pmatrix} m_L \\ A_L \\ V_L \\ \Delta_L e^{i\varphi_L} \end{pmatrix} + \Theta(W > x > 0) \begin{pmatrix} m_C \\ A_C \\ V_C \\ 0 \end{pmatrix} + \Theta(x - W) \begin{pmatrix} m_R \\ A_R \\ V_R \\ \Delta_R e^{i\varphi_R} \end{pmatrix}.$$

Note that the left ($-\infty < x < 0$) and the right ($+\infty > x > W$) regions host superconductors, while the central – tunneling – region $W > x > 0$ is normal ($|\Delta_C| = 0$). Introducing the labelling of the regions in terms of the index $\lambda = L, C, R$ (note that in the previous section it corresponds to the index m , but here and below we use λ to avoid confusion with the mass notation), we further specify that m_λ and $\Delta_{L,R}$ take positive real values, while the hermitian matrices $A_\lambda = A_\lambda^{(0)} + \sum_{j=x,y,z} A_\lambda^{(j)} \sigma_j \equiv A_\lambda^{(0)} + \vec{A}_\lambda \cdot \vec{\sigma}$ and $V_\lambda = V_\lambda^{(0)} + \sum_{j=x,y,z} V_\lambda^{(j)} \sigma_j \equiv V_\lambda^{(0)} + \vec{V}_\lambda \cdot \vec{\sigma}$ are spanned by the Pauli matrices σ_j and the identity matrix with the real-valued coefficients $A_\lambda^{(0,j)}$ and $V_\lambda^{(0,j)}$.

Introducing the extended Nambu spinor

$$\hat{\Psi}(x) = \begin{pmatrix} \hat{\psi}(x) \\ i\sigma_y(\hat{\psi}^\dagger(x))^T \end{pmatrix}, \quad (3.181)$$

we rewrite the Hamiltonian

$$\mathcal{H} = \int_{-\infty}^{\infty} dx \hat{\psi}^\dagger(x) \left[p \frac{1}{2m(x)} p + \frac{1}{2} \{A(x), p\} + V(x) \right] \hat{\psi}(x) + \frac{1}{2} \int_{-\infty}^{\infty} dx \left[\hat{\psi}^\dagger(x) \hat{\Delta}(x) (\hat{\psi}^\dagger(x))^T + \text{h.c.} \right], \quad (3.182)$$

in the form

$$\mathcal{H} = \frac{1}{2} \int_{-\infty}^{\infty} dx \hat{\Psi}^\dagger(x) H \hat{\Psi}(x), \quad (3.183)$$

$$H = \begin{pmatrix} h^{(0)} & \Delta(x) \\ \Delta^*(x) & -\sigma_y h^{(0)*} \sigma_y \end{pmatrix}, \quad h^{(0)} = p \frac{1}{2m(x)} p + \frac{1}{2} \{A(x), p\} + V(x). \quad (3.184)$$

$$[H] = 4 \times 4$$

As per common practice, we find it convenient to define a new set of Pauli matrices τ_x, τ_y, τ_z , along with an identity τ_0 , acting on the space of particles (upper two components of $\Psi(x)$) and holes (lower two components of $\Psi(x)$). The Hamiltonian now may be written as

$$H = \tau_z \left[p \frac{1}{2m(x)} p + \frac{1}{2} \{ \vec{A}(x) \cdot \vec{\sigma}, p \} + V^{(0)}(x) \right] + \tau_0 \left[\frac{1}{2} \{ A^{(0)}(x), p \} + \vec{V}(x) \cdot \vec{\sigma} \right] + \Delta(x) \tau_+ + \Delta^*(x) \tau_-. \quad (3.185)$$

(I'll write later that here we have a tensor product)

In this decomposition of H , the terms $\tau_z \left[p \frac{1}{2m(x)} p + \frac{1}{2} \{ \vec{A}(x) \cdot \vec{\sigma}, p \} + V^{(0)}(x) \right]$, that is the kinetic energy, the spin-orbit interaction, and the scalar potential, as well as the pairing potential $\Delta(x) \tau_+ + \Delta^*(x) \tau_-$ are even under the standard time-reversal operation $\hat{T} = i\sigma_y K$, with K denoting the complex conjugation. In turn, the terms $\tau_0 \left[\frac{1}{2} \{ A^{(0)}(x), p \} + \vec{V}(x) \cdot \vec{\sigma} \right]$, that is the vector potential and the Zeeman field, are odd under the time-reversal operation. (??? so what?? do we need this info??)

Main formulas, spectral density, current for a general JJ

It is well-known (see e.g. in Ref. [44]) that introducing the extended Nambu representation artificially doubles the Hilbert space assigned to the quantum system. This redundancy of the description has to be removed in the calculation of observable quantities by enforcing a pseudo-reality constraint on the Nambu field operators. Eventually this results in removing the hole-like part of the spectrum residing at negative energies $\omega < 0$. It follows that the excitation spectrum of the system may be directly inferred from the Eqs.

$$\rho(\omega) = \sum_{m=0}^M \rho^m(\omega) - \frac{1}{\pi} \text{Im} \frac{\partial}{\partial \omega} \ln \det d(\omega + i0^+),$$

$$d_{\text{tw bar}} = \begin{pmatrix} G(0,0) & G(0,W) \\ G(W,0) & G(W,W) \end{pmatrix}^{-1} = \begin{pmatrix} G^{(C,0)}(0,0) & G^{(C,0)}(0,W) \\ G^{(C,0)}(W,0) & G^{(C,0)}(W,W) \end{pmatrix}^{-1} - \begin{pmatrix} p_0 & 0 \\ 0 & p_W \end{pmatrix},$$

$$p_0 = \mathcal{U}_1 + \mathcal{L}_L - \mathcal{L}_{L \rightarrow C}, \quad (3.186)$$

$$p_W = \mathcal{U}_2 - \mathcal{L}_R + \mathcal{L}_{R \rightarrow C}, \quad (3.187)$$

$$\begin{aligned} \mathcal{L}_{L \rightarrow C} &= \frac{\mathcal{A}_C}{2} G_1^{(0,C)}(0^-, 0) [G^{(0,C)}(0, 0)]^{-1} + \frac{i}{2} \mathcal{B}_C, \\ \mathcal{L}_{R \rightarrow C} &= \frac{\mathcal{A}_C}{2} G_1^{(0,C)}(0^+, 0) [G^{(0,C)}(0, 0)]^{-1} + \frac{i}{2} \mathcal{B}_C \\ &= \mathcal{L}_{L \rightarrow C} + [G^{(0,C)}(0, 0)]^{-1}. \end{aligned} \quad (3.188)$$

by restricting the spectral range to $\omega > 0$. For the JJ models, we choose the parameters

$$\mathcal{A}_\lambda = \frac{\tau_z}{m_\lambda}, \quad \mathcal{U}_{1,2} = 0. \quad (3.189)$$

Note that the contact potential strengths $\mathcal{U}_{1,2}$ are neglected, since the physical effect of the tunneling barrier in the finite- W setup is accounted for by an appropriate tuning of V_C .

In the short junction limit, that is when the junction width W is much smaller than all other physical length scales in the system, we impose the scaling $V_C = \frac{\mathcal{U}_1}{W}$ on the potential of the central region, while all other terms in H_C are supposed to be of $O(W^0)$. Then $H_C = [\mathcal{U}_1 + O(W)] \delta_W(x)$, with the nascent δ -function $\delta_W(x) = \frac{1}{W} \Theta(W > x > 0)$. In the limit $W \rightarrow 0$, we get the delta-distribution $\delta(x) = \lim_{W \rightarrow 0} \delta_W(x)$ and neglect the $O(W)$ terms accompanying $\mathcal{U}_1$. The resulting model has the single barrier at $x = 0$ with the contact potential $\mathcal{U}_1 \delta(x)$, and the corresponding formulas for the spectral density

$$\rho(\omega) = \rho^L(\omega) + \rho^R(\omega) - \frac{1}{\pi} \text{Im} \frac{\partial}{\partial \omega} \ln \det d(\omega + i0^+),$$

$$d = -\mathcal{U}_1 + \mathcal{L}_R - \mathcal{L}_L$$

$$\mathcal{L}_L = \frac{\mathcal{A}_L}{2} G_1^{(0,L)}(0^-, 0) [G^{(0,L)}(0, 0)]^{-1} + \frac{i}{2} \mathcal{B}_L$$

become applicable.

In the expressions $\rho(\omega) = \rho^L(\omega) + \rho^R(\omega) - \frac{1}{\pi} \text{Im} \frac{\partial}{\partial \omega} \ln \det d(\omega + i0^+)$ and $\rho(\omega) = \sum_{m=0}^M \rho^m(\omega) - \frac{1}{\pi} \text{Im} \frac{\partial}{\partial \omega} \ln \det d(\omega + i0^+)$ for the spectral density $\rho(\omega)$ of the composite system there are terms $\rho^m(\omega)$ expressing the spectral density of the isolated subsystems and the term expressed via d .

This last term represents the correction due to the tunneling between the subsystems. It is the only term containing the dependence on the phase difference φ , which we indicate explicitly:

$$\delta\rho(\omega, \varphi) = -\frac{1}{\pi} \text{Im} \frac{\partial}{\partial \omega} \ln \det d(\omega + i0^+, \varphi). \quad (3.190)$$

The Josephson current may be expressed as the derivative of the Gibbs free energy F with respect to the phase difference across the superconducting leads, $J(\varphi) = \frac{2e}{\hbar} \frac{dF}{d\varphi}$. It is then given in terms of $\delta\rho(\omega, \varphi)$ by

$$J(\varphi) = -\frac{2e}{\hbar\beta} \int_0^\infty d\omega \ln \left(2 \cosh \frac{\beta\omega}{2} \right) \frac{\partial}{\partial \varphi} \delta\rho(\omega, \varphi), \quad (3.191)$$

where $\beta = \frac{1}{k_B T}$ is the inverse temperature. Integrating by parts we reveal an alternative representation

$$J(\varphi) = -\frac{1}{\Phi_0} \text{Im} \int_0^\infty d\omega \, \text{th} \frac{\beta\omega}{2} \frac{\partial}{\partial \varphi} \ln \det d(\omega + i0^+, \varphi), \quad (3.192)$$

where $\Phi_0 = \frac{\hbar}{2e}$ is the superconducting magnetic flux quantum.

With the help of $J(\varphi)$, the identity

$$\frac{\partial}{\partial \varphi} \ln \det d(z, \varphi) = \text{tr} \left\{ [d(z, \varphi)]^{-1} \frac{\partial d(z, \varphi)}{\partial \varphi} \right\} \quad (3.193)$$

and the granted particle-hole symmetry, one can also convert the integral over the real frequency axis into the (quickly convergent) Matsubara sum

$$J(\varphi) = -\frac{\pi}{\Phi_0\beta} \sum_{i\omega_n} \text{tr} \left\{ [d(i\omega_n, \varphi)]^{-1} \frac{\partial d(i\omega_n, \varphi)}{\partial \varphi} \right\} \quad (3.194)$$

$$\xrightarrow{T \rightarrow 0} -\frac{1}{2\Phi_0} \int_{-\infty}^\infty d\omega \, \text{tr} \left\{ [d(i\omega, \varphi)]^{-1} \frac{\partial d(i\omega, \varphi)}{\partial \varphi} \right\}, \quad (3.195)$$

which in the zero-temperature limit goes over into the integral over the imaginary frequency axis. These imaginary frequency representations appear very efficient in numerical calculations of $J(\varphi)$.

Derivation of the current and spectrum of JJ

In the short junction case, in which the tunneling region is described by the contact potential $\propto \delta(x)$, and the results of Section 3.2.2 are then employed.

Let us start by considering the paradigmatic problem of a pair of s -wave superconductors tunnel-coupled with each other across a barrier at $x = 0$. In this case, the extended Nambu representation $\hat{\Psi}(x) = \left(\hat{\psi}(x), i\sigma_y(\hat{\psi}^\dagger(x))^T \right)^T$ is not needed, and the matrix Hamiltonians are expressed in the reduced Nambu basis $\hat{\Psi}(x) = \left(\hat{\psi}_\uparrow(x), \hat{\psi}_\downarrow^\dagger(x) \right)^T$ as

$$H_\lambda = \begin{pmatrix} \frac{p^2}{2m} - \mu & \Delta_\lambda \\ \Delta_\lambda^* & -\frac{p^2}{2m} + \mu \end{pmatrix}, \quad \lambda = R, L, \quad (3.196)$$

where the order parameter is assumed to be homogeneous in magnitude throughout the sample $|\Delta_\lambda| = \Delta_0 > 0$. Choosing the symmetric gauge, we express the phases of Δ_λ via the externally induced phase difference $\varphi = \varphi_R - \varphi_L$ such that

$$\Delta_\lambda = \Delta_0 e^{i\varphi_\lambda}, \quad \varphi_\lambda = \lambda \frac{\varphi}{2}. \quad (3.197)$$

Hereby we conveniently re-labeled right $\lambda = +$ and left $\lambda = -$ subsystems. As is traditionally done, the nature of the barrier is modeled by the contact potential

$$\mathcal{U}(x) = \delta(x) V_0 \tau_z. \quad (3.198)$$

The d -matrix for this model is calculated in Appendix 3.4.1. For $\text{Im } z > 0$ it reads

$$d(z, \varphi) = -V_0 \tau_z + i \frac{k^{(+)} }{m} \tau_z + i \frac{k^{(-)} }{m} \frac{z - \Delta_0 \cos \frac{\varphi}{2} \tau_x}{z \sqrt{1 - (\frac{\Delta_0}{z})^2}}, \quad (3.199)$$

$$k^{(\pm)} := k_F \frac{\sqrt{1 + \frac{z}{\mu} \sqrt{1 - (\frac{\Delta_0}{z})^2}} \mp \sqrt{1 - \frac{z}{\mu} \sqrt{1 - (\frac{\Delta_0}{z})^2}}}{2}, \quad k_F := \sqrt{2m\mu}. \quad (3.200)$$

The general bound state equation $\det d(\omega) = 0$ yields

(!! тут добавлю вывод этого!!!)

$$0 = \frac{m^2}{k_F^2} \det d(\omega + i0^+, \varphi) = \left(\frac{k^{(+)} }{ik_F} + \frac{mV_0}{k_F} \right)^2 + \frac{k^{(-)2}}{k_F^2} \left(1 - \frac{\Delta_0^2}{\Delta_0^2 - \omega^2} \sin^2 \frac{\varphi}{2} \right), \quad (3.201)$$

where $D = 1/[1 + (mV_0/k_F)^2]$ (imagine that it is substituted in the equation by changing the mV_0/k_F term to it) is the transparency of the tunneling barrier. It can be satisfied in the gap $|\omega| < \Delta_0$, where it holds

$$z \sqrt{1 - \left(\frac{\Delta_0}{z} \right)^2} \Big|_{z=\omega+i0^+} = i \sqrt{\Delta_0^2 - \omega^2}, \quad (3.202)$$

and the corresponding replacements are to be made in $k^{(\pm)}$. It follows that equation $0 = \left(\frac{k^{(+)} }{ik_F} + \frac{mV_0}{k_F} \right)^2 + \frac{k^{(-)2}}{k_F^2} \left(1 - \frac{\Delta_0^2}{\Delta_0^2 - \omega^2} \sin^2 \frac{\varphi}{2} \right)$ defines the exact energy-phase relation, valid for all values of μ and D .

Evaluating the JC on the basis of $J(\varphi) = -\frac{1}{\Phi_0} \text{Im} \int_0^\infty d\omega \, \text{th} \frac{\beta\omega}{2} \frac{\partial}{\partial \varphi} \ln \det d(\omega + i0^+, \varphi)$ [additionally multiplying it by the factor 2 to account the two copies of $d(z, \phi)$ needed to reproduce the extended Nambu representation] we note that the sub-gap contribution appears as the residuum

$$J_{\text{ABS}}(\varphi) = \frac{2\pi}{\Phi_0} \text{th} \frac{\beta\omega_A}{2} \frac{\partial_\varphi \det d(\omega_A, \varphi)}{\partial_\omega \det d(\omega_A, \varphi)} = -\frac{2\pi}{\Phi_0} \text{th} \frac{\beta\omega_A(\varphi)}{2} \frac{d\omega_A(\varphi)}{d\varphi} \quad (3.203)$$

at the pole $\omega = \omega_A$, given by the ABS energy found from the equation $0 = \left(\frac{k^{(+)} }{ik_F} + \frac{mV_0}{k_F} \right)^2 + \frac{k^{(-)2}}{k_F^2} \left(1 - \frac{\Delta_0^2}{\Delta_0^2 - \omega^2} \sin^2 \frac{\varphi}{2} \right)$. In turn, the continuum contribution to $J(\varphi)$ equals

$$J_{\text{cont}}(\varphi) = -\frac{\sin \varphi}{\Phi_0} \int_{\Delta_0}^\infty d\omega \, \text{th} \frac{\beta\omega}{2} \frac{\Delta_0^2}{\omega^2 - \Delta_0^2} \text{Im} \frac{k^{(-)2}}{m^2 \det d(\omega + i0^+, \varphi)}. \quad (3.204)$$

To obtain the known ABS expression

$$\omega_A(\varphi) = \Delta_0 \sqrt{1 - D \sin^2 \frac{\varphi}{2}}, \quad (3.205)$$

we invoke the so-called Andreev approximation relying on $\mu \gg \Delta_0$. In this limit we find that $k^{(+)} \rightarrow 0$ and $k^{(-)} \rightarrow k_F$, simplifying

$$d(z, \varphi) \approx -V_0 \tau_z + \frac{ik_F}{m} \frac{z - \Delta_0 \cos \frac{\varphi}{2} \tau_x}{z \sqrt{1 - (\frac{\Delta_0}{z})^2}}, \quad (3.206)$$

as well as the equation

$$0 = \frac{m^2}{k_F^2} \det d(\omega + i0^+, \varphi) = \left(\frac{k^{(+)} }{ik_F} + \frac{mV_0}{k_F} \right)^2 + \frac{k^{(-)2}}{k_F^2} \left(1 - \frac{\Delta_0^2}{\Delta_0^2 - \omega^2} \sin^2 \frac{\varphi}{2} \right)$$

to the form which admits the solution $\omega_A(\varphi) = \Delta_0 \sqrt{1 - D \sin^2 \frac{\varphi}{2}}$. In addition, we notice that the JC is completely mediated by the ABS, since $k^{(-)}$ and $\det d(\omega + i0^+)$ in $J_{\text{cont}}(\varphi)$ become purely real, and therefore $J_{\text{cont}}(\varphi)$ vanishes in the Andreev limit. Thus, Eq. $J_{\text{ABS}}(\varphi) = -\frac{2\pi}{\Phi_0} \text{th} \frac{\beta \omega_A(\varphi)}{2} \frac{d\omega_A(\varphi)}{d\varphi}$ yields the following universal relation[45, 14]

$$J(\varphi) = \frac{\pi \Delta_0}{2\Phi_0} \frac{D \text{th} \left[\frac{\beta \Delta_0}{2} \sqrt{1 - D \sin^2 \frac{\varphi}{2}} \right]}{\sqrt{1 - D \sin^2 \frac{\varphi}{2}}} \sin \varphi. \quad (3.207)$$

In Fig. 21, in solid lines, we show how the energy-phase and the corresponding current-phase relations approach the universal results of Eqs. $\omega_A(\varphi) = \Delta_0 \sqrt{1 - D \sin^2 \frac{\varphi}{2}}$, $J(\varphi) = \frac{\pi \Delta_0}{2\Phi_0} \frac{D \text{th} \left[\frac{\beta \Delta_0}{2} \sqrt{1 - D \sin^2 \frac{\varphi}{2}} \right]}{\sqrt{1 - D \sin^2 \frac{\varphi}{2}}} \sin \varphi$ upon an increase in the chemical potential μ .

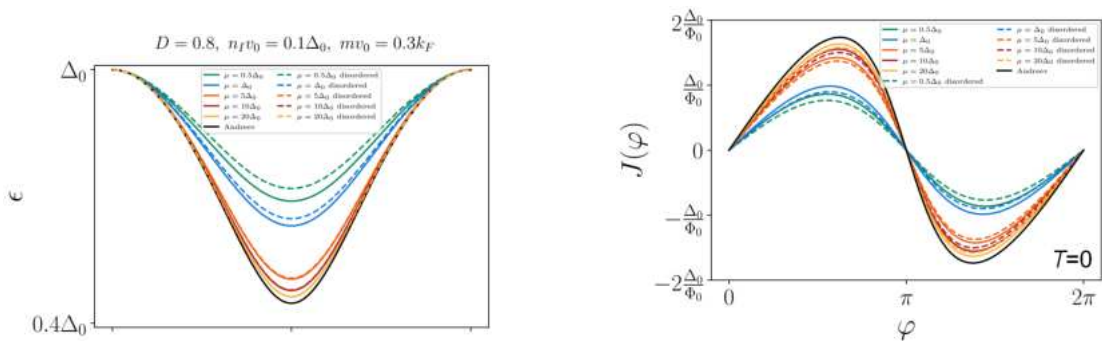


Figure 21: The phase-dispersion of the ABS (top panel) and zero-temperature JC (bottom panel). The results for a clean system are shown in solid lines and are compared to their disordered counterparts plotted in dashed lines. Various colors correspond to different values of the chemical potential, as indicated in the legends. To achieve the Andreev limit universality, we fix the values (see above the top panel) of the barrier transparency D as well as of the energy $n_F v_0$ and momentum $m v_0$ scales associated with the impurity scattering.

It is also worth mentioning that the expression $d(z, \varphi) \approx -V_0\tau_z + \frac{ik_F}{m} \frac{z - \Delta_0 \cos \frac{\varphi}{2} \tau_x}{z \sqrt{1 - (\frac{\Delta_0}{z})^2}}$ multiplied by -1 is analogous by virtue of $G(0, 0) = d^{-1}$ to the self-energy of a quantum dot coupled to superconducting leads in the weak tunneling regime which is further approximated in the wide-band limit, see e.g. Ref. [46] for that self-energy expression (черт его знает, что там авторы имели в виду, какую формулу из той статьи, непонятно это, нужно читать ее, это мб на часа 3 занятие). One has to replace the tunneling rate Γ of that model with $\frac{k_F}{m}$ to gain the formal analogy with the $d(z, \varphi)$. In our treatment, however, an effective Γ is itself of the bandwidth's order of magnitude, that is $\sim \mu$ [cf. the general discussion in Section 3.2.2].

One may further ask whether the JC value beyond the Andreev limit in our model is robust against static disorder, for example. Our approach allows one to get analytical insights into such questions at a modest expense. In particular, we may ignore the effects of disorder on the barrier tunneling, assuming that its imperfection is completely accounted for by the contact potential $\mathcal{U}(x) = \delta(x)V_0\tau_z$. Under such an assumption, we put the self-energy insertions into the bulk propagators $G^{(0,R/L)}$ alone.

Derivation of Position-space Green's functions and d function of bulk s -wave superconductors

The momentum-space Green's functions of the bulk s -wave superconductors are equal at $\varphi = 0$ to

$$G_k^{(0)}(z) = \frac{1}{z^2 - (h_k^{(0)})^2 - \Delta^2} \begin{pmatrix} z + h_k^{(0)} & \Delta_0 \\ \Delta_0 & z - h_k^{(0)} \end{pmatrix}, \quad (3.208)$$

where $h_k^{(0)} = \frac{k^2}{2m} - \mu$.

For the evaluation of

$$G^{(0)}(x, x') = \int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{ik(x-x')} G_k^{(0)}(z) \quad (3.209)$$

we note the following poles of $G_k^{(0)}(z)$

$$k_{1,\pm} = \pm \sqrt{2m\mu + 2mz \sqrt{1 - \left(\frac{\Delta_0}{z}\right)^2}}, \quad k_{2,\pm} = \mp \sqrt{2m\mu - 2mz \sqrt{1 - \left(\frac{\Delta_0}{z}\right)^2}}. \quad (3.210)$$

Considering $\text{Im}(z) > 0$ and deforming appropriately the integration contour in the complex k -plane, we obtain by means of the residue theorem

$$G^{(0)}(x, x') = -\frac{im e^{ik_{1,+}|x-x'|}}{2k_{1,+}} [e^{\tau_x \chi} + \tau_z] + \frac{im e^{ik_{2,+}|x-x'|}}{2k_{2,+}} [e^{\tau_x \chi} - \tau_z], \quad (3.211)$$

where $e^{\tau_x \chi} = \frac{1 + \tau_x \text{th} \chi}{\sqrt{1 - \text{th}^2 \chi}}$ and $\text{th} \chi = \frac{\Delta_0}{z}$.

For $\text{Im}(z) < 0$ it is necessary to replace $k_{1,+} \rightarrow k_{1,-}$ and $k_{2,+} \rightarrow k_{2,-}$ in the above equations. Next, we evaluate

$$G^{(0)}(0, 0) = -\frac{im}{k_{1,+} k_{2,+}} [k^{(+)} \tau_z - k^{(-)} e^{\tau_x \chi}], \quad (3.212)$$

$$[G^{(0)}(0, 0)]^{-1} = -\frac{k^{(+)} \tau_z + k^{(-)} e^{-\tau_x \chi}}{im}, \quad (3.213)$$

where

$$k^{(\pm)} := \frac{k_{1,+} \pm k_{2,+}}{2} \operatorname{sgn} \operatorname{Im}(z), \quad (3.214)$$

$$G_1^{(0)}(0^\pm, 0) = G_2^{(0)}(0, 0^\pm) = \pm m \tau_z. \quad (3.215)$$

With this in hands and noting

$$G^{(0,\lambda)}(x, x') = U_\lambda G^{(0)}(x, x') U_\lambda^\dagger, \quad U_\lambda = e^{\frac{i}{4} \tau_z \lambda \varphi}, \quad (3.216)$$

one can show by

$$d = -\mathcal{U}_1 + \mathcal{L}_R - \mathcal{L}_L, \quad \begin{cases} \mathcal{L}_R = \frac{\mathcal{A}_R}{2} G_1^{(0,R)}(0^+, 0) [G^{(0,R)}(0, 0)]^{-1} + \frac{i}{2} \mathcal{B}_R, \\ \mathcal{L}_L = \frac{\mathcal{A}_L}{2} G_1^{(0,L)}(0^-, 0) [G^{(0,L)}(0, 0)]^{-1} + \frac{i}{2} \mathcal{B}_L, \end{cases}$$

that the matrix d of the Josephson system is given by

$$d(z, \varphi) = -V_0 \tau_z + i \frac{k^{(+)}}{m} \tau_z + i \frac{k^{(-)}}{m} \frac{z - \Delta_0 \cos \frac{\varphi}{2} \tau_x}{z \sqrt{1 - (\frac{\Delta_0}{z})^2}}. \quad (3.217)$$

(!!! add derivation here!!!)

3.4.2 Random non-magnetic impurities in JJ by PPS BGf

Model and results

Let us consider random non-magnetic impurities, characterized by the strength v_0 of the short-range impurity potential and the impurity density n_I . Employing the standard T -matrix approximation [47], we obtain the following disorder-induced self-energy

$$\Sigma^\lambda(z) = U_\lambda \Sigma(z) U_\lambda^\dagger, \quad U_\lambda = e^{\frac{i}{4} \tau_z \lambda \varphi}, \quad (3.218)$$

$$\Sigma(z) = n_I v_0 \tau_z \frac{1}{1 - G^{(0)}(0, 0; z) v_0 \tau_z} + O(n_I^2), \quad (3.219)$$

dressing the momentum-space propagators of the bulk superconductors:

$$\begin{aligned} G_k^{(0,\lambda)}(z) &= U_\lambda G_k^{(0)}(z) U_\lambda^\dagger \quad \rightarrow \quad \tilde{G}_k^{(0,\lambda)}(z) = U_\lambda \tilde{G}_k^{(0)}(z) U_\lambda^\dagger, \\ \tilde{G}_k^{(0)}(z) &= \frac{1}{[G_k^{(0)}(z)]^{-1} - \Sigma(z)}, \end{aligned} \quad (3.220)$$

where $G_k^{(0)}(z) = \frac{1}{z^2 - (h_k^{(0)})^2 - \Delta^2} \begin{pmatrix} z + h_k^{(0)} & \Delta_0 \\ \Delta_0 & z - h_k^{(0)} \end{pmatrix}$. Representing

$$\tilde{G}_k^{(0)}(z) = \frac{1}{\tilde{z}(z) - \tau_z [\frac{k^2}{2m} - \tilde{\mu}(z)] - \tau_x \tilde{\Delta}_0(z)} \quad (3.221)$$

and comparing it with $\tilde{G}_k^{(0)}(z) = \frac{1}{[G_k^{(0)}(z)]^{-1} - \Sigma(z)}$, we find that the effect of the random disorder consists in the following energy-dependent renormalization of the model parameters

$$z \rightarrow \tilde{z}(z) = z \left(1 - n_I v_0 k_{1,+} k_{2,+} \frac{imv_0 k^{(-)}}{(mv_0 k^{(-)})^2 + (k_{1,+} k_{2,+} + imv_0 k^{(+)})^2} \frac{1}{z \sqrt{1 - (\frac{\Delta_0}{z})^2}} \right), \quad (3.222)$$

$$\Delta_0 \rightarrow \tilde{\Delta}_0(z) = \Delta_0 \left(1 - n_I v_0 k_{1,+} k_{2,+} \frac{imv_0 k^{(-)}}{(mv_0 k^{(-)})^2 + (k_{1,+} k_{2,+} + imv_0 k^{(+)})^2} \frac{1}{z \sqrt{1 - (\frac{\Delta_0}{z})^2}} \right), \quad (3.223)$$

$$\mu \rightarrow \tilde{\mu}(z) = \mu - n_I v_0 k_{1,+} k_{2,+} \frac{k_{1,+} k_{2,+} + imv_0 k^{(+)}}{(mv_0 k^{(-)})^2 + (k_{1,+} k_{2,+} + imv_0 k^{(+)})^2}, \quad (3.224)$$

where $k_{1,+}$ and $k_{2,+}$ are

$$k_{1,\pm} = \pm \sqrt{2m\mu + 2mz \sqrt{1 - \left(\frac{\Delta_0}{z}\right)^2}}, \quad k_{2,\pm} = \mp \sqrt{2m\mu - 2mz \sqrt{1 - \left(\frac{\Delta_0}{z}\right)^2}}. \quad (3.225)$$

Since the d -function is eventually expressed via the parameters of the bulk systems, it is sufficient to make the above replacements in

$$d(z, \varphi) = -V_0 \tau_z + i \frac{k^{(+)}}{m} \tau_z + i \frac{k^{(-)}}{m} \frac{z - \Delta_0 \cos \frac{\varphi}{2} \tau_x}{z \sqrt{1 - (\frac{\Delta_0}{z})^2}} \quad (3.226)$$

in order to extract spectral information of the composite system subject to random potential disorder. Using the invariance

$$\frac{z}{\Delta_0} = \frac{\tilde{z}(z)}{\tilde{\Delta}_0(z)}, \quad (3.227)$$

one finds that the dispersion of Andreev levels in Eq.

$$0 = \frac{m^2}{k_F^2} \det d(\omega + i0^+, \varphi) = \left(\frac{k^{(+)}}{ik_F} + \frac{mV_0}{k_F} \right)^2 + \frac{k^{(-)2}}{k_F^2} \left(1 - \frac{\Delta_0^2}{\Delta_0^2 - \omega^2} \sin^2 \frac{\varphi}{2} \right)$$

is modified by the following replacement

$$\frac{k^{(\pm)}}{k_F} \rightarrow \frac{\tilde{k}^{(\pm)}}{k_F} = \frac{1}{2} \sqrt{\frac{\tilde{\mu}(z)}{\mu}} \left(\sqrt{1 + \frac{\tilde{z}(z)}{\tilde{\mu}(z)} \sqrt{1 - (\frac{\Delta_0}{z})^2}} \mp \sqrt{1 - \frac{\tilde{z}(z)}{\tilde{\mu}(z)} \sqrt{1 - (\frac{\Delta_0}{z})^2}} \right)$$

Returning back to the Andreev approximation $\tilde{\mu}(z) \approx \mu \gg \Delta_0$, we find that the results $\omega_A(\varphi) = \Delta_0 \sqrt{1 - D \sin^2 \frac{\varphi}{2}}$ and $J(\varphi) = \frac{\pi \Delta_0}{2\Phi_0} \frac{D \operatorname{th} \left[\frac{\beta \Delta_0}{2} \sqrt{1 - D \sin^2 \frac{\varphi}{2}} \right]}{\sqrt{1 - D \sin^2 \frac{\varphi}{2}}} \sin \varphi$ remain intact.

In Fig. below in dashed lines we show how the energy-phase and the corresponding current-phase relations approach the universal results of Eqs. $\omega_A(\varphi) = \Delta_0 \sqrt{1 - D \sin^2 \frac{\varphi}{2}}$, $J(\varphi) = \frac{\pi \Delta_0}{2\Phi_0} \frac{D \operatorname{th} \left[\frac{\beta \Delta_0}{2} \sqrt{1 - D \sin^2 \frac{\varphi}{2}} \right]}{\sqrt{1 - D \sin^2 \frac{\varphi}{2}}} \sin \varphi$ upon an increase in the chemical potential μ . In addition, Fig. below showcases the effect of disorder on the energy- and current-phase relations away from the universal Andreev limit, showing that static disorder tends to push the Andreev states into

the continuum decreasing the critical current, similar to the effect of decreasing the barrier's transparency.

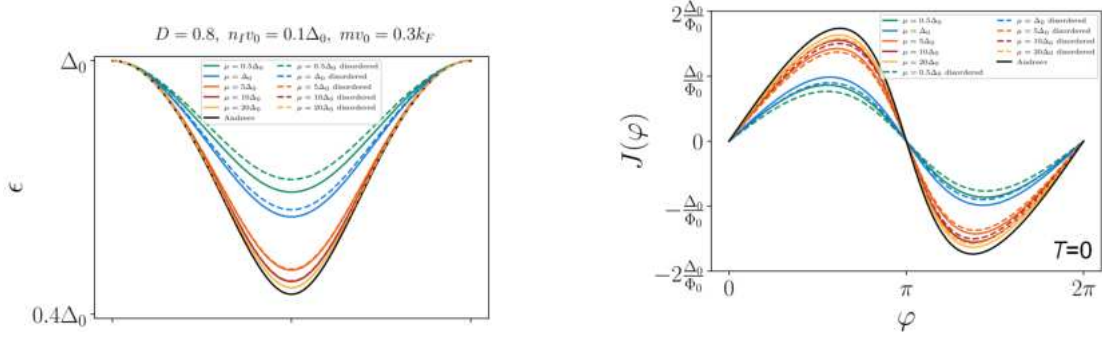


Figure 22: The phase-dispersion of the ABS (top panel) and zero-temperature JC (bottom panel). The results for a clean system are shown in solid lines and are compared to their disordered counterparts plotted in dashed lines. Various colors correspond to different values of the chemical potential, as indicated in the legends. To achieve the Andreev limit universality, we fix the values (see above the top panel) of the barrier transparency D as well as of the energy $n_I v_0$ and momentum $m v_0$ scales associated with the impurity scattering.

3.4.3 The effect of the local Coulomb interaction on JJ by PPS BGf

Model and results

Next, we discuss the effect of the local Coulomb interaction at the point contact of the two superconductors. We model it as the Coulomb repulsion between fluctuations of local spin-up and spin-down densities,

$$V_U = U \left\{ \hat{\psi}_\uparrow^\dagger(0) \hat{\psi}_\uparrow(0) - \langle n_\uparrow(0) \rangle \right\} \left\{ \hat{\psi}_\downarrow^\dagger(0) \hat{\psi}_\downarrow(0) - \langle n_\downarrow(0) \rangle \right\}. \quad (3.228)$$

With this term our model resembles the Anderson-Josephson model of a quantum dot with the on-site Coulomb interaction tunnel-coupled to two superconducting leads (see the review about various studies of this model in Ref. [48]). However, in our case, there is no well-defined quantum dot, since we treat the contact far beyond the tunneling regime.

We also note that the subtraction of the density averages in V_U eventually leads to the energy renormalization in the Cooper channel alone. Performing this subtraction, we get rid of an uninteresting renormalization of the diagonal contact-potential component in the particle-hole basis, which is largely dominated by high-energy normal-state contributions. One can alternatively envisage this subtraction as a result of combining the particle-hole diagonal contribution to the (restricted) Hartree-Fock self-energy with the bare contact potential, which leads to an effective contact potential and defines the contact's physical transparency renormalized by the local Coulomb interaction.

The Hartree-Fock approximation to the local Coulomb interaction V_U is found on the basis of the Wick's theorem

$$V_U \simeq U \langle \hat{\psi}_\downarrow(0) \hat{\psi}_\uparrow(0) \rangle \hat{\psi}_\uparrow^\dagger(0) \hat{\psi}_\downarrow^\dagger(0) + U \langle \hat{\psi}_\uparrow^\dagger(0) \hat{\psi}_\downarrow^\dagger(0) \rangle \hat{\psi}_\downarrow(0) \hat{\psi}_\uparrow(0) - U \langle \hat{\psi}_\uparrow^\dagger(0) \hat{\psi}_\downarrow^\dagger(0) \rangle \langle \hat{\psi}_\downarrow(0) \hat{\psi}_\uparrow(0) \rangle \quad (3.229)$$

$$= \Delta_{\text{loc}} \{ \hat{\psi}_\uparrow^\dagger(0) \hat{\psi}_\downarrow^\dagger(0) + \hat{\psi}_\downarrow(0) \hat{\psi}_\uparrow(0) \} - \frac{\Delta_{\text{loc}}^2}{U}, \quad (3.230)$$

where $\Delta_{\text{loc}} = U \langle \hat{\psi}_\downarrow(0) \hat{\psi}_\uparrow(0) \rangle = U \langle \hat{\psi}_\uparrow^\dagger(0) \hat{\psi}_\downarrow^\dagger(0) \rangle$ is a local superconducting order parameter.

The local Green's function $G_U(0, 0)$ in the Hartree-Fock approximation results from the Dyson equation

$$d_U = [G_U(0, 0)]^{-1} = [G(0, 0)]^{-1} - \Delta_{\text{loc}} \tau_x = d - \Delta_{\text{loc}} \tau_x. \quad (3.231)$$

Using it in the expression $J(\varphi) = -\frac{\pi}{\Phi_0 \beta} \sum_{i\omega_n} \text{tr} \left\{ [d(i\omega_n, \varphi)]^{-1} \frac{\partial d(i\omega_n, \varphi)}{\partial \varphi} \right\}$ [times factor 2 because of the present usage of the reduced Nambu basis] for the JC and accounting for the correction $-\frac{\Delta_{\text{loc}}^2}{U}$ to the free energy occurring in V_U , we establish the mean-field expression for the JC

$$J(\varphi) = -\frac{2\pi}{\Phi_0 \beta} \sum_{i\omega_n} \text{tr} \left\{ [d_U(i\omega_n, \varphi)]^{-1} \frac{\partial d_U(i\omega_n, \varphi)}{\partial \varphi} \right\} - \frac{2\pi}{\Phi_0} \frac{d}{d\varphi} \frac{\Delta_{\text{loc}}^2}{U}. \quad (3.232)$$

With the help of the self-consistency equation (see in Appendix 3.4.3)

$$\Delta_{\text{loc}} = \frac{U}{2\beta} \sum_{i\omega_n} \text{tr} \left\{ [d_U(i\omega_n, \varphi)]^{-1} \tau_x \right\} \quad (3.233)$$

the expression above for $J(\varphi)$ is modified to the form

$$J(\varphi) = -\frac{2\pi}{\Phi_0 \beta} \sum_{i\omega_n} \text{tr} \left\{ [d_U(i\omega_n, \varphi)]^{-1} \frac{\partial d(i\omega_n, \varphi)}{\partial \varphi} \right\}. \quad (3.234)$$

A further application of the self-consistent Hartree-Fock approach to the present model reaches its limitation, emerging in the form of an ultraviolet divergence which is caused by the ultra-local form of the Coulomb interaction (see discussion in Appendix 3.4.3), and one has to refine the method.

A similar problem arises in the bare perturbation theory in the Andreev limit $\mu \gg \Delta_0$: The local superconducting order parameter Δ_{loc} features the logarithmic behavior $\propto \frac{k_F}{m} \ln \frac{\mu}{\Delta_0}$, which hints at the nontrivial competition between the superconductivity and the Kondo effect.[48]

To give a conservative quantitative estimate of the interaction effects on the JC and the Andreev spectrum, we provide below the results of the first-order perturbation theory for $\mu \gtrsim \Delta_0$ and the dimensionless interaction parameter $u = \frac{U m^2 \Delta_0}{\pi k_F^2} \ll 1$. In particular, we approximate the local self-energy term $\Delta_{\text{loc}}(\varphi) = \frac{k_F}{m} \bar{\Delta}_{\text{loc}}(\varphi)$ by its leading $O(u)$ contribution given in

$$\bar{\Delta}_{\text{loc}} \approx -u \int_0^\infty d\bar{\omega} \frac{\bar{k}^{(-)} \cos \frac{\varphi}{2}}{\sqrt{\bar{\omega}^2 + 1}} R(\bar{\omega}) \Big|_{\bar{\Delta}_{\text{loc}}=0},$$

$$R(\bar{\omega}) := \left(\frac{\bar{k}^{(-)2} \bar{\omega}^2}{\bar{\omega}^2 + 1} + \left(\frac{\bar{k}^{(-)} \cos \frac{\varphi}{2}}{\sqrt{\bar{\omega}^2 + 1}} + \bar{\Delta}_{\text{loc}} \right)^2 + \left(\bar{V}_0 - i \bar{k}^{(+)} \right)^2 \right)^{-1}.$$

The resulting d_U is used for the JC evaluation by $J(\varphi) = -\frac{2\pi}{\Phi_0\beta} \sum_{i\omega_n} \text{tr} \left\{ [d_U(i\omega_n, \varphi)]^{-1} \frac{\partial d(i\omega_n, \varphi)}{\partial \varphi} \right\}$ [or $J(\varphi, T=0) = \frac{2\Delta_0 \sin \frac{\varphi}{2}}{\Phi_0} \int_0^\infty d\bar{\omega} \left(\frac{\bar{k}^{(-)} \cos \frac{\varphi}{2}}{\sqrt{\bar{\omega}^2+1}} + \bar{\Delta}_{\text{loc}} \right) \frac{R(\bar{\omega})\bar{k}^{(-)}}{\sqrt{\bar{\omega}^2+1}}$] as well as in the bound state equation

$$\det d_U(\omega, \varphi) = 0. \quad (3.235)$$

Numerical data for the ABS and JC for a junction with $\mu = 3\Delta_0$ are shown in Fig. 23. The panels a) and e) demonstrate the energy-phase relation of the sub-gap states for various values of the interaction parameter $u \in [0, 0.5]$ and contact transparency $D = 0.9$, while the panels b) and f) show the phase-dispersion of the JC at these parameters. We observe that upon increasing u above ~ 0.2 the JC derivative [with respect to φ] changes its sign from positive to negative, indicating the $0 - \pi$ phase transition. It is very analogous to the phase transition which is well-established in superconductor-ferromagnet-superconductor junctions [49], [50]. The panels c) and g) show the JC derivative at $\varphi = 0$ as a function of u and D . In particular, we observe that the phase transition occurs at smaller values of u in more transparent junctions with $D \rightarrow 1$. The panels d) and h) demonstrate the critical current $J_c = \max_\varphi |J(\varphi)|$ as a function of u and D . We reveal that around the phase transition, the critical current displays a non-differentiable cusp in the dependence on the interaction parameter u .

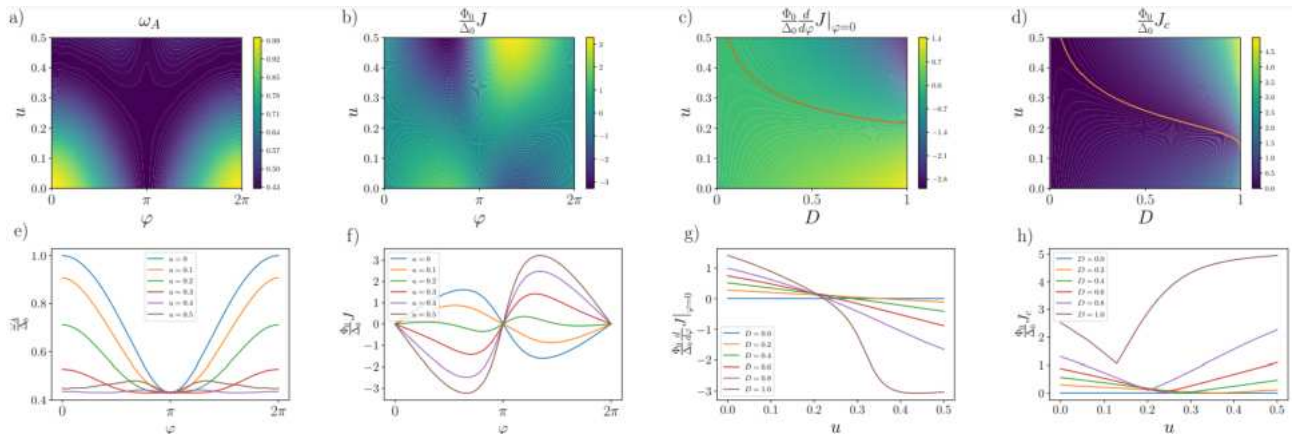


Figure 23: Panels a) and e): Dispersion of sub-gap states as a function of φ and u for a system with $\mu = 3\Delta_0$, $D = 0.9$. Panels b) and f): Josephson current as a function of φ and u for $\mu = 3\Delta_0$, $D = 0.9$. Panels c) and g): derivative of Josephson current with respect to phase difference φ at $\varphi = 0$ for a system with $\mu = 3\Delta_0$. The orange line in panel c) indicates the sign change of $\frac{dJ}{d\varphi}(\varphi = 0)$. Panels d) and h): The critical current as a function of u and D for $\mu = 3\Delta_0$. The orange line in panel d) indicates the location of the cusp in the critical current [explicitly seen in panel h)] as a function of Coulomb interaction strength u .

Derivations of formulas for Local Coulomb interaction in the Hartree-Fock approximation

To compute the expectation values $\langle \psi_\downarrow(0)\psi_\uparrow(0) \rangle$ and $\langle \psi_\uparrow^\dagger(0)\psi_\downarrow^\dagger(0) \rangle$ with respect to the Hartree-Fock Hamiltonian, we define the following imaginary-time Matsubara Green's functions

$$G_{\gamma\delta;U}^{\text{Mat}}(x, x'; \tau) := -\langle T_\tau \hat{\Psi}_\gamma(x, \tau) \hat{\Psi}_\delta^\dagger(x', 0) \rangle = \frac{1}{\beta} \sum_{i\omega_n} e^{-i\omega_n \tau} G_{\gamma\delta;U}(x, x'; i\omega_n), \quad (3.236)$$

where $\hat{\Psi}_1(x) = \hat{\psi}_\uparrow(x)$ and $\hat{\Psi}_2(x) = \hat{\psi}_\downarrow^\dagger(x)$. In this terms

$$\langle \psi_\downarrow(0) \psi_\uparrow(0) \rangle = G_{12,U}^{\text{Mat}}(0, 0; \tau = 0^-) = \frac{1}{2\beta} \sum_{i\omega_n} e^{i\omega_n 0^+} \text{tr} \{ [d_U(i\omega_n)]^{-1} \tau_x \}, \quad (3.237)$$

$$\langle \psi_\uparrow^\dagger(0) \psi_\downarrow^\dagger(0) \rangle = G_{21,U}^{\text{Mat}}(0, 0; \tau = 0^-) = \frac{1}{2\beta} \sum_{i\omega_n} e^{i\omega_n 0^+} \text{tr} \{ [d_U(i\omega_n)]^{-1} \tau_x \}. \quad (3.238)$$

The function $[d_U(z)]^{-1} = [d(z) - \Delta_{\text{loc}} \tau_x]^{-1}$ has the following off-diagonal components

$$\frac{1}{2} \text{tr} \{ [d_U(z)]^{-1} \tau_x \} = \quad (3.239)$$

$$= -\frac{m}{k_F} \left(\frac{\frac{k^{(-)} }{k_F} \frac{i\Delta_0}{z} \cos \frac{\varphi}{2}}{\sqrt{1 - (\frac{\Delta_0}{z})^2}} + \bar{\Delta}_{\text{loc}} \right) \left(\frac{\frac{k^{(-) 2}}{k_F^2}}{1 - (\frac{\Delta_0}{z})^2} + \left(\frac{\frac{k^{(-)} }{k_F} \frac{i\Delta_0}{z} \cos \frac{\varphi}{2}}{\sqrt{1 - (\frac{\Delta_0}{z})^2}} + \bar{\Delta}_{\text{loc}} \right)^2 + \left(\bar{V}_0 + \frac{k^{(+)} }{ik_F} \right)^2 \right)^{-1}, \quad (3.240)$$

where $\bar{V}_0 = \frac{mV_0}{k_F}$ and $\bar{\Delta}_{\text{loc}} = \frac{m\Delta_{\text{loc}}}{k_F}$.

At zero temperature the Matsubara sum in $\langle \psi_\downarrow(0) \psi_\uparrow(0) \rangle$ turns into the imaginary-frequency integral. The integrand $\propto \Delta_0 \cos \frac{\varphi}{2}$ (??? I think it is prop to this in minus 1 degree) in $\frac{1}{2} \text{tr} \{ [d_U(z)]^{-1} \tau_x \}$ decays sufficiently fast at high frequencies (??? because if we substitute z there, for big z the leading term in denominator), so that the convergence factor $e^{i\omega 0^+}$ for it may be omitted. The other integrand $\propto \bar{\Delta}_{\text{loc}}$ does require the convergence factor, which is needed to handle the terms $\sim \frac{1}{z}$ at high frequencies. Collecting all contributions we obtain the following self-consistency equation

$$\bar{\Delta}_{\text{loc}} = -u \int_0^\infty d\bar{\omega} \frac{\bar{k}^{(-)} \cos \frac{\varphi}{2}}{\sqrt{\bar{\omega}^2 + 1}} R(\bar{\omega}) - u \bar{\Delta}_{\text{loc}} \int_0^\infty d\bar{\omega} [R(\bar{\omega}) - R_0(\bar{\omega})] - u \bar{\Delta}_{\text{loc}} \int_0^\infty d\bar{\omega} R_0(\bar{\omega}) \cos(\bar{\omega} + 0^+), \quad (3.241)$$

where u is a dimensionless interaction parameter,

$$u = \frac{Um^2\Delta_0}{\pi k_F^2}, \quad (3.242)$$

$$R(\bar{\omega}) = \left(\frac{\bar{k}^{(-) 2} \bar{\omega}^2}{\bar{\omega}^2 + 1} + \left(\frac{\bar{k}^{(-)} \cos \frac{\varphi}{2}}{\sqrt{\bar{\omega}^2 + 1}} + \bar{\Delta}_{\text{loc}} \right)^2 + (\bar{V}_0 - i\bar{k}^{(+)})^2 \right)^{-1}, \quad (3.243)$$

$$R_0(\bar{\omega}) = \left(\bar{k}_0^{(-) 2} + (\bar{V}_0 - i\bar{k}_0^{(+)})^2 \right)^{-1}, \quad (3.244)$$

$$\bar{k}^{(\pm)} = \frac{1}{2} \left(\sqrt{1 + i\frac{\Delta_0}{\mu} \sqrt{\bar{\omega}^2 + 1}} \mp \sqrt{1 - i\frac{\Delta_0}{\mu} \sqrt{\bar{\omega}^2 + 1}} \right), \quad (3.245)$$

$$\bar{k}_0^{(\pm)} = \frac{1}{2} \left(\sqrt{1 + i\frac{\Delta_0}{\mu} \bar{\omega}} \mp \sqrt{1 - i\frac{\Delta_0}{\mu} \bar{\omega}} \right). \quad (3.246)$$

To perform the integral in $\bar{\Delta}_{\text{loc}}$ we rotate the integration contour to the real axis

$$u \bar{\Delta}_{\text{loc}} i \int_{i0}^{i\infty} d\bar{z} R_0(-i\bar{z}) \frac{e^{\bar{z}+0^+} + e^{-\bar{z}+0^+}}{2} = -\frac{i}{2} u \bar{\Delta}_{\text{loc}} \int_0^\infty d\bar{\omega} [R_0(i\bar{\omega} + 0^+) - R_0(-i\bar{\omega} + 0^+)] \quad (3.247)$$

$$= -u \bar{\Delta}_{\text{loc}} \int_{\mu/\Delta_0}^\infty \frac{d\bar{\omega} e^{-\bar{\omega}+0^+} \sqrt{\frac{\Delta_0}{\mu} \bar{\omega} + 1}}{\left(\sqrt{\frac{\Delta_0}{\mu} \bar{\omega} - 1} + \bar{V}_0 \right) \left(\frac{\Delta_0}{\mu} \bar{\omega} + 1 + \bar{V}_0^2 \right)}. \quad (3.248)$$

The resulting integral has the ultraviolet logarithmic divergence which is not cut off at the scale μ . The origin of this divergence is rooted in the ultra-local form of the Coulomb interaction in our model, and for the integral's regularization it is necessary to introduce the cut-off scale $\omega_c \gg \mu$, whose inverse v_F/ω_c captures the microscopic length scale of the interaction. Thus

$$\bar{\Delta}_{\text{loc}} = -u \int_0^\infty d\bar{\omega} \frac{\bar{k}^{(-)} \cos \frac{\varphi}{2}}{\sqrt{\bar{\omega}^2 + 1}} R(\bar{\omega}) - u \bar{\Delta}_{\text{loc}} \int_0^\infty d\bar{\omega} [R(\bar{\omega}) - R_0(\bar{\omega})] - u \bar{\Delta}_{\text{loc}} \int_0^\infty d\bar{\omega} R_0(\bar{\omega}) \cos(\bar{\omega} + 0^+), \quad (3.249)$$

is estimated by $-u \bar{\Delta}_{\text{loc}} \frac{\mu}{\Delta_0} \ln \frac{\omega_c}{\Delta_0}$. This behaviour signals the necessity to refine the model's consideration, e.g. by resorting to renormalization group methods.[\[48\]](#)

In turn, in the bare perturbation theory the local superconducting order parameter is given by the convergent integral

$$\bar{\Delta}_{\text{loc}} \approx -u \int_0^\infty d\bar{\omega} \frac{\bar{k}^{(-)} \cos \frac{\varphi}{2}}{\sqrt{\bar{\omega}^2 + 1}} R(\bar{\omega}) \Big|_{\bar{\Delta}_{\text{loc}}=0}, \quad (3.250)$$

$$R(\bar{\omega}) := \left(\frac{\bar{k}^{(-)2} \bar{\omega}^2}{\bar{\omega}^2 + 1} + \left(\frac{\bar{k}^{(-)} \cos \frac{\varphi}{2}}{\sqrt{\bar{\omega}^2 + 1}} + \bar{\Delta}_{\text{loc}} \right)^2 + (\bar{V}_0 - i\bar{k}^{(+)})^2 \right)^{-1}. \quad (3.251)$$

It is remarkable that in the Andreev limit $\mu \gg \Delta_0$ it also features the logarithmic dependence of the high energy scale μ :

$$\bar{\Delta}_{\text{loc}} \approx -Du \cos \frac{\varphi}{2} \ln \frac{\mu}{\Delta_0}, \quad D = \frac{1}{1 + \bar{V}_0^2}. \quad (3.252)$$

In addition, we provide the zero temperature limit of the JC formula $J(\varphi) = -\frac{2\pi}{\Phi_0 \beta} \sum_{i\omega_n} \text{tr} \left\{ [d_U(i\omega_n, \varphi)]^{-1} \frac{\partial d(i\omega_n, \varphi)}{\partial \varphi} \right\}$:

$$J(\varphi) = \frac{2\Delta_0 \sin \frac{\varphi}{2}}{\Phi_0} \int_0^\infty d\bar{\omega} \left(\frac{\bar{k}^{(-)} \cos \frac{\varphi}{2}}{\sqrt{\bar{\omega}^2 + 1}} + \bar{\Delta}_{\text{loc}} \right) \frac{R(\bar{\omega}) \bar{k}^{(-)}}{\sqrt{\bar{\omega}^2 + 1}}. \quad (3.253)$$

3.4.4 Junction of two Majorana wires by PPS BGf

Spectral properties of junction of two Majorana wires

Let us now apply our formalism to the famous Majorana wire problem [\[34\],\[35\]](#). In particular, we consider a pair of two semiconducting wires with the strong spin-orbit interaction α and induced superconducting correlations, which are additionally submersed into the external magnetic field B pointing in the wires' direction. In the extended Nambu basis, introduced in [Section 3.4.1](#), they are described by the following Bogoliubov-de-Gennes Hamiltonians

$$H_\lambda = \begin{pmatrix} h_p^{(0)} + \alpha p \sigma_z - B \sigma_x & \Delta_\lambda \\ \Delta_\lambda^* & -h_p^{(0)} - \alpha p \sigma_z - B \sigma_x \end{pmatrix}, \quad (3.254)$$

$$h_p^{(0)} = \frac{p^2}{2m} - \mu; \quad \lambda = R, L \equiv +, -. \quad (3.255)$$

Assuming the wires to be semi-infinite, we bring them in contact[\[32, 33\]](#) at $x = 0$ enabling the tunneling across the contact potential

$$\mathcal{U}(x) = \delta(x) V_0 \tau_z \sigma_0, \quad (3.256)$$

where σ_0 is the identity matrix in the spin space.

As in the previous consideration, we consider again for simplicity the case of the isospectral junction with $|\Delta_\lambda| = \Delta_0 > 0$, with a symmetrically induced phase difference $\Delta_\lambda = \Delta_0 e^{\frac{i}{2}\lambda\varphi}$ across the interface.

The bulk Green's functions

$$G^{(0,\lambda)}(x, x') = U_\lambda g(x - x') U_\lambda^\dagger, \quad U_\lambda = e^{\frac{i}{4}\tau_z \lambda \varphi}, \quad (3.257)$$

associated with the Hamiltonians H_λ , are evaluated in Appendix 3.4.4. This information appears sufficient for establishing the composite Green's function $G(x, x') = G^L(x, x') + G^R(x, x') + F(x)d^{-1}\bar{F}(x')$ and the d -matrix $d = -\mathcal{U}_1 + \mathcal{L}_R - \mathcal{L}_L$ containing spectral properties relevant for the study of the ABS and JC. For the present model we obtain

$$d = \frac{\tau_z}{2m} \left\{ -2mV_0 + U_+ g'(0^+) g^{-1}(0) U_- - U_- g'(0^-) g^{-1}(0) U_+ \right\}. \quad (3.258)$$

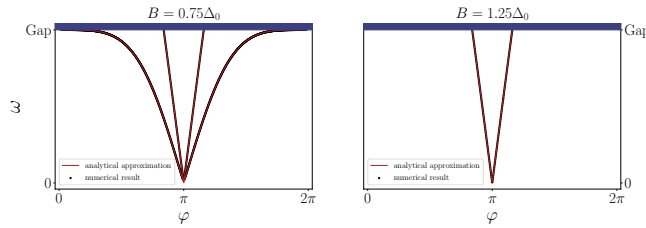


Figure 24: The ABS dispersion in the short junction of the two Majorana wires at zero chemical potential $\mu = 0$ and large spin-orbit energy $E_{\text{SO}} = \frac{m\alpha^2}{2} = 60 \Delta_0$ dominating over Δ_0 and B . Left and right panels refer to the non-topological ($B = 0.75 \Delta_0 < \Delta_0$) and topological ($B = 1.25 \Delta_0 > \Delta_0$) regimes, respectively. As one can see, the numerical result based on the exact equation $\det d = 0$ perfectly agrees with the approximate low-energy result derived in Appendix 3.4.4 [see Eqs. $\cos^2 \frac{\varphi}{2} = \frac{\omega^2 + \Delta_0^2 - B^2 + C_{++}C_{+-}}{\omega^2 + \Delta_0^2 - B^2 - C_{++}C_{+-}}$, $\cos^2 \frac{\varphi}{2} = \frac{\omega^2}{\Delta_0^2}$, which coincides with that of Ref. [33].

The calculation of the JC by means of Eqs.

$$J(\varphi) = -\frac{1}{\Phi_0} \text{Im} \int_0^\infty d\omega \, \text{th} \frac{\beta\omega}{2} \frac{\partial}{\partial \varphi} \ln \det d(\omega + i0^+, \varphi)$$

$$J(\varphi) = -\frac{\pi}{\Phi_0 \beta} \sum_{i\omega_n} \text{tr} \left\{ [d(i\omega_n, \varphi)]^{-1} \frac{\partial d(i\omega_n, \varphi)}{\partial \varphi} \right\} \xrightarrow{T \rightarrow 0} -\frac{1}{2\Phi_0} \int_{-\infty}^\infty d\omega \, \text{tr} \left\{ [d(i\omega, \varphi)]^{-1} \frac{\partial d(i\omega, \varphi)}{\partial \varphi} \right\} \quad (3.259)$$

requires the additional derivative $\partial_\varphi d$. Since the φ -dependence of d in the above formula enters only via the gauging matrices $U_\pm$, we note the following useful relation

$$\partial_\varphi d = \frac{i\tau_z}{8m} \left\{ U_+ [\tau_z, g'(0^+) g^{-1}(0)] U_- + U_- [\tau_z, g'(0^-) g^{-1}(0)] U_+ \right\}. \quad (3.260)$$

On the basis of the exact expressions for $g(0)$ and $g'(0^\pm)$ (see in Appendix 3.4.4) one can address various limits of the model's parameters. For example, we show in Appendix 3.4.4 how the recent results of Ref. [33] for the energy spectrum and the JC in the regime of dominating spin-orbit energy $E_{\text{SO}} \equiv \frac{m\alpha^2}{2} \gg \Delta_0, B, \mu$ can be analytically recovered from our

general expression for the d -function. We also use it to numerically evaluate the ABS dispersion at large $E_{\text{SO}} = 60 \Delta_0$ [see in Fig. 24 to benchmark the present exact interface Green's function approach versus the scattering approach of Ref. [33] relying on low-energy approximations in the spin-orbit dominated regime.

The demonstration presented in Appendix 3.4.4 serves a more general purpose of explaining how one can derive low-energy approximations for $G(x, x')$ in arbitrary heterostructures. Since $G(x, x')$ can be always expressed according to our present findings in terms of bulk Green's functions, it suffices to make a low-energy approximation for these functions. This approximation typically relies on the bulk spectrum linearization near Fermi points (see also Ref. [41] for a similar discussion), and the approximate bulk Green's functions are evaluated much easier than their exact counterparts.

At the same time, it is no longer needed to investigate how the low-energy approximation affects the matching condition $\frac{\mathcal{A}_m \Psi'(x_m^+) - \mathcal{A}_{m-1} \Psi'(x_m^-)}{2} = [\mathcal{U}_m - i \frac{\mathcal{B}_m - \mathcal{B}_{m-1}}{2}] \Psi(x_m)$ (with derivatives of the wavefunctions). This is usually a subtle problem since the spectrum linearization lowers by one the order of the Schrödinger differential equation, and this requires relaxing the condition on the first derivatives. On the other hand, this matching condition expresses the current density conservation, which must be somehow accounted for in the construction of the eigenfunctions. Our approach circumvents this problem, since it allows us to make the low-energy approximation directly for $G(x, x')$, skipping any intermediate approximate treatment of Eq. $\frac{\mathcal{A}_m \Psi'(x_m^+) - \mathcal{A}_{m-1} \Psi'(x_m^-)}{2} = [\mathcal{U}_m - i \frac{\mathcal{B}_m - \mathcal{B}_{m-1}}{2}] \Psi(x_m)$.

The generality of our approach allows one to go beyond the low-energy approximation relying on the dominance of the spin-orbit interaction energy E_{SO} in the present model. Thus it enables exploring arbitrary parameter regimes of the Majorana junction model, as demonstrated in the following.

In particular, in Fig. 25 we display the dispersion of the sub-gap states for the moderate spin-orbit energy $E_{\text{SO}} = 4 \Delta_0$, at zero chemical potential $\mu = 0$. We note that in the non-topological regime $B < \Delta_0$, addressed in the left panel of Fig. 25, Andreev levels are being pushed into the continuum upon an increase of the contact potential strength V_0 , akin to the behavior in ordinary JJs. In the topological regime, shown in the right panel of Fig. 25, we find that the increase of V_0 results in the unwrapping of the Andreev mode, such that its tails near $\varphi = 0, 2\pi$ are pushed towards zero energy.

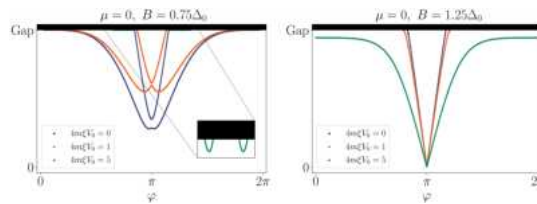


Figure 25: The ABS dispersion in the short Majorana junction at zero chemical potential $\mu = 0$ and moderate spin-orbit interaction strength $E_{\text{SO}} = \frac{m\alpha^2}{2} = 4\Delta_0$. Left and right panels refer to the non-topological ($B < \Delta_0$) and topological ($B > \Delta_0$) regimes, respectively. Blue, orange, and green colors are used to mark three different strengths of the contact potential $4m\xi V_0 = 0$, $4m\xi V_0 = 1$, and $4m\xi V_0 = 5$, respectively, which are expressed in terms of the Ginzburg–Landau coherence length $\xi = 1/\sqrt{2m\Delta_0}$.

The results for a yet different parameter regime with $E_{\text{SO}} = 4\Delta_0$ and – more essentially – nonzero $\mu = \sqrt{3}\Delta_0$ are shown in Fig. 26. As before, the left and right panels refer to the non-topological and topological regimes, respectively. The nonzero μ replaces the boundary value B of the topological phase transition from Δ_0 to the larger value $\sqrt{\Delta_0^2 + \mu^2}$ (in our specific example it is $2\Delta_0$).

Apart from that, we observe the apparent increase in the value of V_0 required to push the Andreev states into and away from the continuum of the scattering states in the non-topological and topological phases, respectively.

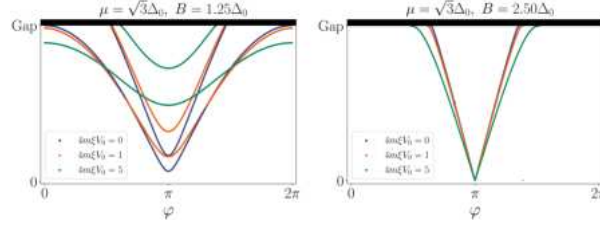


Figure 26: The ABS dispersion in the short Majorana junction at $\mu = \sqrt{3}\Delta_0$ and moderate spin-orbit interaction strength $E_{\text{SO}} = \frac{m\alpha^2}{2} = 4\Delta_0$. Left and right panels refer to the non-topological ($B < \sqrt{\Delta_0^2 + \mu^2}$) and topological ($B > \sqrt{\Delta_0^2 + \mu^2}$) regimes, respectively. Blue, orange, and green colors are used to mark three different strengths of the contact potential $4m\xi V_0 = 0$, $4m\xi V_0 = 1$, and $4m\xi V_0 = 5$, respectively.

Junction of two Majorana wires: Josephson current

An example of the zero-temperature JC computed for the model parameters $E_{\text{SO}} = 4\Delta_0$, $\mu = \sqrt{3}\Delta_0$ is shown in Fig. 27, with the left and right panels referring to the non-topological and topological phases, respectively. In the non-topological phase, the JC demonstrates a sinusoidal behavior with a tilt towards the high-symmetry point $\varphi = \pi$. On the contrary, in the topological phase we see that the JC exhibits a sharp step-like discontinuity at the high-symmetry point $\varphi = \pi$ (see the left panel of Fig. 27). This property relates to the discontinuity in the derivative of the ground state energy (see Fig. 26) with respect to the phase difference $\partial_\varphi E_{\text{GS}}(\varphi) \propto \text{sgn}(\varphi - \pi)$, $\varphi \approx \pi$. Such a result is nonphysical and arises from the non-commutativity of the zero-energy and zero-temperature limits, as is most easily seen in the $J(\varphi) = -\frac{1}{\Phi_0} \text{Im} \int_0^\infty d\omega \, \text{th} \frac{\beta\omega}{2} \frac{\partial}{\partial \varphi} \ln \det d(\omega + i0^+, \varphi)$:

$$\lim_{T \rightarrow 0^+} \lim_{\omega \rightarrow 0^+} \text{th} \left[\frac{\omega}{2k_B T} \right] = 0, \quad (3.261)$$

$$\lim_{\omega \rightarrow 0^+} \lim_{T \rightarrow 0^+} \text{th} \left[\frac{\omega}{2k_B T} \right] = 1. \quad (3.262)$$

This implies that any arbitrarily small but finite temperature will smear the sharp step (see Fig. 28). Additionally, Fig. 27 demonstrates the effect of the barrier imperfection $V_0 \neq 0$ on the JC, quite conventionally implying the reduction of the critical current $J_c = \max_\varphi |J(\varphi)|$ with the increase in the back-scattering strength V_0 .

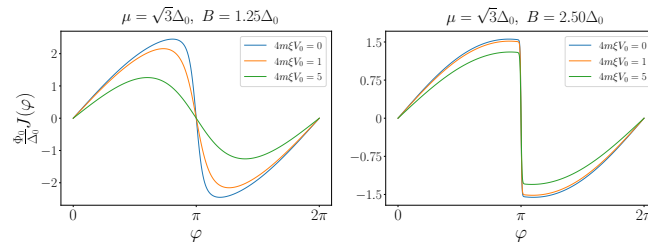


Figure 27: The zero-temperature ($T = 0$) limit of the JC calculated at finite chemical potential $\mu = \sqrt{3}\Delta_0$ and relatively small spin-orbit energy $E_{\text{SO}} = \frac{m\alpha^2}{2} = 4\Delta_0$. Left and right panels refer to the non-topological ($B = 1.25\Delta_0 < \sqrt{\Delta_0^2 + \mu^2} = 2\Delta_0$) and topological phases ($B = 2.50\Delta_0 > 2\Delta_0$), respectively.

The finite-temperature effects are shown in Fig. 28, with the left and right panels referring to the non-topological and topological phases, as before. In the non-topological phase, the key effects of thermodynamic fluctuations are the simultaneous decrease in the critical current and demolition of the aforementioned π -tilt of the current-phase relation. As is mentioned in the previous paragraph, in the topological phase the non-zero temperature destroys the sharp step in the current profile, further affecting the current-phase relation in a manner typical for the non-topological phase.

Double barrier example: Long junction of two Majorana wires

Now we assume that our system consists of the three layers: the right (R) and the left (L) superconducting leads, and the central (C) normal region which is free of superconducting correlations, as is outlined in Section 3.4.1. The Hamiltonians of the right and left superconducting leads and in the central region are given by Eq.

$$H_\lambda = \begin{pmatrix} h_p^{(0)} + \alpha p \sigma_z - B \sigma_x & \Delta_\lambda \\ \Delta_\lambda^* & -h_p^{(0)} - \alpha p \sigma_z - B \sigma_x \end{pmatrix}, \quad \lambda = R, L \equiv +, -;$$

$$H_C = \tau_z h_p^{(0,C)} + \tau_z \alpha_C p \sigma_z - B_C \sigma_x, \quad (3.263)$$

with $h_p^{(0,C)} := h_p^{(0)} = \frac{p^2}{2m} - \mu$, with $\mu \rightarrow \mu_C$ and $m_C = m$.

The bound state equation $\det d(\omega) = 0$, with

$$d_{\text{tw bar}} = \begin{pmatrix} G(0,0) & G(0,W) \\ G(W,0) & G(W,W) \end{pmatrix}^{-1} = \begin{pmatrix} G^{(C,0)}(0,0) & G^{(C,0)}(0,W) \\ G^{(C,0)}(W,0) & G^{(C,0)}(W,W) \end{pmatrix}^{-1} - \begin{pmatrix} p_0 & 0 \\ 0 & p_W \end{pmatrix}$$

can be conveniently written (????) as

$$\det \begin{pmatrix} 1 - \hat{p}_0 g^C(0) & -\hat{p}_0 g^C(-W) e^{\frac{i}{2} \tau_z \varphi} \\ -\hat{p}_W g^C(W) e^{-\frac{i}{2} \tau_z \varphi} & 1 - \hat{p}_W g^C(0) \end{pmatrix} = 0, \quad (3.264)$$

where the translationally invariant Green's function $g^C(x - x')$ for the normal central region is evaluated in Appendix 3.4.4. We also shifted the phase dependence from $p_0 = U_- \hat{p}_0 U_+$ and $p_W = U_+ \hat{p}_W U_-$ to the off-diagonal elements, responsible for the quantum coherence in the normal section, in order to elucidate its importance for supporting the JC. In the present model, the non-decaying terms $g^C(\pm W)$ are achieved due to the gapless spectrum of the Hamiltonian $H_C = \tau_z h_p^{(0,C)} + \tau_z \alpha_C p \sigma_z - B_C \sigma_x$ (in particular, its outer branches do not gap out due to $\Delta_0 = 0$). The matrices

$$\hat{p}_{0,W} = \pm \frac{\tau_z}{2m} g'(0^\mp) [g(0)]^{-1} \mp \frac{\tau_z}{2m} g^C(0^\mp) [g^C(0)]^{-1} \pm \frac{i}{2} (\alpha - \alpha_C) \tau_z \sigma_z \quad (3.265)$$

do not carry the φ -dependence and expressed with the help of the Green's function $g(x - x')$ of the bulk superconductors evaluated in Appendix 3.4.4.

In Fig. 29 we show the phase-dependence of the junction-localised bound states and of the JC at zero and finite temperatures. We focus on the case of the isospectral junction with $B_C = B$, $\mu_C = \mu = \sqrt{3} \Delta_0$, $\alpha_C = \alpha = 4 \Delta_0 \xi$ varying the junction's width $W = 0.5 \xi, 2.5 \xi, 5 \xi, 10 \xi$ [in the units of the coherence length $\xi = 1/\sqrt{2m\Delta_0}$] at the two values of the magnetic field $B = 1.25 \Delta_0, 2.5 \Delta_0$.

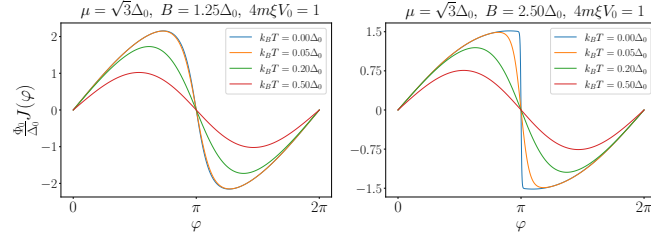


Figure 28: The effect of the finite temperature on the JC, calculated at finite chemical potential $\mu = \sqrt{3}\Delta_0$ and relatively small spin-orbit energy $E_{SO} = \frac{m\alpha^2}{2} = 4\Delta_0$. Left and right panels refer to the non-topological ($B = 1.25\Delta_0 < \sqrt{\Delta_0^2 + \mu^2} = 2\Delta_0$) and topological phases ($B = 2.50\Delta_0 > 2\Delta_0$), respectively.

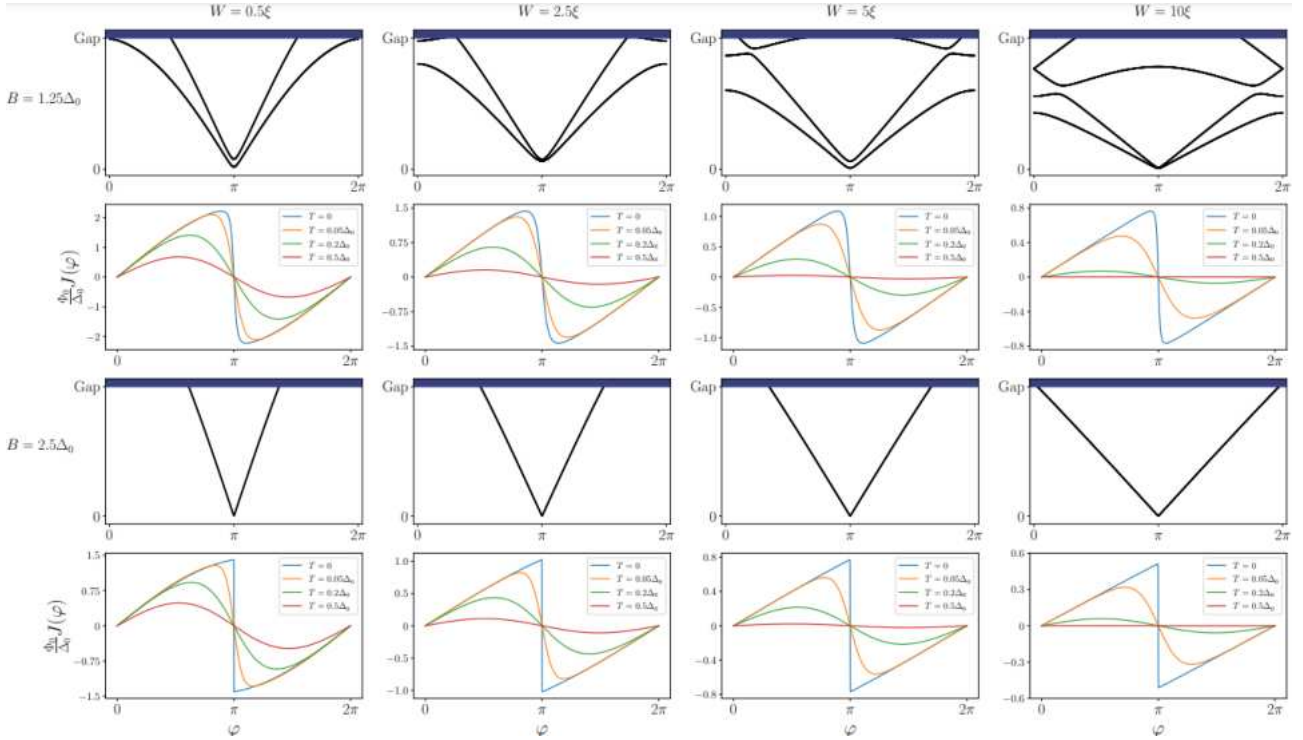


Figure 29: The phase dependence of the ABS and JC for long JJs with $\alpha_C = \alpha = 4\Delta_0\xi$, $\mu_C = \mu = \sqrt{3}\Delta_0$, and $B_C = B$, for a variety of the junction's widths $W = 0.5\xi$, 2.5ξ , 5ξ , and $W = 10\xi$ (four different columns). The two upper rows correspond to the non-topological regime $B = 1.25\Delta_0 < \sqrt{\Delta_0^2 + \mu^2} = 2\Delta_0$, while the two lower rows correspond to the topological phase $B = 2.5\Delta_0 > 2\Delta_0$. In the second and fourth rows, different colors are used to mark different temperatures $T = 0$, $0.05\Delta_0$, $0.2\Delta_0$, and $0.5\Delta_0$.

Specifically, the first and third rows demonstrate the evolution of the energy dispersion with an increase in the width of the central region. We see that in the non-topological phase ($B = 1.25\Delta_0 < \sqrt{\Delta_0^2 + \mu^2} = 2\Delta_0$) the addition of the normal segment between the two Majorana wires has the effect of lowering the energy of the sub-gap states, as well as increasing their number. In turn, in the topological phase ($B = 2.5\Delta_0 > 2\Delta_0$) the only bound state mode tends to spread over the whole phase interval featuring the piece-wise linear branches of its phase dispersion.

As for the JC shown in the second and fourth rows of the same figure, we find that it features qualitatively similar behavior to its short-junction counterpart, with the major effect of the finite width being the decrease in the critical current.

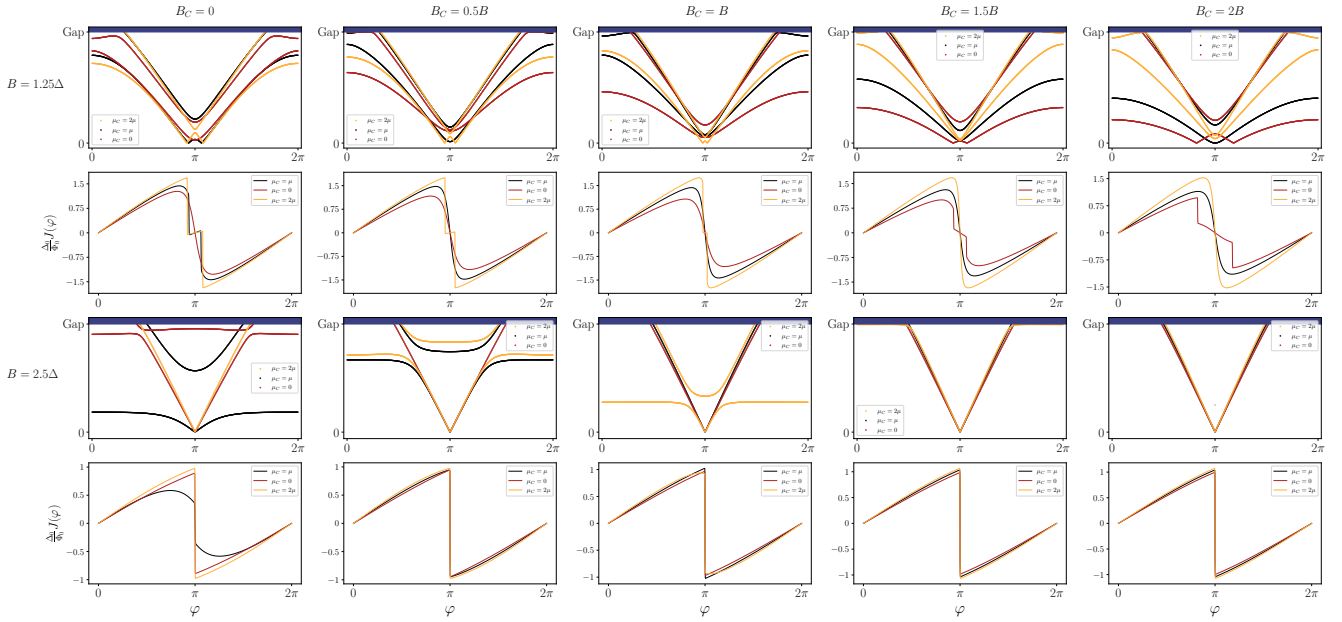


Figure 30: The phase dependence of the ABS and JC for long JJs for the long ($W = 2.5\xi$) JJ with $\alpha_C = \alpha = 4\Delta_0\xi$. The two upper rows correspond to the non-topological phase $B = 1.25\Delta_0$, while the two lower rows describe the topological regime $B = 2.5\Delta_0$. The five different columns correspond to five different values $B_C = 0, 0.5B, B, 1.5B$, and $B_C = 2B$ of the Zeeman field in the central region. Three distinct colors mark three different values $\mu_C = 0, \mu, 2\mu$ of the chemical potential in the central region.

Next, we consider an example of the model's realization, in which some of the parameters of the central region are distinct from the corresponding ones in the superconducting leads. For instance, let us study the effect of modifying the parameters of Abelian (μ_C) and non-Abelian (B_C) parts of the scalar potential. Experimentally this may be achieved by applying an appropriate gate voltage (to affect μ_C) and by bringing the wire in proximity with a ferromagnet (to affect B_C). (???? и где это рассмотрено? нигде?)

In Fig. 30, we present the results for the ABS and the zero temperature JC in the model with $\alpha_C = \alpha = 4\xi\Delta_0$, $W = 2.5\xi$, and $\mu = \sqrt{3}\Delta_0$. The five different columns correspond to five different values of $B_C = 0, 0.5B, B, 1.5B$, and $2B$. Like in Fig. 29, the first and third rows of Fig. 30 show the phase-dispersion of the ABS in the non-topological ($B = 1.25\Delta_0$) and topological ($B = 2.5\Delta_0$) regimes of the Majorana wires, respectively, while the second and fourth rows give the corresponding zero-temperature current-phase relations. Three distinct colors are used to indicate three different values of the chemical potential in the normal region: $\mu_C = 0$ (red), $\mu_C = \mu$ (black), and $\mu_C = 2\mu$ (yellow).

In the non-topological phase ($B = 1.25\Delta_0$, first and second rows), we observe that, by varying the parameters of the central region, it is possible to push the ABS energies downwards as much as needed for creating the crossing points at zero energy. They are similar to the one observed in the topological phase in the short junction. However, these crossing points are not topological in nature, and arise from the crossing of two particle-like and hole-like Andreev bands, that are symmetric around the $\varphi = \pi$ point, and hence come in pairs. Note that one can envisage the situation in which the two bands touch one another at exactly zero energy. In this case, however, the states around zero energy are expected to disperse quadratically, as opposed to the linear dispersion found in the topological regime.

In turn, the crossing at zero energy in the topological phase arises from the intersection

of two $\varphi = \pi$ -asymmetric bands, belonging to two distinct parity branches, and is topologically protected [5]. We note that an analogous effect was also observed in superconductor-ferromagnet-superconductor heterostructures in the presence of spin-orbit coupling in Ref. [6].

As for the zero-temperature JC in the non-topological phase, we reveal that the appearance of the zero-energy-touching Andreev branches results in the discontinuities of the current-phase relations, akin to the topological regime. Like the touching points themselves, the step-like discontinuities also come in pairs and are notably falling behind the critical current in size. The latter effect has to do with the presence of additional strongly dispersing Andreev states, which significantly contribute to the current.

In the topological phase ($B = 2.5\Delta_0$, third and fourth rows), we see that independently of the parameter values inside the central region, there is always only one bound state touching the zero energy at $\varphi = \pi$. In addition, for the cases of the equal ($\mu_C = \mu$, black curve) and enhanced ($\mu_C = 2\mu_C$, yellow curve) chemical potentials of the central segment we observe that an additional bound state appears for $B_C < B$. As in the previous consideration of the topological regime, we find that the current-phase relation again features sharp step-like discontinuity at the high-symmetry point $\varphi = \pi$. We note that for most parameter values the size of this discontinuity is inappreciably smaller than the critical current.

Derivation of the bulk Green's functions for a short Majorana junction

In this Appendix, we evaluate bulk Green's functions and their spatial derivatives which are needed for constructing the d -matrix in the Majorana junction model (both in the short $W \rightarrow 0$ and finite W cases). The derived expressions are used for plotting energy- and current-phase relations in the numerical examples presented in the main text.

The bulk Green's function $g(x)$ occurring in

$$G^{(0,\lambda)}(x, x') = U_\lambda g(x - x') U_\lambda^\dagger, \quad U_\lambda = e^{\frac{i}{4}\tau_z \lambda \varphi} \quad (3.266)$$

is defined by

$$g(x) = \int \frac{dk}{2\pi} e^{ikx} g_k, \quad (3.267)$$

where

$$g_k = \begin{pmatrix} a_k^{(+)} & -\Delta_0 \\ -\Delta_0 & a_k^{(-)} \end{pmatrix}^{-1}, \quad (3.268)$$

$$a_k^{(\pm)} = z \mp h_k^{(0)} \mp \alpha k \sigma_z + B \sigma_x, \quad (3.269)$$

$$h_k^{(0)} = \frac{k^2}{2m} - \mu. \quad (3.270)$$

The inverse in g_K may be written as

$$g_k = \frac{P_k}{Q_k} = \frac{1}{Q_k} \begin{pmatrix} a_k^{(-)} A_k^{(+)} & \Delta_0 A_k^{(-)} \\ \Delta_0 A_k^{(+)} & a_k^{(+)} A_k^{(-)} \end{pmatrix}, \quad (3.271)$$

where

$$\begin{aligned} A_k^{(+)} &= \text{adj} [a_k^{(+)} a_k^{(-)} - \Delta_0^2] \\ &= z^2 - \Delta_0^2 + B^2 - (h_k^{(0)})^2 - \alpha^2 k^2 - 2zB\sigma_x + 2\alpha k h_k^{(0)} \sigma_z + 2iB\alpha k \sigma_y, \end{aligned} \quad (3.272)$$

$$\begin{aligned} A_k^{(-)} &= \text{adj} [a_k^{(+)} a_k^{(-)} - \Delta_0^2] \\ &= z^2 - \Delta_0^2 + B^2 - (h_k^{(0)})^2 - \alpha^2 k^2 - 2zB\sigma_x + 2\alpha k h_k^{(0)} \sigma_z - 2iB\alpha k \sigma_y \end{aligned} \quad (3.273)$$

and

$$\begin{aligned}
Q_k &= \det [a_k^{(+)} a_k^{(-)} - \Delta_0^2] \\
&= \det [a_k^{(-)} a_k^{(+)} - \Delta_0^2] \\
&= [z^2 - \Delta_0^2 + B^2 - (h_k^{(0)})^2 - \alpha^2 k^2]^2 - 4z^2 B^2 - 4\alpha^2 k^2 (h_k^{(0)})^2 + 4\alpha^2 B^2 k^2.
\end{aligned} \tag{3.274}$$

Setting $z = \omega + i0^+$ we find on the basis of

$$\begin{aligned}
G^{(0,m)}(x, x'; \omega + i0^+) &= \sum_{s=1}^{N_c} \frac{i\Theta(x - x') e^{ik_{s,+}|x-x'|}}{\det(\frac{1}{2}\mathcal{A}_m) \prod_{s' \neq s} (k_{s,+} - k_{s',+}) \prod_{s'} (k_{s,+} - k_{s',-})} P(k_{s,+}; \omega) \\
&\quad - \sum_{s=1}^{N_c} \frac{i\Theta(x' - x) e^{-ik_{s,-}|x-x'|}}{\det(\frac{1}{2}\mathcal{A}_m) \prod_{s'} (k_{s,-} - k_{s',+}) \prod_{s' \neq s} (k_{s,-} - k_{s',-})} P(k_{s,-}; \omega).
\end{aligned} \tag{3.275}$$

that

$$\begin{aligned}
g_{\text{Maj fin}}(x) &= i\Theta(x) \sum_{s=1}^4 \frac{(2m)^4 e^{ik_{s,+}x}}{\prod_{s' \neq s} (k_{s,+} - k_{s',+}) \prod_{s'} (k_{s,+} - k_{s',-})} \begin{pmatrix} a_k^{(-)} A_k^{(+)} & \Delta_0 A_k^{(-)} \\ \Delta_0 A_k^{(+)} & a_k^{(+)} A_k^{(-)} \end{pmatrix} \Big|_{k=k_{s,+}} \\
&\quad - i\Theta(-x) \sum_{s=1}^4 \frac{(2m)^4 e^{ik_{s,-}x}}{\prod_{s'} (k_{s,-} - k_{s',+}) \prod_{s' \neq s} (k_{s,-} - k_{s',-})} \begin{pmatrix} a_k^{(-)} A_k^{(+)} & \Delta_0 A_k^{(-)} \\ \Delta_0 A_k^{(+)} & a_k^{(+)} A_k^{(-)} \end{pmatrix} \Big|_{k=k_{s,-}},
\end{aligned} \tag{3.276}$$

where $k_{s,+}$ and $k_{s,-}$ are the roots of the equation $Q_k = 0$ with positive and imaginary parts, respectively.

The expression $g_{\text{Maj fin}}(x)$ allows one to easily establish

$$g'(0^+) = - \sum_{s=1}^4 \frac{(2m)^4 k_{s,+}}{\prod_{s' \neq s} (k_{s,+} - k_{s',+}) \prod_{s'} (k_{s,+} - k_{s',-})} \begin{pmatrix} a_k^{(-)} A_k^{(+)} & \Delta_0 A_k^{(-)} \\ \Delta_0 A_k^{(+)} & a_k^{(+)} A_k^{(-)} \end{pmatrix} \Big|_{k=k_{s,+}}, \tag{3.277}$$

$$g'(0^-) = \sum_{s=1}^4 \frac{(2m)^4 k_{s,-}}{\prod_{s'} (k_{s,-} - k_{s',+}) \prod_{s' \neq s} (k_{s,-} - k_{s',-})} \begin{pmatrix} a_k^{(-)} A_k^{(+)} & \Delta_0 A_k^{(-)} \\ \Delta_0 A_k^{(+)} & a_k^{(+)} A_k^{(-)} \end{pmatrix} \Big|_{k=k_{s,-}}. \tag{3.278}$$

Expansion in the spin-orbit dominated regime

(??? what is the spin-orbit dominate regime???)

To reproduce the results of Ref. [33] in the spin-orbit dominated regime, we perform the expansion of

$$\begin{aligned}
g_{\text{Maj fin}}(x) &= i\Theta(x) \sum_{s=1}^4 \frac{(2m)^4 e^{ik_{s,+}x}}{\prod_{s' \neq s} (k_{s,+} - k_{s',+}) \prod_{s'} (k_{s,+} - k_{s',-})} \begin{pmatrix} a_k^{(-)} A_k^{(+)} & \Delta_0 A_k^{(-)} \\ \Delta_0 A_k^{(+)} & a_k^{(+)} A_k^{(-)} \end{pmatrix} \Big|_{k=k_{s,+}} \\
&\quad - i\Theta(-x) \sum_{s=1}^4 \frac{(2m)^4 e^{ik_{s,-}x}}{\prod_{s'} (k_{s,-} - k_{s',+}) \prod_{s' \neq s} (k_{s,-} - k_{s',-})} \begin{pmatrix} a_k^{(-)} A_k^{(+)} & \Delta_0 A_k^{(-)} \\ \Delta_0 A_k^{(+)} & a_k^{(+)} A_k^{(-)} \end{pmatrix} \Big|_{k=k_{s,-}},
\end{aligned} \tag{3.279}$$

to the leading order $1/\alpha$. This gives

$$\begin{aligned}
g(x) \approx & \frac{e^{iC_0|x|/\alpha} e^{-i2m\alpha x\sigma_z}}{2i\alpha} \left[\frac{z + \Delta_0\tau_x}{C_0} - \text{sgn}(x)\tau_z\sigma_z \right] \\
& + \frac{e^{iC_{++}|x|/\alpha}}{2i\alpha} \left[\frac{z - (B - \Delta_0)\sigma_x}{C_{++}} + \text{sgn}(x)\tau_z\sigma_z \right] \frac{1 + \tau_x\sigma_x}{2} \\
& + \frac{e^{iC_{+-}|x|/\alpha}}{2i\alpha} \left[\frac{z - (B + \Delta_0)\sigma_x}{C_{+-}} + \text{sgn}(x)\tau_z\sigma_z \right] \frac{1 - \tau_x\sigma_x}{2},
\end{aligned} \tag{3.280}$$

where

$$C_0 = z\sqrt{1 - \frac{\Delta_0^2}{z^2}}, \quad C_{\tau\sigma} = z\sqrt{1 - \frac{(\sigma B - \tau\Delta_0)^2}{z^2}}, \quad \tau, \sigma = \pm. \tag{3.281}$$

Next, we establish the quantities relevant for the calculation of the d -matrix:

$$g(0^\pm) = g(0) = \frac{1}{2i\alpha} \sum_{\tau=\pm} \sum_{\sigma=\pm} \bar{g}_{\tau\sigma} \frac{1 + \tau\tau_x}{2} \frac{1 + \sigma\sigma_x}{2}, \tag{3.282}$$

$$\frac{1}{m} g'(0^\pm) = \pm\tau_z - \frac{z + \tau_x\Delta_0}{C_0} \sigma_z, \tag{3.283}$$

where

$$\bar{g}_{\tau\sigma} = \frac{z + \tau\Delta_0}{C_0} + \frac{z - \sigma B + \tau\Delta_0}{C_{\tau\sigma}}. \tag{3.284}$$

Inserting these results at $z = \omega + i0^+$ into

$$d = \frac{\tau_z}{2m} \left\{ -2mV_0 + U_+ g'(0^+) g^{-1}(0) U_- - U_- g'(0^-) g^{-1}(0) U_+ \right\}.$$

after a simple albeit tedious calculation (!?!?) we recover the equations for the sub-gap bound states at the absent contact potential ($V_0 = 0$)

$$\cos^2 \frac{\varphi}{2} = \frac{\omega^2 + \Delta_0^2 - B^2 + C_{++}C_{+-}}{\omega^2 + \Delta_0^2 - B^2 - C_{++}C_{+-}}, \tag{3.285}$$

$$\cos^2 \frac{\varphi}{2} = \frac{\omega^2}{\Delta_0^2}, \tag{3.286}$$

previously reported in Ref. [33].

We note in passing that for an isolated Majorana wire a boundary Green's function of the form discussed in Appendix 3.3.1 suggests the following equation for the bound states

$$\det g(0) = \frac{1}{(2\alpha)^4} \prod_{\tau,\sigma} \bar{g}_{\tau\sigma} = 0, \tag{3.287}$$

which factorizes into the four equations $\bar{g}_{\tau\sigma} = 0$ for $\tau, \sigma = \pm$. The Majorana zero mode $\omega = 0$ appears at $B > \Delta_0$ as a solution satisfying the two equations $\bar{g}_{++} = \bar{g}_{--} = 0$.

Green's function for the Long Majorana junction

In the case of the long junction we, in addition, need the bulk position space Green's function of the normal central region. Setting $\Delta_0 = 0$ in $g_k = \begin{pmatrix} a_k^{(+)} & -\Delta_0 \\ -\Delta_0 & a_k^{(-)} \end{pmatrix}^{-1}$, we represent it as

$$g_k^C = \sum_{\tau=\pm} \frac{\text{adj } a_k^{(\tau)}}{\det a_k^{(\tau)}} \frac{1 + \tau\tau_z}{2} = \sum_{\tau=\pm} g_k^{(\tau)} \frac{1 + \tau\tau_z}{2}, \tag{3.288}$$

where

$$\text{adj} a_k^{(\tau)} = z - \tau h_k^{(0)} + \tau \alpha k \sigma_z - B \sigma_x, \quad (3.289)$$

$$\det a_k^{(\tau)} = [z - \tau h_k^{(0)}]^2 - \alpha^2 k^2 - B^2. \quad (3.290)$$

It follows

$$g^C(x) = \sum_{\tau=\pm} g^{(\tau)}(x) \frac{1 + \tau \tau_z}{2}, \quad g^{(\tau)}(x) = \int \frac{dk}{2\pi} e^{ikx} g_k^{(\tau)}. \quad (3.291)$$

By the virtue of

$$\begin{aligned} G^{(0,m)}(x, x'; \omega + i0^+) &= \sum_{s=1}^{N_c} \frac{i\Theta(x - x') e^{ik_{s,+}|x-x'|}}{\det(\frac{1}{2}\mathcal{A}_m) \prod_{s' \neq s} (k_{s,+} - k_{s',+}) \prod_{s'} (k_{s,+} - k_{s',-})} P(k_{s,+}; \omega) \\ &\quad - \sum_{s=1}^{N_c} \frac{i\Theta(x' - x) e^{-ik_{s,-}|x-x'|}}{\det(\frac{1}{2}\mathcal{A}_m) \prod_{s'} (k_{s,-} - k_{s',+}) \prod_{s' \neq s} (k_{s,-} - k_{s',-})} P(k_{s,-}; \omega). \end{aligned} \quad (3.292)$$

we find for $z = \omega + i0^+$

$$g^{(\tau)}(x) = -\frac{(2m\alpha)^2}{k_{1,+}^2 - k_{2,+}^2} \sum_{s=1}^2 (-1)^{s+1} \frac{e^{ik_{s,+}|x|}}{2i\alpha} \left[\frac{\omega - \tau h_{k_{s,+}}^{(0)} - B\sigma_x}{\alpha k_{s,+}} + \text{sgn}(x) \tau \sigma_z \right], \quad (3.293)$$

where $k_{s,+} = -k_{s,-}$ are the roots of the bi-quadratic equation $\det a_k^{(\tau)} = 0$ with $\text{Im } k_{s,+} > 0$.

We consequently find

$$g^{(\tau)}(0^\pm) = -\frac{(2m\alpha)^2}{k_{1,+}^2 - k_{2,+}^2} \sum_{s=1}^2 \frac{(-1)^{s+1}}{2i\alpha} \frac{\omega - \tau h_{k_{s,+}}^{(0)} - B\sigma_x}{\alpha k_{s,+}}$$

and

$$g^{(\tau)'}(0^\pm) = m\tau \left[\pm 1 - \frac{2m\alpha\sigma_z}{k_{1,+} + k_{2,+}} \right], \quad (3.294)$$

from which the functions in

$$\mathcal{L}_{L \rightarrow C} = \frac{\mathcal{A}_C}{2} G_1^{(0,C)}(0^-, 0) [G^{(0,C)}(0, 0)]^{-1} + \frac{i}{2} \mathcal{B}_C, \quad (3.295)$$

$$\mathcal{L}_{R \rightarrow C} = \frac{\mathcal{A}_C}{2} G_1^{(0,C)}(0^+, 0) [G^{(0,C)}(0, 0)]^{-1} + \frac{i}{2} \mathcal{B}_C \quad (3.296)$$

may be deduced. Hence the d matrix for the model of Section 3.4.4 may be constructed on the basis of Eqs.

$$d_{\text{tw bar}} = \begin{pmatrix} G(0, 0) & G(0, W) \\ G(W, 0) & G(W, W) \end{pmatrix}^{-1} = \begin{pmatrix} G^{(C,0)}(0, 0) & G^{(C,0)}(0, W) \\ G^{(C,0)}(W, 0) & G^{(C,0)}(W, W) \end{pmatrix}^{-1} - \begin{pmatrix} p_0 & 0 \\ 0 & p_W \end{pmatrix},$$

$$p_0 = \mathcal{U}_1 + \mathcal{L}_L - \mathcal{L}_{L \rightarrow C}, \quad (3.297)$$

$$p_W = \mathcal{U}_2 - \mathcal{L}_R + \mathcal{L}_{R \rightarrow C}. \quad (3.298)$$

It is worth noting the useful formula

$$\partial_\varphi d = \frac{i}{4} \begin{pmatrix} [\tau_z, \mathcal{L}_L] & 0 \\ 0 & [\tau_z, \mathcal{L}_R] \end{pmatrix}, \quad (3.299)$$

which is applicable for the JC calculation by means of Eqs. (??? see no connection btw the stated here eq-s and the above one, so I deleted them)

4 Basic Theory of Multiterminal Junctions

5 Green's functions for a multi-terminal model

5.1 Model's description

Let us introduce a model with consists of the three leads, each of which is described by an isolated Green's function $G_{n_m, n'_m}^m(z)$, with $m = 1, 2, 3$ and $1 \leq n_m, n'_m < \infty$. Note that for each lead m we have a local site labelling $\{n_m\}$. We also introduce three finite segments described by isolated Green's functions $G_{n_{\bar{m}}, n'_{\bar{m}}}^{\bar{m}}(z)$, with $\bar{m} = \bar{1}, \bar{2}, \bar{3}$ and $1 \leq n_{\bar{m}}, n'_{\bar{m}} \leq N_{\bar{m}}$. The local labeling of sites in each segment has an increasing order in the direction of arrows shown in Fig. 31. The finite segments are connected to the semi-infinite leads via the junction sites by means of the nearest-neighbour hoppings with the amplitudes $-t'_m, -t'_{(m, \bar{m})}$ (in the direction of the arrows) and $-t_m^\dagger, -t_{(m, \bar{m})}^\dagger$ (against the direction of arrows). The isolated junction sites are characterized by the on-site (matrix-valued) potentials W^m , and we also assign the site labels $n_m = 0$ to the junction sites, envisaging them as a continuation of the corresponding leads. Thus the Green's function of the isolated junction sites read $g^m(z) = \frac{|0\rangle_m \langle 0|}{z - W^m}$.

Figure 31: Sketch of the model.

5.2 Green's function

Our goal is to establish the Green's function $G_{n_l, n'_l}^{(l, l')}$ of the full model, where $l = 1, 2, 3, \bar{1}, \bar{2}, \bar{3}$, and the site labeling ranges as $0 \leq n_{1,2,3} < \infty$ and $1 \leq n_{\bar{1}, \bar{2}, \bar{3}} \leq N_{\bar{1}, \bar{2}, \bar{3}}$. We set up the Dyson equation $G = \bar{G} + \bar{G}VG$, where $\bar{G}$ is a collection of the Green's function of the isolated subsystems, and V is a coupling term (the above described hoppings):

$$\begin{aligned} V = & -|1\rangle_1 (t'_1)_1 \langle 0| - |N_{\bar{2}}\rangle_{\bar{2}} (t'_{(1, \bar{2})})_1 \langle 0| - |1\rangle_{\bar{3}} (t'_{(1, \bar{3})})_1 \langle 0| + h.c. \\ & -|1\rangle_2 (t'_2)_2 \langle 0| - |N_{\bar{3}}\rangle_{\bar{3}} (t'_{(2, \bar{3})})_2 \langle 0| - |1\rangle_{\bar{1}} (t'_{(2, \bar{1})})_2 \langle 0| + h.c. \\ & -|1\rangle_3 (t'_3)_3 \langle 0| - |N_{\bar{1}}\rangle_{\bar{1}} (t'_{(3, \bar{1})})_3 \langle 0| - |1\rangle_{\bar{2}} (t'_{(3, \bar{2})})_3 \langle 0| + h.c. \end{aligned} \quad (5.1)$$

Component-wise we have the following Dyson equation:

$$\begin{aligned} G_{n_l, n'_l}^{(l, l')} = & \bar{G}_{n_l, n'_l}^{(l, l')} \\ & - \left[\delta_{l,1} G_{n_{1,1}}^1 t'_1 + \delta_{l,\bar{2}} G_{n_{\bar{2}}, N_{\bar{2}}}^{\bar{2}} t'_{(1, \bar{2})}^\dagger + \delta_{l,\bar{3}} G_{n_{\bar{3}}, 1}^{\bar{3}} t'_{(1, \bar{3})} \right] G_{0, n'_l}^{(1, l')} \\ & - \frac{\delta_{l,1} \delta_{n_{1,0}}}{z - W^1} \left[t_1^\dagger G_{1, n'_l}^{(1, l')} + t'_{(1, \bar{2})} G_{N_{\bar{2}}, n'_l}^{(\bar{2}, l')} + t'_{(1, \bar{3})} G_{1, n'_l}^{(\bar{3}, l')} \right] \\ & - \left[\delta_{l,2} G_{n_{2,1}}^2 t'_2 + \delta_{l,\bar{3}} G_{n_{\bar{3}}, N_{\bar{3}}}^{\bar{3}} t'_{(2, \bar{3})}^\dagger + \delta_{l,\bar{1}} G_{n_{\bar{1}}, 1}^{\bar{1}} t'_{(2, \bar{1})} \right] G_{0, n'_l}^{(2, l')} \\ & - \frac{\delta_{l,2} \delta_{n_{2,0}}}{z - W^2} \left[t_2^\dagger G_{1, n'_l}^{(2, l')} + t'_{(2, \bar{3})} G_{N_{\bar{3}}, n'_l}^{(\bar{3}, l')} + t'_{(2, \bar{1})} G_{1, n'_l}^{(\bar{1}, l')} \right] \\ & - \left[\delta_{l,3} G_{n_{3,1}}^3 t'_3 + \delta_{l,\bar{1}} G_{n_{\bar{1}}, N_{\bar{1}}}^{\bar{1}} t'_{(3, \bar{1})}^\dagger + \delta_{l,\bar{2}} G_{n_{\bar{2}}, 1}^{\bar{2}} t'_{(3, \bar{2})} \right] G_{0, n'_l}^{(3, l')} \\ & - \frac{\delta_{l,3} \delta_{n_{3,0}}}{z - W^3} \left[t_3^\dagger G_{1, n'_l}^{(3, l')} + t'_{(3, \bar{1})} G_{N_{\bar{1}}, n'_l}^{(\bar{1}, l')} + t'_{(3, \bar{2})} G_{1, n'_l}^{(\bar{2}, l')} \right]. \end{aligned} \quad (5.2)$$

There are 12 unknown quantities in the rhs: 1) $G_{0, n'_l}^{(1, l')}, G_{0, n'_l}^{(2, l')}, G_{0, n'_l}^{(3, l')}$; 2) $G_{1, n'_l}^{(1, l')}, G_{N_{\bar{2}}, n'_l}^{(\bar{2}, l')}, G_{1, n'_l}^{(\bar{3}, l')}$; 3) $G_{1, n'_l}^{(2, l')}, G_{N_{\bar{3}}, n'_l}^{(\bar{3}, l')}, G_{1, n'_l}^{(\bar{1}, l')}$; 4) $G_{1, n'_l}^{(3, l')}, G_{N_{\bar{1}}, n'_l}^{(\bar{1}, l')}, G_{1, n'_l}^{(\bar{2}, l')}$.

Let s first set $l = 1$ in the above equation. We obtain

$$\begin{aligned} G_{n_1, n'_1}^{(1, l')} &= \delta_{l', 1} \bar{G}_{n_1, n'_1}^{(1, 1)} - G_{n_1, 1}^1 t'_1 G_{0, n'_1}^{(1, l')} \\ &\quad - \frac{\delta_{n_1, 0}}{z - W^1} \left[t_1'^{\dagger} G_{1, n'_1}^{(1, l')} + t'_{(1, 2)} G_{N_2, n'_1}^{(2, l')} + t'_{(1, 3)}^{\dagger} G_{1, n'_1}^{(3, l')} \right]. \end{aligned} \quad (5.3)$$

Considering this relation at $n_1 = 0$ and $n_1 = 1$, we derive

$$G_{0, n'_1}^{(1, l')} = \frac{1}{z - W^1} \left[\delta_{l', 1} \delta_{n'_1, 0} - t_1'^{\dagger} G_{1, n'_1}^{(1, l')} - t'_{(1, 2)} G_{N_2, n'_1}^{(2, l')} - t'_{(1, 3)}^{\dagger} G_{1, n'_1}^{(3, l')} \right], \quad (5.4)$$

$$G_{1, n'_1}^{(1, l')} = \delta_{l', 1} G_{1, n'_1}^1 - G_{1, 1}^1 t'_1 G_{0, n'_1}^{(1, l')}. \quad (5.5)$$

Set $l = 2$ and obtain

$$\begin{aligned} G_{n_2, n'_2}^{(2, l')} &= \delta_{l', 2} \bar{G}_{n_2, n'_2}^{(2, 2)} - G_{n_2, 1}^2 t'_2 G_{0, n'_2}^{(2, l')} \\ &\quad - \frac{\delta_{n_2, 0}}{z - W^2} \left[t_2'^{\dagger} G_{1, n'_2}^{(2, l')} + t'_{(2, 3)} G_{N_3, n'_2}^{(3, l')} + t'_{(2, 1)}^{\dagger} G_{1, n'_2}^{(1, l')} \right]. \end{aligned} \quad (5.6)$$

Considering this relation at $n_2 = 0$ and $n_2 = 1$, we derive

$$G_{0, n'_2}^{(2, l')} = \frac{1}{z - W^2} \left[\delta_{l', 2} \delta_{n'_2, 0} - t_2'^{\dagger} G_{1, n'_2}^{(2, l')} - t'_{(2, 3)} G_{N_3, n'_2}^{(3, l')} - t'_{(2, 1)}^{\dagger} G_{1, n'_2}^{(1, l')} \right], \quad (5.7)$$

$$G_{1, n'_2}^{(2, l')} = \delta_{l', 2} G_{1, n'_2}^2 - G_{1, 1}^2 t'_2 G_{0, n'_2}^{(2, l')}. \quad (5.8)$$

Set $l = 3$ and obtain

$$\begin{aligned} G_{n_3, n'_3}^{(3, l')} &= \delta_{l', 3} \bar{G}_{n_3, n'_3}^{(3, 3)} - G_{n_3, 1}^3 t'_3 G_{0, n'_3}^{(3, l')} \\ &\quad - \frac{\delta_{n_3, 0}}{z - W^3} \left[t_3'^{\dagger} G_{1, n'_3}^{(3, l')} + t'_{(3, 1)} G_{N_1, n'_3}^{(1, l')} + t'_{(3, 2)}^{\dagger} G_{1, n'_3}^{(2, l')} \right]. \end{aligned} \quad (5.9)$$

Considering this relation at $n_3 = 0$ and $n_3 = 1$, we derive

$$G_{0, n'_3}^{(3, l')} = \frac{1}{z - W^3} \left[\delta_{l', 3} \delta_{n'_3, 0} - t_3'^{\dagger} G_{1, n'_3}^{(3, l')} - t'_{(3, 1)} G_{N_1, n'_3}^{(1, l')} - t'_{(3, 2)}^{\dagger} G_{1, n'_3}^{(2, l')} \right], \quad (5.10)$$

$$G_{1, n'_3}^{(3, l')} = \delta_{l', 3} G_{1, n'_3}^3 - G_{1, 1}^3 t'_3 G_{0, n'_3}^{(3, l')}. \quad (5.11)$$

Next we set $l = \bar{1}$. This yields

$$G_{n_{\bar{1}}, n'_{\bar{1}}}^{(\bar{1}, l')} = \delta_{l', \bar{1}} \bar{G}_{n_{\bar{1}}, n'_{\bar{1}}}^{\bar{1}} - G_{n_{\bar{1}}, 1}^{\bar{1}} t'_{(2, \bar{1})} G_{0, n'_{\bar{1}}}^{(2, l')} - G_{n_{\bar{1}}, N_{\bar{1}}}^{\bar{1}} t'_{(3, \bar{1})}^{\dagger} G_{0, n'_{\bar{1}}}^{(3, l')}. \quad (5.12)$$

Considering this relation at $n_{\bar{1}} = 1$ and $n_{\bar{1}} = N_{\bar{1}}$ we derive

$$G_{1, n'_{\bar{1}}}^{(\bar{1}, l')} = \delta_{l', \bar{1}} \bar{G}_{1, n'_{\bar{1}}}^{\bar{1}} - G_{1, 1}^{\bar{1}} t'_{(2, \bar{1})} G_{0, n'_{\bar{1}}}^{(2, l')} - G_{1, N_{\bar{1}}}^{\bar{1}} t'_{(3, \bar{1})}^{\dagger} G_{0, n'_{\bar{1}}}^{(3, l')}, \quad (5.13)$$

$$G_{N_{\bar{1}}, n'_{\bar{1}}}^{(\bar{1}, l')} = \delta_{l', \bar{1}} \bar{G}_{N_{\bar{1}}, n'_{\bar{1}}}^{\bar{1}} - G_{N_{\bar{1}}, 1}^{\bar{1}} t'_{(2, \bar{1})} G_{0, n'_{\bar{1}}}^{(2, l')} - G_{N_{\bar{1}}, N_{\bar{1}}}^{\bar{1}} t'_{(3, \bar{1})}^{\dagger} G_{0, n'_{\bar{1}}}^{(3, l')}. \quad (5.14)$$

Setting $l = \bar{2}$ we obtain

$$G_{n_{\bar{2}}, n'_{\bar{2}}}^{(\bar{2}, l')} = \delta_{l', \bar{2}} \bar{G}_{n_{\bar{2}}, n'_{\bar{2}}}^{\bar{2}} - G_{n_{\bar{2}}, N_{\bar{2}}}^{\bar{2}} t'_{(1, \bar{2})}^{\dagger} G_{0, n'_{\bar{2}}}^{(1, l')} - G_{n_{\bar{2}}, 1}^{\bar{2}} t'_{(3, \bar{2})} G_{0, n'_{\bar{2}}}^{(3, l')}. \quad (5.15)$$

Considering this relation at $n_{\bar{2}} = 1$ and $n_{\bar{2}} = N_{\bar{2}}$ we derive

$$G_{1, n'_{\bar{2}}}^{(\bar{2}, l')} = \delta_{l', \bar{2}} \bar{G}_{1, n'_{\bar{2}}}^{\bar{2}} - G_{1, N_{\bar{2}}}^{\bar{2}} t'_{(1, \bar{2})}^{\dagger} G_{0, n'_{\bar{2}}}^{(1, l')} - G_{1, 1}^{\bar{2}} t'_{(3, \bar{2})} G_{0, n'_{\bar{2}}}^{(3, l')}, \quad (5.16)$$

$$G_{N_{\bar{2}}, n'_{\bar{2}}}^{(\bar{2}, l')} = \delta_{l', \bar{2}} \bar{G}_{N_{\bar{2}}, n'_{\bar{2}}}^{\bar{2}} - G_{N_{\bar{2}}, N_{\bar{2}}}^{\bar{2}} t'_{(1, \bar{2})}^{\dagger} G_{0, n'_{\bar{2}}}^{(1, l')} - G_{N_{\bar{2}}, 1}^{\bar{2}} t'_{(3, \bar{2})} G_{0, n'_{\bar{2}}}^{(3, l')}. \quad (5.17)$$

Setting $l = \bar{3}$ we obtain

$$G_{n_{\bar{3}}, n'_{l'}}^{(\bar{3}, l')} = \delta_{l', \bar{3}} G_{n_{\bar{3}}, n'_{\bar{3}}}^{\bar{3}} - G_{n_{\bar{3}}, 1}^{\bar{3}} t'_{(1, \bar{3})} G_{0, n'_{l'}}^{(1, l')} - G_{n_{\bar{3}}, N_{\bar{3}}}^{\bar{3}} t'^{\dagger}_{(2, \bar{3})} G_{0, n'_{l'}}^{(2, l')}. \quad (5.18)$$

Considering this relation at $n_{\bar{3}} = 1$ and $n_{\bar{3}} = N_{\bar{3}}$ we derive

$$G_{1, n'_{l'}}^{(\bar{3}, l')} = \delta_{l', \bar{3}} G_{1, n'_{\bar{3}}}^{\bar{3}} - G_{1, 1}^{\bar{3}} t'_{(1, \bar{3})} G_{0, n'_{l'}}^{(1, l')} - G_{1, N_{\bar{3}}}^{\bar{3}} t'^{\dagger}_{(2, \bar{3})} G_{0, n'_{l'}}^{(2, l')}, \quad (5.19)$$

$$G_{N_{\bar{3}}, n'_{l'}}^{(\bar{3}, l')} = \delta_{l', \bar{3}} G_{N_{\bar{3}}, n'_{\bar{3}}}^{\bar{3}} - G_{N_{\bar{3}}, 1}^{\bar{3}} t'_{(1, \bar{3})} G_{0, n'_{l'}}^{(1, l')} - G_{N_{\bar{3}}, N_{\bar{3}}}^{\bar{3}} t'^{\dagger}_{(2, \bar{3})} G_{0, n'_{l'}}^{(2, l')}. \quad (5.20)$$

Inserting $G_{1, n'_{l'}}^{(1, l')}$, $G_{1, n'_{l'}}^{(2, l')}$, $G_{1, n'_{l'}}^{(3, l')}$ as well as $G_{1, n'_{l'}}^{(\bar{1}, l')}$, $G_{N_{\bar{1}}, n'_{l'}}^{(\bar{1}, l')}$, $G_{1, n'_{l'}}^{(\bar{2}, l')}$, $G_{N_{\bar{2}}, n'_{l'}}^{(\bar{2}, l')}$, $G_{1, n'_{l'}}^{(\bar{3}, l')}$, $G_{N_{\bar{3}}, n'_{l'}}^{(\bar{3}, l')}$ into the equations for $G_{0, n'_{l'}}^{(1, l')}$, $G_{0, n'_{l'}}^{(2, l')}$, $G_{0, n'_{l'}}^{(3, l')}$, we obtain

$$D_{11} G_{0, n'_{l'}}^{(1, l')} + D_{12} G_{0, n'_{l'}}^{(2, l')} + D_{13} G_{0, n'_{l'}}^{(3, l')} = -\bar{F}_{n'_{l'}}^{(1, l')}, \quad (5.21)$$

$$D_{21} G_{0, n'_{l'}}^{(1, l')} + D_{22} G_{0, n'_{l'}}^{(2, l')} + D_{23} G_{0, n'_{l'}}^{(3, l')} = -\bar{F}_{n'_{l'}}^{(2, l')}, \quad (5.22)$$

$$D_{31} G_{0, n'_{l'}}^{(1, l')} + D_{32} G_{0, n'_{l'}}^{(2, l')} + D_{33} G_{0, n'_{l'}}^{(3, l')} = -\bar{F}_{n'_{l'}}^{(3, l')}, \quad (5.23)$$

where

$$D_{11} = z - W^1 - t_1'^{\dagger} G_{1, 1}^1 t'_1 - t'_{(1, \bar{2})} G_{N_{\bar{2}}, N_{\bar{2}}}^{\bar{2}} t'^{\dagger}_{(1, \bar{2})} - t'^{\dagger}_{(1, \bar{3})} G_{1, 1}^{\bar{3}} t'_{(1, \bar{3})}, \quad (5.24)$$

$$D_{12} = -t'^{\dagger}_{(1, \bar{3})} G_{1, N_{\bar{3}}}^{\bar{3}} t'^{\dagger}_{(2, \bar{3})}, \quad (5.25)$$

$$D_{13} = -t'_{(1, \bar{2})} G_{N_{\bar{2}}, 1}^{\bar{2}} t'_{(3, \bar{2})}, \quad (5.26)$$

$$D_{21} = -t'_{(2, \bar{3})} G_{N_{\bar{3}}, 1}^{\bar{3}} t'_{(1, \bar{3})}, \quad (5.27)$$

$$D_{22} = z - W^2 - t_2'^{\dagger} G_{1, 1}^2 t'_2 - t'_{(2, \bar{3})} G_{N_{\bar{3}}, N_{\bar{3}}}^{\bar{3}} t'^{\dagger}_{(2, \bar{3})} - t'^{\dagger}_{(2, \bar{1})} G_{1, 1}^{\bar{1}} t'_{(2, \bar{1})}, \quad (5.28)$$

$$D_{23} = -t'^{\dagger}_{(2, \bar{1})} G_{1, N_{\bar{1}}}^{\bar{1}} t'^{\dagger}_{(3, \bar{1})}, \quad (5.29)$$

$$D_{31} = -t'^{\dagger}_{(3, \bar{2})} G_{1, N_{\bar{2}}}^{\bar{2}} t'^{\dagger}_{(1, \bar{2})}, \quad (5.30)$$

$$D_{32} = -t'_{(3, \bar{1})} G_{N_{\bar{1}}, 1}^{\bar{1}} t'_{(2, \bar{1})}, \quad (5.31)$$

$$D_{33} = z - W^3 - t_3'^{\dagger} G_{1, 1}^3 t'_3 - t'_{(3, \bar{1})} G_{N_{\bar{1}}, N_{\bar{1}}}^{\bar{1}} t'^{\dagger}_{(3, \bar{1})} - t'^{\dagger}_{(3, \bar{2})} G_{1, 1}^{\bar{2}} t'_{(3, \bar{2})}, \quad (5.32)$$

and

$$\bar{F}_{n'_{l'}}^{(1, l')} = -\delta_{l', 1} \delta_{n'_{1, 0}} + \delta_{l', 1} t_1'^{\dagger} G_{1, n'_1}^1 + \delta_{l', 2} t'_{(1, \bar{2})} G_{N_{\bar{2}}, n'_2}^{\bar{2}} + \delta_{l', 3} t'^{\dagger}_{(1, \bar{3})} G_{1, n'_3}^{\bar{3}}, \quad (5.33)$$

$$\bar{F}_{n'_{l'}}^{(2, l')} = -\delta_{l', 2} \delta_{n'_{2, 0}} + \delta_{l', 2} t_2'^{\dagger} G_{1, n'_2}^2 + \delta_{l', 3} t'_{(2, \bar{3})} G_{N_{\bar{3}}, n'_3}^{\bar{3}} + \delta_{l', \bar{1}} t'^{\dagger}_{(2, \bar{1})} G_{1, n'_1}^{\bar{1}}, \quad (5.34)$$

$$\bar{F}_{n'_{l'}}^{(3, l')} = -\delta_{l', 3} \delta_{n'_{3, 0}} + \delta_{l', 3} t_3'^{\dagger} G_{1, n'_3}^3 + \delta_{l', \bar{1}} t'_{(3, \bar{1})} G_{N_{\bar{1}}, n'_1}^{\bar{1}} + \delta_{l', 2} t'^{\dagger}_{(3, \bar{2})} G_{1, n'_2}^{\bar{2}}. \quad (5.35)$$

In the compact form, the equation for $G_{0, n'_{l'}}^{(\mu', l')}$ reads

$$\sum_{\mu'=1}^3 D_{\mu, \mu'} G_{0, n'_{l'}}^{(\mu', l')} = -\bar{F}_{n'_{l'}}^{(\mu, l')}, \quad \mu = 1, 2, 3. \quad (5.36)$$

Inverting the block matrix D , we obtain the solution

$$G_{0, n'_{l'}}^{(\mu, l')} = -\sum_{\mu'=1}^3 (D^{-1})_{\mu, \mu'} \bar{F}_{n'_{l'}}^{(\mu', l')}. \quad (5.37)$$

Next, we express the full Green's functions in terms of $G_{0,n'_l}^{(\mu',l')}$, with $\mu' = 1, 2, 3$:

$$\begin{aligned} G_{n_l, n'_l}^{(l,l')} &= \bar{G}_{n_l, n'_l}^{(l,l')} - \frac{\delta_{l,1}\delta_{n_1,0}\delta_{l',1}\delta_{n'_1,0}}{z - W^1} - \frac{\delta_{l,2}\delta_{n_2,0}\delta_{l',2}\delta_{n'_2,0}}{z - W^2} - \frac{\delta_{l,3}\delta_{n_3,0}\delta_{l',3}\delta_{n'_3,0}}{z - W^3} \\ &- F_{n_l}^{(l,1)}G_{0,n'_l}^{(1,l')} - \frac{\delta_{l,1}\delta_{n_1,0}}{z - W^1} \left[\bar{F}_{n'_l}^{(1,l')} + D_{11}G_{0,n'_l}^{(1,l')} + D_{12}G_{0,n'_l}^{(2,l')} + D_{13}G_{0,n'_l}^{(3,l')} \right] \\ &- F_{n_l}^{(l,2)}G_{0,n'_l}^{(2,l')} - \frac{\delta_{l,2}\delta_{n_2,0}}{z - W^2} \left[\bar{F}_{n'_l}^{(2,l')} + D_{21}G_{0,n'_l}^{(1,l')} + D_{22}G_{0,n'_l}^{(2,l')} + D_{23}G_{0,n'_l}^{(3,l')} \right] \\ &- F_{n_l}^{(l,3)}G_{0,n'_l}^{(3,l')} - \frac{\delta_{l,3}\delta_{n_3,0}}{z - W^3} \left[\bar{F}_{n'_l}^{(3,l')} + D_{31}G_{0,n'_l}^{(1,l')} + D_{32}G_{0,n'_l}^{(2,l')} + D_{33}G_{0,n'_l}^{(3,l')} \right], \end{aligned} \quad (5.38)$$

where

$$F_{n_l}^{(l,1)} = -\delta_{l,1}\delta_{n_1,0} + \delta_{l,1}G_{n_1,1}^1 t'_1 + \delta_{l,2}G_{n_2,N_2}^2 t'_{(1,2)} + \delta_{l,3}G_{n_3,1}^3 t'_{(1,3)}, \quad (5.39)$$

$$F_{n_l}^{(l,2)} = -\delta_{l,2}\delta_{n_2,0} + \delta_{l,2}G_{n_2,1}^2 t'_2 + \delta_{l,3}G_{n_3,N_3}^3 t'_{(2,3)} + \delta_{l,1}G_{n_1,1}^1 t'_{(2,1)}, \quad (5.40)$$

$$F_{n_l}^{(l,3)} = -\delta_{l,3}\delta_{n_3,0} + \delta_{l,3}G_{n_3,1}^3 t'_3 + \delta_{l,1}G_{n_1,N_1}^1 t'_{(3,1)} + \delta_{l,2}G_{n_2,1}^2 t'_{(3,2)}. \quad (5.41)$$

Taking into account the relations (5.36) and (5.37), we establish that

$$G_{n_l, n'_l}^{(l,l')} = \delta_{l,l'} G_{n_l, n'_l}^l - \sum_{\mu=1}^3 F_{n_l}^{(l,\mu)} G_{0,n'_l}^{(\mu,l')} \quad (5.42)$$

$$= \delta_{l,l'} G_{n_l, n'_l}^l + \sum_{\mu, \mu'=1}^3 F_{n_l}^{(l,\mu)} (D^{-1})_{\mu, \mu'} \bar{F}_{n'_l}^{(\mu', l')}. \quad (5.43)$$

5.3 Global DOS and the bound state equation

Consider

$$\sum_l \sum_{n_l} \text{tr} \left[G_{n_l, n_l}^{(l,l)} - G_{n_l, n_l}^l \right] = \sum_l \sum_{n_l} \sum_{\mu, \mu'=1}^3 \text{tr} \left[(D^{-1})_{\mu, \mu'} \bar{F}_{n_l}^{(\mu', l)} F_{n_l}^{(l, \mu)} \right]. \quad (5.44)$$

Let us focus on

$$P_{\mu', \mu} = \sum_l \sum_{n_l} \bar{F}_{n_l}^{(\mu', l)} F_{n_l}^{(l, \mu)}. \quad (5.45)$$

Using the property of the isolated Green's function of the l -th subsystem

$$\sum_{n'' \in \{n_l\}} G_{n, n''}^l G_{n'', n'}^l = -\frac{\partial G_{n, n'}^l}{\partial z}, \quad n, n' \in \{n_l\}, \quad (5.46)$$

which is most easily proven on the basis of the Lehmann representation, we inspect one by one the components of $P_{\mu', \mu}$. So we get

$$\begin{aligned} P_{11} &= \sum_l \sum_{n_l} \left(-\delta_{l,1}\delta_{n_1,0} + \delta_{l,1}t_1'^{\dagger} G_{1,n_1}^1 + \delta_{l,2}t'_{(1,2)} G_{N_2, n_2}^2 + \delta_{l,3}t'_{(1,3)} G_{1, n_3}^3 \right) \\ &\times \left(-\delta_{l,1}\delta_{n_1,0} + \delta_{l,1}G_{n_1,1}^1 t'_1 + \delta_{l,2}G_{n_2, N_2}^2 t'_{(1,2)} + \delta_{l,3}G_{n_3,1}^3 t'_{(1,3)} \right) \end{aligned} \quad (5.47)$$

$$= 1 + \sum_{n_1} t_1'^{\dagger} G_{1, n_1}^1 G_{n_1, 1}^1 t'_1 + \sum_{n_2} t'_{(1,2)} G_{N_2, n_2}^2 G_{n_2, N_2}^2 t'_{(1,2)} + \sum_{n_3} t'_{(1,3)} G_{1, n_3}^3 G_{n_3, 1}^3 t'_{(1,3)} \quad (5.48)$$

$$= \frac{\partial D_{11}}{\partial z}, \quad (5.49)$$

$$P_{12} = \sum_l \sum_{n_l} \left(-\delta_{l,1} \delta_{n_1,0} + \delta_{l,1} t_1'^{\dagger} G_{1,n_1}^1 + \delta_{l,2} t'_{(1,2)} G_{N_2,n_2}^2 + \delta_{l,3} t'_{(1,3)}^{\dagger} G_{1,n_3}^3 \right) \\ \times \left(-\delta_{l,2} \delta_{n_2,0} + \delta_{l,2} G_{n_2,1}^2 t'_2 + \delta_{l,3} G_{n_3,N_3}^3 t'_{(2,3)}^{\dagger} + \delta_{l,1} G_{n_1,1}^1 t'_{(2,1)} \right) \quad (5.50)$$

$$= \sum_{n_3} t'_{(1,3)}^{\dagger} G_{1,n_3}^3 G_{n_3,N_3}^3 t'_{(2,3)}^{\dagger} = \frac{\partial D_{12}}{\partial z}, \quad (5.51)$$

$$P_{13} = \sum_l \sum_{n_l} \left(-\delta_{l,1} \delta_{n_1,0} + \delta_{l,1} t_1'^{\dagger} G_{1,n_1}^1 + \delta_{l,2} t_{(1,2)} G_{N_2,n_2}^2 + \delta_{l,3} t'_{(1,3)}^{\dagger} G_{1,n_3}^3 \right) \\ \times \left(-\delta_{l,3} \delta_{n_3,0} + \delta_{l,3} G_{n_3,1}^3 t'_3 + \delta_{l,1} G_{n_1,N_1}^1 t'_{(3,1)}^{\dagger} + \delta_{l,2} G_{n_2,1}^2 t'_{(3,2)} \right) \quad (5.52)$$

$$= \sum_{n_2} t_{(1,2)} G_{N_2,n_2}^2 G_{n_2,1}^2 t'_{(3,2)} = \frac{\partial D_{13}}{\partial z}, \quad (5.53)$$

$$P_{21} = \sum_l \sum_{n_l} \left(-\delta_{l,2} \delta_{n_2,0} + \delta_{l,2} t_2'^{\dagger} G_{1,n_2}^2 + \delta_{l,3} t'_{(2,3)} G_{N_3,n_3}^3 + \delta_{l,1} t'_{(2,1)}^{\dagger} G_{1,n_1}^1 \right) \\ \times \left(-\delta_{l,1} \delta_{n_1,0} + \delta_{l,1} G_{n_1,1}^1 t'_1 + \delta_{l,2} G_{n_2,N_2}^2 t'_{(1,2)}^{\dagger} + \delta_{l,3} G_{n_3,1}^3 t'_{(1,3)} \right) \quad (5.54)$$

$$= \sum_{n_3} t'_{(2,3)} G_{N_3,n_3}^3 G_{n_3,1}^3 t'_{(1,3)} = \frac{\partial D_{21}}{\partial z}, \quad (5.55)$$

$$P_{22} = \sum_l \sum_{n_l} \left(-\delta_{l,2} \delta_{n_2,0} + \delta_{l,2} t_2'^{\dagger} G_{1,n_2}^2 + \delta_{l,3} t'_{(2,3)} G_{N_3,n_3}^3 + \delta_{l,1} t'_{(2,1)}^{\dagger} G_{1,n_1}^1 \right) \\ \times \left(-\delta_{l,2} \delta_{n_2,0} + \delta_{l,2} G_{n_2,1}^2 t'_2 + \delta_{l,3} G_{n_3,N_3}^3 t'_{(2,3)}^{\dagger} + \delta_{l,1} G_{n_1,1}^1 t'_{(2,1)} \right) \quad (5.56)$$

$$= 1 + \sum_{n_2} t_2'^{\dagger} G_{1,n_2}^2 G_{n_2,1}^2 t'_2 + \sum_{n_3} t'_{(2,3)} G_{N_3,n_3}^3 G_{n_3,N_3}^3 t'_{(2,3)}^{\dagger} + \sum_{n_1} t'_{(2,1)}^{\dagger} G_{1,n_1}^1 G_{n_1,1}^1 t'_{(2,1)} \quad (5.57)$$

$$= \frac{\partial D_{22}}{\partial z}, \quad (5.58)$$

$$P_{23} = \sum_l \sum_{n_l} \left(-\delta_{l,2} \delta_{n_2,0} + \delta_{l,2} t_2'^{\dagger} G_{1,n_2}^2 + \delta_{l,3} t'_{(2,3)} G_{N_3,n_3}^3 + \delta_{l,1} t'_{(2,1)}^{\dagger} G_{1,n_1}^1 \right) \\ \times \left(-\delta_{l,3} \delta_{n_3,0} + \delta_{l,3} G_{n_3,1}^3 t'_3 + \delta_{l,1} G_{n_1,N_1}^1 t'_{(3,1)}^{\dagger} + \delta_{l,2} G_{n_2,1}^2 t'_{(3,2)} \right) \quad (5.59)$$

$$= \sum_{\bar{n}_1} t'_{(2,1)}^{\dagger} G_{1,n_1}^1 G_{n_1,N_1}^1 t'_{(3,1)}^{\dagger} = \frac{\partial D_{23}}{\partial z}, \quad (5.60)$$

$$P_{31} = \sum_l \sum_{n_l} \left(-\delta_{l,3} \delta_{n_3,0} + \delta_{l,3} t_3'^{\dagger} G_{1,n_3}^3 + \delta_{l,1} t'_{(3,1)} G_{N_1,n_1}^1 + \delta_{l,2} t'_{(3,2)}^{\dagger} G_{1,n_2}^2 \right) \\ \times \left(-\delta_{l,1} \delta_{n_1,0} + \delta_{l,1} G_{n_1,1}^1 t'_1 + \delta_{l,2} G_{n_2,N_2}^2 t'_{(1,2)}^{\dagger} + \delta_{l,3} G_{n_3,1}^3 t'_{(1,3)} \right) \quad (5.61)$$

$$= \sum_{n_2} t'_{(3,2)}^{\dagger} G_{1,n_2}^2 G_{n_2,N_2}^2 t'_{(1,2)}^{\dagger} = \frac{\partial D_{31}}{\partial z}, \quad (5.62)$$

$$P_{32} = \sum_l \sum_{n_l} \left(-\delta_{l,3} \delta_{n_3,0} + \delta_{l,3} t_3'^{\dagger} G_{1,n_3}^3 + \delta_{l,1} t'_{(3,1)} G_{N_1,n_1}^1 + \delta_{l,2} t'_{(3,2)}^{\dagger} G_{1,n_2}^2 \right) \\ \times \left(-\delta_{l,2} \delta_{n_2,0} + \delta_{l,2} G_{n_2,1}^2 t'_2 + \delta_{l,3} G_{n_3,N_3}^3 t'_{(2,3)}^{\dagger} + \delta_{l,1} G_{n_1,1}^1 t'_{(2,1)} \right) \quad (5.63)$$

$$= \sum_{n_1} t'_{(3,1)} G_{N_1,n_1}^1 G_{n_1,1}^1 t'_{(2,1)} = \frac{\partial D_{32}}{\partial z}, \quad (5.64)$$

$$P_{33} = \sum_l \sum_{n_l} \left(-\delta_{l,3} \delta_{n_3,0} + \delta_{l,3} t_3'^{\dagger} G_{1,n_3}^3 + \delta_{l,\bar{1}} t'_{(3,\bar{1})} G_{N_{\bar{1}},n_{\bar{1}}}^{\bar{1}} + \delta_{l,\bar{2}} t'_{(3,\bar{2})} G_{1,n_{\bar{2}}}^{\bar{2}} \right) \\ \times \left(-\delta_{l,3} \delta_{n_3,0} + \delta_{l,3} G_{n_3,1}^3 t_3' + \delta_{l,\bar{1}} G_{n_{\bar{1}},N_{\bar{1}}}^{\bar{1}} t'_{(3,\bar{1})} + \delta_{l,\bar{2}} G_{n_{\bar{2}},1}^{\bar{2}} t'_{(3,\bar{2})} \right) \quad (5.65)$$

$$= 1 + \sum_{n_3} t_3'^{\dagger} G_{1,n_3}^3 G_{n_3,1}^3 + \sum_{n_{\bar{1}}} t'_{(3,\bar{1})} G_{N_{\bar{1}},n_{\bar{1}}}^{\bar{1}} G_{n_{\bar{1}},N_{\bar{1}}}^{\bar{1}} t'_{(3,\bar{1})} + \sum_{n_{\bar{2}}} t'_{(3,\bar{2})} G_{1,n_{\bar{2}}}^{\bar{2}} G_{n_{\bar{2}},1}^{\bar{2}} t'_{(3,\bar{2})} \quad (5.66)$$

$$= \frac{\partial D_{33}}{\partial z}. \quad (5.67)$$

Thus we conclude that

$$P_{\mu',\mu} = \frac{\partial D_{\mu',\mu}}{\partial z}. \quad (5.68)$$

It follows

$$\sum_l \sum_{n_l} \text{tr} \left[G_{n_l,n_l}^{(l,l)} - G_{n_l,n_l}^l \right] = \sum_{\mu,\mu'=1}^3 \text{tr} \left[(D^{-1})_{\mu,\mu'} \frac{\partial D_{\mu',\mu}}{\partial z} \right] \quad (5.69)$$

$$= \text{Tr} \left[D^{-1} \frac{\partial D}{\partial z} \right] = \frac{\partial}{\partial z} \ln \text{Det } D. \quad (5.70)$$

In this expression, both Tr and Det are understood as operations with the block-formed matrix D .

The correction to the global density of states (DOS) is then given by

$$\delta\rho(\omega) = -\frac{1}{\pi} \text{Im} \frac{\partial}{\partial z} \ln \text{Det } D(z) \Big|_{z=\omega+i0^+}. \quad (5.71)$$

The equation

$$\text{Det } D(\omega) = 0 \quad (5.72)$$

determines bound states present in the coupled systems (NB: it might also have roots corresponding to bound states of the decoupled systems, so a careful sorting is necessary!). Indeed, whenever the following relation holds

$$\text{Det } D(\omega + i0^+) \approx \frac{\omega + i0^+ - E_A}{a_A}, \quad (5.73)$$

the corresponding contribution to (5.71) reads

$$\delta\rho_A(\omega) = -\frac{1}{\pi} \text{Im} \frac{1}{\omega + i0^+ - E_A} = \delta(\omega - E_A). \quad (5.74)$$

5.4 Josephson currents

We consider the following Hamiltonians H_m , $H_{\bar{m}}$ of the isolated subsystems S_m , $\bar{S}_{\bar{m}}$ in the extended coordinate Nambu representation $\Psi_j = (c_{j,\uparrow}, c_{j,\downarrow}, c_{j,\downarrow}^\dagger, -c_{j,\uparrow}^\dagger)^T \equiv (c_j, i\sigma_y c_j^\dagger)^T$:

$$H_m = \frac{1}{2} \sum_{n_m, n'_m \in S_m} \Psi_{n_m}^\dagger \begin{pmatrix} (H_m^e)_{n_m, n'_m} & \Delta_m \delta_{n_m, n'_m} \\ \Delta_m^* \delta_{n_m, n'_m} & (H_m^h)_{n_m, n'_m} \end{pmatrix} \Psi_{n'_m} \equiv \frac{1}{2} \Psi^\dagger h_m \Psi, \quad (5.75)$$

$$H_{\bar{m}} = \frac{1}{2} \sum_{n_{\bar{m}}, n'_{\bar{m}} \in S_{\bar{m}}} \Psi_{n_{\bar{m}}}^\dagger \begin{pmatrix} (H_{\bar{m}}^e)_{n_{\bar{m}}, n'_{\bar{m}}} & 0 \\ 0 & (H_{\bar{m}}^h)_{n_{\bar{m}}, n'_{\bar{m}}} \end{pmatrix} \Psi_{n'_{\bar{m}}}. \quad (5.76)$$

The (particle) Hamiltonians $(H_m^e)_{n_m, n'_m}$ of the superconducting leads contain only spin-independent nearest-neighbour hopping amplitudes $-t$ and onsite chemical potentials $-\mu$. In turn, the (particle) Hamiltonians $(H_m^e)_{n_m, n'_m}$ of the semiconducting segments contain spin-dependent nearest-neighbour hopping matrix $-\bar{t}$ and spin-dependent onsite potentials (accounting the effect of the Zeeman field). The corresponding hole Hamiltonians are obtained by the particle-hole symmetry

$$H^h = -\sigma_y (H^e)^* \sigma_y. \quad (5.77)$$

The superconducting order parameters $\Delta_m = \Delta_0 e^{i\varphi_m}$ are characterized by the same absolute value Δ_0 and the lead-specific phases φ_m .

The operator of the current flowing out of the m -th lead and its average stationary value are given by

$$\hat{J}_m = -\frac{ie}{\hbar} [H, \hat{N}_m] = -\frac{ie}{\hbar} \sum_{\sigma} \left(c_{m,1,\sigma}^\dagger t'_{m,\sigma\sigma'} c_{m,0,\sigma'} - c_{m,0,\sigma'}^\dagger t_{m,\sigma\sigma'}^* c_{m,1,\sigma} \right), \quad (5.78)$$

$$J_m \equiv \langle J_m(t) \rangle = -\frac{ie}{\hbar} \sum_{\sigma} \left(t'_{m,\sigma\sigma'} \langle c_{m,1,\sigma}^\dagger(t) c_{m,0,\sigma'}(t) \rangle - t_{m,\sigma\sigma'}^* \langle c_{m,0,\sigma'}^\dagger(t) c_{m,1,\sigma}(t) \rangle \right). \quad (5.79)$$

Applying the Meir-Wingreen approach, we establish (see another note) we obtain the expression

$$J_m = \frac{2e}{\hbar} \text{Re} \int d\omega f(\omega) \text{tr} \left\{ \frac{1 + \tau_z}{2} \left[G_{0,1}^{(m,m)}(\omega + i0^+) \hat{t}'_m - \tilde{t}_m^\dagger G_{1,0}^{(m,m)}(\omega + i0^+) \right] \right\} \quad (5.80)$$

$$= \frac{2e}{\hbar} \text{Re} \int d\omega f(\omega) \text{tr} \left\{ \frac{\tau_z}{2} \left[G_{0,1}^{(m,m)}(\omega + i0^+) \hat{t}'_m - \tilde{t}_m^\dagger G_{1,0}^{(m,m)}(\omega + i0^+) \right] \right\} \quad (5.81)$$

in terms of the full system's Nambu-space (retarded) Green's function (5.43). Hereby $f(\omega)$ is the Fermi distribution function, and

$$\hat{t}'_m = \begin{pmatrix} t'_m & 0 \\ 0 & -\sigma_y t_{m'}^* \sigma_y \end{pmatrix} = -C \tilde{t}_m^* C, \quad C = \tau_y \sigma_y. \quad (5.82)$$

In the following we omit hats identifying t'_m with its Nambu representation.

In particular, on the basis of (5.43) and the definitions of the ingredients entering this formula we find the following expressions

$$G_{0,1}^{(m,m)} = \sum_{\mu, \mu'=1}^3 F_0^{(m,\mu)}(D^{-1})_{\mu, \mu'} \bar{F}_1^{(\mu', m)}, \quad (5.83)$$

$$G_{1,0}^{(m,m)} = \sum_{\mu, \mu'=1}^3 F_1^{(m,\mu)}(D^{-1})_{\mu, \mu'} \bar{F}_0^{(\mu', m)}. \quad (5.84)$$

Representing

$$J_m = \frac{2e}{\hbar} \text{Re} \int d\omega f(\omega) \sum_{\mu, \mu'=1}^3 \text{tr} \left\{ (D^{-1})_{\mu, \mu'} R_{\mu', \mu}^m \right\} \quad (5.85)$$

in terms of

$$R_{\mu', \mu}^m = \bar{F}_1^{(\mu', m)} t'_m \frac{1 + \tau_z}{2} F_0^{(m, \mu)} - \bar{F}_0^{(\mu', m)} \frac{1 + \tau_z}{2} t_m^\dagger F_1^{(m, \mu)}, \quad (5.86)$$

we evaluate the latter matrix using

$$\bar{F}_1^{(\mu',m)} = \delta_{\mu',m} t_m'^{\dagger} G_{1,1}^m, \quad (5.87)$$

$$\bar{F}_0^{(\mu',m)} = -\delta_{\mu',m}, \quad (5.88)$$

$$F_1^{(m,\mu)} = \delta_{m,\mu} G_{1,1}^m t_m', \quad (5.89)$$

$$F_0^{(m,\mu)} = -\delta_{m,\mu}. \quad (5.90)$$

It follows

$$R_{\mu',\mu}^m = -\delta_{\mu',m} \delta_{m,\mu} \frac{1}{2} t_m'^{\dagger} [G_{1,1}^m, \tau_z] t_m'. \quad (5.91)$$

Next we note that the phase dependence of G_{11}^m has the form

$$G_{11}^m = e^{\frac{i}{2}\varphi_m \tau_z} G_{11}^{m,0} e^{-\frac{i}{2}\varphi_m \tau_z}, \quad (5.92)$$

where $G_{11}^{m,0}$ is evaluated at $\varphi_m = 0$, as ensuing from the relation $h_m = e^{\frac{i}{2}\varphi_m \tau_z} h_{m,0} e^{-\frac{i}{2}\varphi_m \tau_z}$. It follows

$$\frac{\partial G_{11}^m}{\partial \varphi_m} = -\frac{i}{2} [G_{1,1}^m, \tau_z]. \quad (5.93)$$

Hence

$$R_{\mu',\mu}^m = i \frac{\partial D_{\mu',\mu}}{\partial \varphi_m} \quad (5.94)$$

and

$$J_m = -\frac{2e}{h} \text{Im} \int d\omega f(\omega) \text{Tr} \left\{ D^{-1} \frac{\partial D}{\partial \varphi_m} \right\} \quad (5.95)$$

$$= -\frac{2e}{h} \frac{\partial}{\partial \varphi_m} \text{Im} \int d\omega f(\omega) \ln \text{Det } D = -\frac{2e}{h} \frac{\partial F}{\partial \varphi_m}. \quad (5.96)$$

In fact, the free energy depends on the two gauge invariant combinations (tuned by two independent external magnetic fluxes)

$$\varphi_L = \varphi_3 - \varphi_1, \quad (5.97)$$

$$\varphi_R = \varphi_2 - \varphi_3. \quad (5.98)$$

Thus

$$J_1 = -\frac{2e}{h} \frac{\partial F}{\partial \varphi_1} = \frac{2e}{h} \frac{\partial F}{\partial \varphi_L} \equiv J_L, \quad (5.99)$$

$$J_2 = -\frac{2e}{h} \frac{\partial F}{\partial \varphi_2} = -\frac{2e}{h} \frac{\partial F}{\partial \varphi_R} \equiv -J_R, \quad (5.100)$$

$$J_3 = -\frac{2e}{h} \frac{\partial F}{\partial \varphi_3} = -\frac{2e}{h} \frac{\partial F}{\partial \varphi_L} + \frac{2e}{h} \frac{\partial F}{\partial \varphi_R} \equiv -J_L + J_R. \quad (5.101)$$

5.5 Continuum model

As one can see from the previous consideration, the whole spectral information is contained in the quantity $\text{Det } D$. The matrix D receives additive contributions from all the subsystems. Therefore this representation is very convenient for study a continuum limit. It can be taken

in two different ways: 1) approximating the scattering region by just three sites with spin-dependent hoppings; 2) by taking the continuum limit in all subsystems (both superconducting leads and normal segments) as in the recent PRB paper. In the following we consider the both options.

In every case, it is possible to decompose

$$D = \begin{pmatrix} D^{ee} & D^{eh} \\ D^{he} & D^{hh} \end{pmatrix}, \quad (5.102)$$

which simplifies the calculation of

$$\text{Det } D = \text{Det } (D^{ee}) \text{Det } (D^{hh}) \text{Det } [1 - (D^{ee})^{-1} D^{eh} (D^{hh})^{-1} D^{he}]. \quad (5.103)$$

In this expression, only the last factor depends on the phases φ_m .

5.5.1 Three quantum dots

In this case the matrix D becomes the inverse Green's function of the three-dot subsystem:

$$D = G_{\text{dots}}^{-1} = \begin{pmatrix} z - W^1 - t_1'^{\dagger} G_{1,1}^1 t_1' & t_{21}'^{\dagger} & t_{13}' \\ t_{21}' & z - W^2 - t_2'^{\dagger} G_{1,1}^2 t_2' & t_{32}'^{\dagger} \\ t_{13}'^{\dagger} & t_{32}' & z - W^3 - t_3'^{\dagger} G_{1,1}^3 t_3' \end{pmatrix}, \quad (5.104)$$

where t_{ij}' describe a direct spin-dependent hopping from the site j to the site i , with $i, j = 1, 2, 3$.

The leads induced self-energy depends on $G_{1,1}^m$. We assume that all leads have the same structure, besides the phases φ_m . Thus we can write

$$G_{1,1}^m = e^{\frac{i}{2}\varphi_m\tau_z} G_{1,1} e^{-\frac{i}{2}\varphi_m\tau_z}, \quad (5.105)$$

where

$$G_{1,1} = G_0^{(0)} - G_1^{(0)} [G_0^{(0)}]^{-1} G_{-1}^{(0)} \quad (5.106)$$

is the superconductor's boundary Green's function expressed in terms of the translationally invariant superconductor's Green's function

$$G_{n-n'}^{(0)}(z) = \int_{-\pi}^{\pi} \frac{dk}{2\pi} \frac{e^{ik(n-n')}}{z - (-2t \cos k - \mu)\tau_z - \Delta_0\tau_x}. \quad (5.107)$$

Considering this integral in the wide-band limit $t \gg \Delta_0$, we approximate it by linearizing near the Fermi points $\mu = 0$, $k = \pm \frac{\pi}{2}$:

$$\begin{aligned} G_{n-n'}^{(0)}(z) &\approx e^{i\frac{\pi}{2}(n-n')} \int_{-\infty}^{\infty} \frac{dq}{2\pi} \frac{e^{iq(n-n')}}{z - 2tq\tau_z - \Delta_0\tau_x} \\ &+ e^{-i\frac{\pi}{2}(n-n')} \int_{-\infty}^{\infty} \frac{dq}{2\pi} \frac{e^{iq(n-n')}}{z + 2tq\tau_z - \Delta_0\tau_x}. \end{aligned} \quad (5.108)$$

The relevant components are

$$G_0^{(0)}(z) \approx \frac{1}{t} \int_{-\infty}^{\infty} \frac{d\bar{q}}{2\pi} \frac{z + \Delta_0\tau_x}{z^2 - \bar{q}^2 - \Delta_0^2} = -\frac{1}{t} \int_{-\infty}^{\infty} \frac{d\bar{q}}{2\pi} \frac{z + \Delta_0\tau_x}{(\bar{q} - k_+)(\bar{q} - k_-)}. \quad (5.109)$$

Hereby

$$k_{\pm}(z) = \pm z \sqrt{1 - \frac{\Delta_0^2}{z^2}}, \quad (5.110)$$

such that for $z = \omega + i0^+$ it holds $\text{Im } k_+ > 0$ and $\text{Im } k_- < 0$. In particular,

$$k_+(\omega + i0^+) = \begin{cases} \sqrt{(\omega + i0^+)^2 - \Delta_0^2}, & \omega > \Delta_0, \\ i\sqrt{\Delta_0^2 - \omega^2}, & |\omega| < \Delta_0, \\ -\sqrt{(\omega + i0^+)^2 - \Delta_0^2}, & \omega < -\Delta_0. \end{cases} \quad (5.111)$$

Closing the integration contour in the upper half-plane, we obtain

$$G_0^{(0)}(z) \approx -\frac{i}{2t} \frac{z + \Delta_0 \tau_x}{k_+} = -\frac{i}{2t} \frac{1 + \frac{\Delta_0}{z} \tau_x}{\sqrt{1 - \frac{\Delta_0^2}{z^2}}}. \quad (5.112)$$

We also establish

$$[G_0^{(0)}(z)]^{-1} = 2ti \frac{1 - \frac{\Delta_0}{z} \tau_x}{\sqrt{1 - \frac{\Delta_0^2}{z^2}}}. \quad (5.113)$$

Analogously we evaluate

$$G_1^{(0)}(z) \approx \frac{i\tau_z}{t} \int_{-\infty}^{\infty} \frac{d\bar{q}}{2\pi} \frac{\bar{q} e^{i\frac{\bar{q}}{2t}}}{z^2 - \bar{q}^2 - \Delta_0^2} \quad (5.114)$$

$$= -\frac{i\tau_z}{t} \int_{-\infty}^{\infty} \frac{d\bar{q}}{2\pi} \frac{\bar{q} e^{i\frac{\bar{q}}{2t}}}{(\bar{q} - k_+)(\bar{q} - k_-)} \approx \frac{\tau_z}{2t}. \quad (5.115)$$

Finally, we find

$$G_{-1}^{(0)}(z) \approx -\frac{i\tau_z}{t} \int_{-\infty}^{\infty} \frac{d\bar{q}}{2\pi} \frac{\bar{q} e^{-i\frac{\bar{q}}{2t}}}{z^2 - \bar{q}^2 - \Delta_0^2} = G_1^{(0)}(z), \quad (5.116)$$

where the last equality is obtained by re-defining the integration variable $\bar{q} \rightarrow -\bar{q}$.

Collecting together all the contribution to $G_{1,1}$, we reveal

$$G_{1,1} = -\frac{i}{2t} \frac{1 + \frac{\Delta_0}{z} \tau_x}{\sqrt{1 - \frac{\Delta_0^2}{z^2}}} - \frac{\tau_z}{2t} 2ti \frac{1 - \frac{\Delta_0}{z} \tau_x}{\sqrt{1 - \frac{\Delta_0^2}{z^2}}} \frac{\tau_z}{2t} \quad (5.117)$$

$$= -\frac{i}{t} \frac{1 + \frac{\Delta_0}{z} \tau_x}{\sqrt{1 - \frac{\Delta_0^2}{z^2}}}. \quad (5.118)$$

Assuming that for all three leads $t'_m = t' \tau_z$, we find

$$t_m'^{\dagger} G_{1,1}^m t'_m = e^{\frac{i}{2} \varphi_m \tau_z} \left[-i\Gamma_S \frac{1 - \frac{\Delta_0}{z} \tau_x}{\sqrt{1 - \frac{\Delta_0^2}{z^2}}} \right] e^{-\frac{i}{2} \varphi_m \tau_z} \equiv -i\Gamma_S e^{\frac{i}{2} \varphi_m \tau_z} F(z) e^{-\frac{i}{2} \varphi_m \tau_z}, \quad (5.119)$$

where $\Gamma_S = \frac{(t')^2}{t}$, and

$$F(z) = \frac{1 - \frac{\Delta_0}{z} \tau_x}{\sqrt{1 - \frac{\Delta_0^2}{z^2}}}. \quad (5.120)$$

We note in passing that the phase dependence of D can be shifted to from the diagonal to the off-diagonal elements by means of the transformation

$$\tilde{D} = \begin{pmatrix} e^{-\frac{i}{2}\varphi_1\tau_z} & 0 & 0 \\ 0 & e^{-\frac{i}{2}\varphi_2\tau_z} & 0 \\ 0 & 0 & e^{-\frac{i}{2}\varphi_3\tau_z} \end{pmatrix} D \begin{pmatrix} e^{\frac{i}{2}\varphi_1\tau_z} & 0 & 0 \\ 0 & e^{\frac{i}{2}\varphi_2\tau_z} & 0 \\ 0 & 0 & e^{\frac{i}{2}\varphi_3\tau_z} \end{pmatrix}. \quad (5.121)$$

As a result, one can explicitly see the dependence of $\text{Det } D$ on the gauge-invariant phase differences φ_L and φ_R .

5.5.2 Modelling of the hopping amplitudes

We continue modelling the hopping amplitudes in (5.104). Let us focus on the particle sector. In the absence of the magnetic field we have the time-reversal (TR) invariance described by the anti-unitary operator $T = i\sigma_y K$, where K is a complex conjugation. This symmetry constrains the hopping amplitudes to the form

$$t'_{21} = t_d e^{-i\varphi_{SO}\vec{e}_3 \cdot \vec{\sigma}}, \quad (5.122)$$

$$t'_{32} = t_d e^{-i\varphi_{SO}\vec{e}_1 \cdot \vec{\sigma}}, \quad (5.123)$$

$$t'_{13} = t_d e^{-i\varphi_{SO}\vec{e}_2 \cdot \vec{\sigma}}, \quad (5.124)$$

where $\varphi_{SO} = m\alpha L$ is a spin precession phase due to spin orbit coupling, and L is an effective segment's length. Choosing

$$\vec{e}_3 = \vec{e}_x, \quad (5.125)$$

$$\vec{e}_1 = \cos(-\varphi_0)\vec{e}_x + \sin(-\varphi_0)\vec{e}_y, \quad (5.126)$$

$$\vec{e}_2 = \cos(-2\varphi_0)\vec{e}_x + \sin(-2\varphi_0)\vec{e}_y, \quad (5.127)$$

where $\varphi_0 = \frac{2\pi}{3}$, and making use of the expression

$$e^{-i\frac{\varphi}{2}\sigma_z}\sigma_x e^{i\frac{\varphi}{2}\sigma_z} = \cos \varphi \sigma_x + \sin \varphi \sigma_y, \quad (5.128)$$

we represent

$$t'_{21} = t_d e^{-i\varphi_{SO}\sigma_x} \equiv \hat{t}_d, \quad (5.129)$$

$$t'_{32} = t_d e^{i\frac{\varphi_0}{2}\sigma_z} e^{-i\varphi_{SO}\sigma_x} e^{-i\frac{\varphi_0}{2}\sigma_z} = U \hat{t}_d U^\dagger, \quad U \equiv e^{i\frac{\varphi_0}{2}\sigma_z}, \quad (5.130)$$

$$t'_{13} = t_d e^{i\frac{2\varphi_0}{2}\sigma_z} e^{-i\varphi_{SO}\sigma_x} e^{-i\frac{2\varphi_0}{2}\sigma_z} = U^2 \hat{t}_d (U^\dagger)^2. \quad (5.131)$$

The matrix $D = D_N + D_S$ for this model is given as the sum of the terms

$$D_N = \begin{pmatrix} z - W^1 & \tau_z t_d e^{i\varphi_{SO}\sigma_x} & \tau_z t_d U^2 e^{-i\varphi_{SO}\sigma_x} (U^\dagger)^2 \\ \tau_z t_d e^{-i\varphi_{SO}\sigma_x} & z - W^2 & \tau_z t_d U e^{i\varphi_{SO}\sigma_x} U^\dagger \\ \tau_z t_d U^2 e^{i\varphi_{SO}\sigma_x} (U^\dagger)^2 & \tau_z t_d U e^{-i\varphi_{SO}\sigma_x} U^\dagger & z - W^3 \end{pmatrix}, \quad (5.132)$$

$$D_S = i\Gamma_S \begin{pmatrix} e^{\frac{i}{2}\varphi_1\tau_z} F(z) e^{-\frac{i}{2}\varphi_1\tau_z} & 0 & 0 \\ 0 & e^{\frac{i}{2}\varphi_2\tau_z} F(z) e^{-\frac{i}{2}\varphi_2\tau_z} & 0 \\ 0 & 0 & e^{\frac{i}{2}\varphi_3\tau_z} F(z) e^{-\frac{i}{2}\varphi_3\tau_z} \end{pmatrix}. \quad (5.133)$$

Neglecting the onsite potentials W^i (which may generally have a spin structure) and the coupling to the leads ($\Gamma_S = 0$), we express $D = z - \tau_z H_d$ in terms of the three-dot Hamiltonian

$$H_d = - \sum_{l=1}^N U^{l-1} |l+1\rangle \hat{t}_d \langle l| (U^\dagger)^{l-1} - \sum_{l=1}^N U^{l-1} |l\rangle \hat{t}_d^\dagger \langle l+1| (U^\dagger)^{l-1}, \quad (5.134)$$

where $N = 3$ in the present case. This Hamiltonian is TR-invariant, since both $\hat{t}_d$ and U are TR-invariant. So the spectrum must be twofold degenerate.

We try the following eigenfunction ansatz

$$|\Psi_k\rangle = \frac{1}{\sqrt{N}} \sum_{l=1}^N e^{ikl} U^{l-1} |l\rangle |\chi_k\rangle. \quad (5.135)$$

Substituting it into (5.134), we obtain

$$H_d |\Psi_k\rangle = \frac{1}{\sqrt{N}} \sum_{l=1}^N U^{l-1} e^{ikl} |l\rangle \left(-e^{-ik} U^\dagger \hat{t}_d - e^{ik} \hat{t}_d^\dagger U \right) |\chi_k\rangle. \quad (5.136)$$

Claiming

$$\left(-e^{-ik} U^\dagger \hat{t}_d - e^{ik} \hat{t}_d^\dagger U \right) |\chi_k\rangle = \varepsilon_k |\chi_k\rangle, \quad (5.137)$$

we also get

$$H_d |\Psi_k\rangle = \varepsilon_k |\Psi_k\rangle. \quad (5.138)$$

Rewriting the eigenvalue problem (5.137)

$$-t_d \left(e^{-ik} U^\dagger e^{-i\varphi_{SO}\sigma_x} + e^{ik} e^{i\varphi_{SO}\sigma_x} U \right) |\chi_k\rangle = \varepsilon_k |\chi_k\rangle, \quad (5.139)$$

we notice that the matrix

$$U_{SO} = e^{i\varphi_{SO}\sigma_x} U \quad (5.140)$$

is unitary, and $\det U_{SO} = 1$. Therefore it is diagonalized by some unitary transformation V to the form

$$V^\dagger U_{SO} V = \begin{pmatrix} e^{i\bar{\varphi}} & 0 \\ 0 & e^{-i\bar{\varphi}} \end{pmatrix}. \quad (5.141)$$

This implies

$$\begin{pmatrix} -2t_d \cos(k + \bar{\varphi}) & 0 \\ 0 & -2t_d \cos(k - \bar{\varphi}) \end{pmatrix} V^\dagger |\chi_k\rangle = \varepsilon_k V^\dagger |\chi_k\rangle, \quad (5.142)$$

and

$$\varepsilon_{k,+} = -2t_d \cos(k + \bar{\varphi}), \quad |\chi_{k,+}\rangle = V |\uparrow\rangle, \quad (5.143)$$

$$\varepsilon_{k,-} = -2t_d \cos(k - \bar{\varphi}), \quad |\chi_{k,-}\rangle = V |\downarrow\rangle. \quad (5.144)$$

For the possible values of $k = -\frac{2\pi}{3}, 0, \frac{2\pi}{3}$ we observe the following twofold degeneracy due to the TR symmetry

$$\varepsilon_{\pm \frac{2\pi}{3},+} = \varepsilon_{\mp \frac{2\pi}{3},-}, \quad (5.145)$$

$$\varepsilon_{0,+} = \varepsilon_{0,-}. \quad (5.146)$$

To evaluate the value of $\bar{\varphi}$ we consider the trace

$$\text{tr } U_{SO} = 2 \cos \bar{\varphi} \quad (5.147)$$

$$= \text{tr } e^{i\varphi_{SO}\sigma_x} e^{i\frac{\pi}{3}\sigma_z} = 2 \cos \varphi_{SO} \cos \frac{\pi}{3} = \cos \varphi_{SO}. \quad (5.148)$$

So it holds

$$2 \cos \bar{\varphi} = \cos \varphi_{SO}. \quad (5.149)$$

Varying φ_{SO} we can tune some of the degenerate levels close to zero energy. These states are potentially interesting for a peculiar spin-splitting after coupling the dots to the superconducting leads.

5.5.3 Evaluation of Det D

To diagonalize the matrix D given by (5.132), (5.133), we first assume that the on-site potentials W^i are all non-magnetic, that is

$$W^i = \varepsilon_i \tau_z. \quad (5.150)$$

Then

$$D'_N = \begin{pmatrix} 1 & 0 & 0 \\ 0 & U^\dagger & 0 \\ 0 & 0 & (U^\dagger)^2 \end{pmatrix} D_N \begin{pmatrix} 1 & 0 & 0 \\ 0 & U & 0 \\ 0 & 0 & U^2 \end{pmatrix}, \quad (5.151)$$

where

$$D'_N = \begin{pmatrix} z - \varepsilon_1 \tau_z & \tau_z t_d U_{SO} & \tau_z t_d U_{SO}^\dagger \\ \tau_z t_d U_{SO}^\dagger & z - \varepsilon_2 \tau_z & \tau_z t_d U_{SO} \\ \tau_z t_d U_{SO} & \tau_z t_d U_{SO}^\dagger & z - \varepsilon_3 \tau_z \end{pmatrix}, \quad (5.152)$$

and U_{SO} is given by (5.140). In turn, the matrix D_S remains invariant under this spin transformation.

Moreover, we know that the matrix U_{SO} is diagonalized by the unitary transformation (5.141). So we get

$$D''_N = V^\dagger D'_N V = \begin{pmatrix} z - \varepsilon_1 \tau_z & \tau_z t_d e^{i\sigma_z \bar{\varphi}} & \tau_z t_d e^{-i\sigma_z \bar{\varphi}} \\ \tau_z t_d e^{-i\sigma_z \bar{\varphi}} & z - \varepsilon_2 \tau_z & \tau_z t_d e^{i\sigma_z \bar{\varphi}} \\ \tau_z t_d e^{i\sigma_z \bar{\varphi}} & \tau_z t_d e^{-i\sigma_z \bar{\varphi}} & z - \varepsilon_3 \tau_z \end{pmatrix}. \quad (5.153)$$

Again, the matrix D_S remains invariant under the V -transformation.

Since D''_N is diagonal in the spin basis, we can replace σ_z by its eigenvalues $\nu = \pm$ and consider in the following $D^{(\nu)} = D''_N + D_S$ with

$$D^{(\nu)} = \begin{pmatrix} z - \varepsilon_1 \tau_z & \tau_z t_d e^{i\nu \bar{\varphi}} & \tau_z t_d e^{-i\nu \bar{\varphi}} \\ \tau_z t_d e^{-i\nu \bar{\varphi}} & z - \varepsilon_2 \tau_z & \tau_z t_d e^{i\nu \bar{\varphi}} \\ \tau_z t_d e^{i\nu \bar{\varphi}} & \tau_z t_d e^{-i\nu \bar{\varphi}} & z - \varepsilon_3 \tau_z \end{pmatrix}. \quad (5.154)$$

The equations $\det D^{(+)}(\{\varphi_i\}) = 0$ and $\det D^{(-)}(\{-\varphi_i\}) = 0$ give the same spectra, since $D^{(+)}(\{\varphi_i\}) = (D^{(-)}(\{-\varphi_i\}))^T$.

In the following we study $\det D^{(+)}$. Our next step is shifting the superconducting phase dependence from D_S to $D_N^{(+)}$ by means of the transformation (5.121)

$$\tilde{D}^{(+)} = \begin{pmatrix} z - \varepsilon_1 \tau_z & t_d \tau_z e^{-\frac{i}{2}(\varphi_1 - \varphi_2)\tau_z} e^{i\bar{\varphi}} & t_d \tau_z e^{-\frac{i}{2}(\varphi_1 - \varphi_3)\tau_z} e^{-i\bar{\varphi}} \\ t_d \tau_z e^{-\frac{i}{2}(\varphi_2 - \varphi_1)\tau_z} e^{-i\bar{\varphi}} & z - \varepsilon_2 \tau_z & t_d \tau_z e^{-\frac{i}{2}(\varphi_2 - \varphi_3)\tau_z} e^{i\bar{\varphi}} \\ t_d \tau_z e^{-\frac{i}{2}(\varphi_3 - \varphi_1)\tau_z} e^{i\bar{\varphi}} & t_d \tau_z e^{-\frac{i}{2}(\varphi_3 - \varphi_2)\tau_z} e^{-i\bar{\varphi}} & z - \varepsilon_3 \tau_z \end{pmatrix} + i\Gamma_S F(z) \mathbb{1}_3 \quad (5.155)$$

$$= \tilde{z} + \tau_z \begin{pmatrix} -\varepsilon_1 & t_d e^{\frac{i}{2}(\varphi_L + \varphi_R)\tau_z} e^{i\bar{\varphi}} & t_d e^{\frac{i}{2}\varphi_L \tau_z} e^{-i\bar{\varphi}} \\ t_d e^{-\frac{i}{2}(\varphi_L + \varphi_R)\tau_z} e^{-i\bar{\varphi}} & -\varepsilon_2 & t_d e^{-\frac{i}{2}\varphi_R \tau_z} e^{i\bar{\varphi}} \\ t_d e^{-\frac{i}{2}\varphi_L \tau_z} e^{i\bar{\varphi}} & t_d e^{\frac{i}{2}\varphi_R \tau_z} e^{-i\bar{\varphi}} & -\varepsilon_3 \end{pmatrix} - i\Gamma_S \frac{\frac{\Delta_0}{z} \tau_x}{\sqrt{1 - \frac{\Delta_0^2}{z^2}}}. \quad (5.156)$$

Hereby

$$\tilde{z} = z + \frac{i\Gamma_S}{\sqrt{1 - \frac{\Delta_0^2}{z^2}}}. \quad (5.157)$$

Taking into account the property (5.103) we represent

$$\det \tilde{D}^{(+)} = \det \tilde{D}^{(+),ee} \det \tilde{D}^{(+),hh} \det \left(1 - \frac{\Gamma_S^2 \Delta_0^2}{\Delta_0^2 - z^2} (\tilde{D}^{(+),ee})^{-1} (\tilde{D}^{(+),hh})^{-1} \right) \quad (5.158)$$

$$= \det \left(\tilde{D}^{(+),ee} \tilde{D}^{(+),hh} - \frac{\Gamma_S^2 \Delta_0^2}{\Delta_0^2 - z^2} \right), \quad (5.159)$$

with

$$\tilde{D}^{(+),ee} = \tilde{z} + \begin{pmatrix} -\varepsilon_1 & t_d e^{\frac{i}{2}(\varphi_L + \varphi_R)} e^{i\bar{\varphi}} & t_d e^{\frac{i}{2}\varphi_L} e^{-i\bar{\varphi}} \\ t_d e^{-\frac{i}{2}(\varphi_L + \varphi_R)} e^{-i\bar{\varphi}} & -\varepsilon_2 & t_d e^{-\frac{i}{2}\varphi_R} e^{i\bar{\varphi}} \\ t_d e^{-\frac{i}{2}\varphi_L} e^{i\bar{\varphi}} & t_d e^{\frac{i}{2}\varphi_R} e^{-i\bar{\varphi}} & -\varepsilon_3 \end{pmatrix} \quad (5.160)$$

and

$$\tilde{D}^{(+),hh} = \tilde{z} - \begin{pmatrix} -\varepsilon_1 & t_d e^{-\frac{i}{2}(\varphi_L + \varphi_R)} e^{i\bar{\varphi}} & t_d e^{-\frac{i}{2}\varphi_L} e^{-i\bar{\varphi}} \\ t_d e^{\frac{i}{2}(\varphi_L + \varphi_R)} e^{-i\bar{\varphi}} & -\varepsilon_2 & t_d e^{\frac{i}{2}\varphi_R} e^{i\bar{\varphi}} \\ t_d e^{\frac{i}{2}\varphi_L} e^{i\bar{\varphi}} & t_d e^{-\frac{i}{2}\varphi_R} e^{-i\bar{\varphi}} & -\varepsilon_3 \end{pmatrix}. \quad (5.161)$$

5.5.4 Bound state equation

To determine bound states at $|\omega| < \Delta_0$ it is necessary to solve the equation

$$\det \left(\tilde{D}^{(+),ee} \tilde{D}^{(+),hh} - \frac{\Gamma_S^2 \Delta_0^2}{\Delta_0^2 - \omega^2} \right) = 0, \quad (5.162)$$

(as well as the analogous equation with $\tilde{D}^{(-),ee}$ and $\tilde{D}^{(-),hh}$). In this evaluation we have to replace

$$\tilde{z} = \omega \left(1 + \frac{\Gamma_S}{\sqrt{\Delta_0^2 - \omega^2}} \right). \quad (5.163)$$

Low-energy analysis

We are mostly interested in solutions with $\omega = 0$, i.e. for which phases' configuration a crossing of the zero energy level is possible. Setting for simplicity $\varepsilon_i = 0$ and undoing the transformation (5.121), we derive the following equation

$$\det \left(\tilde{d}_0^{(+)} U_\varphi \tilde{d}_0^{(+)} U_\varphi^{-1} + \lambda^2 \right) = 0, \quad (5.164)$$

where

$$\tilde{d}_0^{(+)} = \begin{pmatrix} 0 & e^{i\bar{\varphi}} & e^{-i\bar{\varphi}} \\ e^{-i\bar{\varphi}} & 0 & e^{i\bar{\varphi}} \\ e^{i\bar{\varphi}} & e^{-i\bar{\varphi}} & 0 \end{pmatrix}, \quad U_\varphi = \begin{pmatrix} e^{i\varphi_1} & 0 & 0 \\ 0 & e^{i\varphi_2} & 0 \\ 0 & 0 & e^{i\varphi_3} \end{pmatrix}, \quad (5.165)$$

and the dimensionless parameter $\lambda = \frac{\Gamma_S}{t_d}$ is introduced.

Evaluating the determinant in (5.164), we obtain

$$f(\lambda^2) = \lambda^6 + 2C\lambda^4 + (2C + 3)\lambda^2 + 4\cos^2 3\bar{\varphi} = 0, \quad (5.166)$$

where

$$C = \cos \varphi_L + \cos \varphi_R + \cos(\varphi_L + \varphi_R). \quad (5.167)$$

Let us analyze under which conditions the function

$$f(x) = x^3 + 2Cx^2 + (2C + 3)x + 4\cos^2 3\bar{\varphi} \quad (5.168)$$

may have a positive root. Since $f(0) = 0$ and $f(-\infty) \rightarrow -\infty$, there is at least one negative root (say, $x_3 < 0$). The product of all three roots $x_1 x_2 x_3 = -4\cos^2 3\bar{\varphi} < 0$, therefore $x_1 x_2 > 0$. There are three possible cases: 1) two positive real roots; 2) two negative real roots; 3) a pair of complex valued roots. In order to have the favourable first scenario, it is necessary (and sufficient) to ensure that the function $f(x)$ possesses a local minimum at $x_{\min} > 0$ with $f(x_{\min}) < 0$.

Solving

$$f'(x) = 3x^2 + 4Cx + 2C + 3 = 0, \quad (5.169)$$

we find the extrema values

$$x_{1,2}^e = \frac{-2C \pm \sqrt{4C^2 - 6C - 9}}{3}, \quad (5.170)$$

and in particular

$$x_{\min} = \frac{-2C + \sqrt{4C^2 - 6C - 9}}{3}, \quad (5.171)$$

provided that $4C^2 - 6C - 9 > 0$. This condition along with the requirement $x_{\min} > 0$ imply that

$$C < -\frac{3}{4}(\sqrt{5} - 1) \approx -0.927. \quad (5.172)$$

Since the minimally possible value of C is -1.5 , the last condition can be satisfied (in particular, for $\varphi_L = \varphi_R$ the minimal value -1.5 is achieved at $\frac{2\pi}{3}$ and $\frac{4\pi}{3}$, while for $\varphi_L = -\varphi_R$ the minimal value -1 is achieved at $\varphi_L = \pi$).

Next, we apply the condition

$$f(x_{\min}) = -\frac{2}{9}(4C^2 - 6C - 9)x_{\min} - \frac{2}{9}C(2C + 3) + 4\cos^2 3\bar{\varphi} < 0. \quad (5.173)$$

Defining the function

$$g(C) = \frac{2}{9}(4C^2 - 6C - 9)x_{\min}(C) + \frac{2}{9}C(2C + 3), \quad (5.174)$$

we reformulate the above condition as

$$g(C) > 4\cos^2 3\bar{\varphi}. \quad (5.175)$$

Studying the behavior of $g(C)$ on the interval $[-1.5, -\frac{3}{4}(\sqrt{5} - 1)]$ we establish that 1) it is monotonically decaying; 2) it is positive for $C < -1$; 3) it takes values $g(-1.5) = 4$ and $g(-1) = 0$.

From this we can draw the following conclusions (note that in the gauge $\varphi_3 = 0$ we have the equality $\varphi_L = -\varphi_1$ and $\varphi_R = \varphi_2$).

- To satisfy the condition (5.175), the values of C must be in the range $[-1.5, -1)$.
- The condition (5.175) can not be satisfied at $\varphi_L = -\varphi_R$ (i.e $\varphi_1 = \varphi_2$) since the minimal value of C along this cut is -1 .

- The absence of SO coupling implies $\cos \bar{\varphi} = \frac{1}{2}$ (see Eq. (5.149)), and hence $\cos^2 3\bar{\varphi} = 1$. In this case we can satisfy (5.166) only by fine-tuning $\lambda^2 = 2$ and $C = -1.5$.
- Notice that $f(x, C_1) - f(x, C_2) = 2(C_1 - C_2)x(x + 1) > 0$ for $C_1 > C_2$ and $x > 0$. This means that the positive roots of $f(x, C_1)$ lie between the positive roots of $f(x, C_2)$. Increasing C we make the roots coalesce at some critical value. This consideration allows us to establish the phase diagram for the existence of zero-energy solution. It is given by the inequalities $0 < x < 3$ (whose limiting points are the roots of $f(x)$ at $C = -1.5$ and $4\cos^2 3\bar{\varphi} = 0$) and $4\cos^2 3\bar{\varphi} < 3x^2 - x^3$ (which is equivalent to $f(x, C = -1.5) = x^3 - 3x^2 + 4\cos^2 3\bar{\varphi} < 0$).

We also note that the condition $C < -1$ is equivalent to the geometric condition that the triangle between the points on a circle with the phases $\varphi_1, \varphi_2, \varphi_3$ encloses the circle's center. In particular, setting $\varphi_3 = 0$ one can check that the inequality

$$\cos \varphi_1 + \cos \varphi_2 + \cos(\varphi_1 - \varphi_2) < -1 \quad (5.176)$$

is fulfilled for $\{0 < \varphi_1 < \pi \cap \pi < \varphi_2 < \varphi_1 + \pi\} \cup \{\pi < \varphi_1 < 2\pi \cap \varphi_1 - \pi < \varphi_2 < \pi\}$. To see this, we rewrite the above inequality as

$$(1 + \cos \varphi_1)(1 + \cos \varphi_2) < -\sin \varphi_1 \sin \varphi_2 \iff 1 < -\tan \frac{\varphi_1}{2} \tan \frac{\varphi_2}{2}. \quad (5.177)$$

5.5.5 Continuum limit everywhere

In this case, the results from the PRB paper can be directly transferred. In particular, the boundary operators $t_m'^{\dagger} G_{1,1}^m t_m'$ related to the superconductors have the same form in the continuum limit as discussed in the paper. The other contributions to D require the knowledge of the semiconductor's Green's functions G^m . To this end, we need a model describing the semiconducting segments (especially focusing on modelling the spin-orbit coupling term).

Most generally, we can model the one-dimensional semiconductor Hamiltonian as

$$H = \frac{p^2}{2m} + \alpha p \vec{\sigma} \cdot \vec{n}_\alpha + B \vec{\sigma} \cdot \vec{n}_B + V, \quad (5.178)$$

where $\vec{n}_\alpha$ and $\vec{n}_B$ are two unit vectors defining the direction of the spin-orbit coupling and of the magnetic field, respectively, and V is a constant scalar potential. Assuming that

$$\vec{n}_\alpha \cdot \vec{n}_B = \cos \psi, \quad 0 \leq \psi \leq \pi, \quad (5.179)$$

we define the unit vector

$$\vec{n}_\perp = \frac{\vec{n}_B - \vec{n}_\alpha \cos \psi}{\sin \psi} \quad (5.180)$$

and rewrite

$$H = \frac{p^2}{2m} + \alpha(p + p_0) \vec{\sigma} \cdot \vec{n}_\alpha + \tilde{B} \vec{\sigma} \cdot \vec{n}_\perp + V, \quad (5.181)$$

where

$$p_0 = \frac{B \cos \psi}{\alpha}, \quad (5.182)$$

$$\tilde{B} = B \sin \psi. \quad (5.183)$$

By a global unitary transformation in spin space we can rotate H into

$$H = \frac{p^2}{2m} + \alpha(p + p_0)\sigma_z + \tilde{B}\sigma_x + V. \quad (5.184)$$

In the two cases $\psi = 0$ and $\psi = \frac{\pi}{2}$ a translationally invariant Green's function of this model can be analytically evaluated in explicit form.

In the absence of the B -field The Green's function of the Hamiltonian (5.184) on the finite interval $0 \leq x \leq L$ can be evaluate in the following way. First, we represent

$$\begin{aligned} H &= \frac{(p + m\alpha\sigma_x)^2}{2m} + V - \frac{m^2\alpha^2}{2} = \frac{(p + m\alpha\sigma_x)^2}{2m} + \tilde{V} \\ &= e^{-i\sigma_x m\alpha x} H_0 e^{i\sigma_x m\alpha x}, \end{aligned} \quad (5.185)$$

$$H_0 = \frac{p^2}{2m} + \tilde{V}, \quad \tilde{V} = V - \frac{m^2\alpha^2}{2}. \quad (5.186)$$

Hence

$$G(x, x') = e^{-i\sigma_x m\alpha x} G_0(x, x') e^{i\sigma_x m\alpha x'}, \quad (5.187)$$

where $G_0(x, x')$ is the Green's function corresponding to the Hamiltonian H_0 (note that the above transformation does not change the vanishing boundary conditions).

Second, we establish $G_0(x, x')$. Defining $\bar{z} = z - \tilde{V}$, we apply the method of the boundary Green's function (or, alternatively, the Lehmann representation) to derive the following result

$$G_0(x, x') = \frac{m \cos[k_0(L - |x - x'|)] - \cos[k_0(L - x - x')]}{k_0 \sin k_0 L}, \quad k_0 = \sqrt{2m\bar{z}}. \quad (5.188)$$

The validity of this result can be verified by checking: 1) the boundary values $G_0(0, x') = G_0(L, x') = 0$; 2) the fulfilment of the defining equation $(\bar{z} - \frac{p^2}{2m})G_0(x, x') = \delta(x - x')$; 3) the jump value of the first derivative $G_1(x' + 0^+, x') - G_1(x' + 0^-, x') = 2m$. Note that the equation $\sin k_0 L = 0$ determines the eigenenergies of this auxiliary model.

Note that this Green's function describes particles only. By definition we have

$$G_0^{(+)}(x, x') = -\frac{2}{L} \sum_n \frac{\sin(\frac{\pi n x}{L}) \sin(\frac{\pi n x'}{L})}{\frac{1}{2m} \left(\frac{\pi n}{L}\right)^2 - (z + \tilde{V})}, \quad (5.189)$$

$$G_0^{(-)}(x, x') = \frac{2}{L} \sum_n \frac{\sin(\frac{\pi n x}{L}) \sin(\frac{\pi n x'}{L})}{\frac{1}{2m} \left(\frac{\pi n}{L}\right)^2 - (\tilde{V} - z)}, \quad (5.190)$$

$$G_0^{(-)}(x, x') = -G_0^{(+)}(x, x') \Big|_{\bar{z}=\tilde{V}-z}, \quad (5.191)$$

such that

$$G_0^{(\pm)}(x, x') = \pm \frac{m \cos[k_{\pm}(L - |x - x'|)] - \cos[k_{\pm}(L - x - x')]}{\sin k_{\pm} L}, \quad k_{\pm} = \sqrt{2m(\tilde{V} \pm z)} \quad (5.192)$$

It is necessary to remark that we measure energies from the value μ of the chemical potential in the leads. This means that effectively we should write

$$\bar{z} = z + \mu - \tilde{V}. \quad (5.193)$$

Since $\mu \gg |\omega|, \tilde{V}$, we can approximate

$$k_0(z) = \sqrt{2m\tilde{z}} = \sqrt{2m(z + \mu + \varepsilon_N - \tilde{V} - \varepsilon_N)} \quad (5.194)$$

$$= k_N \sqrt{1 + \frac{z - \varepsilon_N}{\mu + \varepsilon_N - \tilde{V}}} \approx k_N + m \frac{z - \varepsilon_N}{k_N} \approx k_N + \frac{z - \varepsilon_N}{v_F}, \quad (5.195)$$

where

$$k_N = \frac{2\pi N}{3L} = \sqrt{2m(\mu + \varepsilon_N - \tilde{V})} \quad (5.196)$$

is given by the large integer N , defining the eigenstate at the energy ε_N above μ , and $v_F = \sqrt{2m\mu}/m$ (see below for the quantization condition on k_N).

In the following we have to approximate (questionable in view of the new condition on k_N)

$$\sin k_0 L \approx \sin \left(k_N L + \frac{z - \varepsilon_N}{v_F} L \right) = (-1)^N \sin \frac{z - \varepsilon_N}{v_F} L, \quad (5.197)$$

$$\cos k_0 L \approx \cos \left(k_N L + \frac{z - \varepsilon_N}{v_F} L \right) = (-1)^N \cos \frac{z - \varepsilon_N}{v_F} L. \quad (5.198)$$

Assuming a short junction $L \ll \frac{v_F}{\Delta}$, we further approximate

$$\sin k_0 L \approx (-1)^N \frac{z - \varepsilon_N}{v_F} L, \quad (5.199)$$

$$\cos k_0 L \approx (-1)^N. \quad (5.200)$$

Next, we evaluate

$$-\frac{1}{8m^2} \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G(x, x) = -\frac{1}{8m^2} \lim_{x \rightarrow 0^+} \frac{d^2}{dx^2} G_0(x, x) = -\frac{k_0}{2m} \cot k_0 L, \quad (5.201)$$

$$-\frac{1}{8m^2} \lim_{x \rightarrow L^-} \frac{d^2}{dx^2} G(x, x) = -\frac{1}{8m^2} \lim_{x \rightarrow L^-} \frac{d^2}{dx^2} G_0(x, x) = -\frac{k_0}{2m} \cot k_0 L, \quad (5.202)$$

$$\frac{1}{4m^2} \lim_{x \rightarrow 0^+} \lim_{x' \rightarrow L^-} G_{12}(x, x') = \frac{e^{i\sigma_x m \alpha L}}{4m^2} \lim_{x \rightarrow 0^+} \lim_{x' \rightarrow L^-} G_{0,12}(x, x') = \frac{k_0}{2m} \frac{e^{i\sigma_x m \alpha L}}{\sin k_0 L}, \quad (5.203)$$

$$\frac{1}{4m^2} \lim_{x \rightarrow L^-} \lim_{x' \rightarrow 0^+} G_{12}(x, x') = \frac{-e^{i\sigma_x m \alpha L}}{4m^2} \lim_{x \rightarrow L^-} \lim_{x' \rightarrow 0^+} G_{0,12}(x, x') = \frac{k_0}{2m} \frac{e^{-i\sigma_x m \alpha L}}{\sin k_0 L}. \quad (5.204)$$

With the help of these expressions we construct the continuum limit

$$d = \lim_{a \rightarrow 0} aD = \begin{pmatrix} -\mathcal{U}_1 - \tau_z \frac{k_0}{2m_N} \mathbf{2} \cot k_0 L & \tau_z \frac{k_0}{2m_N} \frac{e^{i\varphi_{SO}\sigma x}}{\sin k_0 L} & \tau_z \frac{k_0}{2m_N} U^2 \frac{e^{-i\varphi_{SO}\sigma x}}{\sin k_0 L} (U^\dagger)^2 \\ \tau_z \frac{k_0}{2m_N} \frac{e^{-i\varphi_{SO}\sigma x}}{\sin k_0 L} & -\mathcal{U}_2 - \tau_z \frac{k_0}{2m_N} \mathbf{2} \cot k_0 L & \tau_z \frac{k_0}{2m_N} U \frac{e^{i\varphi_{SO}\sigma x}}{\sin k_0 L} U^\dagger \\ \tau_z \frac{k_0}{2m_N} U^2 \frac{e^{i\varphi_{SO}\sigma x}}{\sin k_0 L} (U^\dagger)^2 & \tau_z \frac{k_0}{2m_N} U \frac{e^{-i\varphi_{SO}\sigma x}}{\sin k_0 L} U^\dagger & -\mathcal{U}_3 - \tau_z \frac{k_0}{2m_N} \mathbf{2} \cot k_0 L \end{pmatrix} \quad (5.205)$$

$$+ i \frac{v_{F,S}}{2} \begin{pmatrix} e^{\frac{i}{2}\varphi_1 \tau_z} F(z) e^{-\frac{i}{2}\varphi_1 \tau_z} & 0 & 0 \\ 0 & e^{\frac{i}{2}\varphi_2 \tau_z} F(z) e^{-\frac{i}{2}\varphi_2 \tau_z} & 0 \\ 0 & 0 & e^{\frac{i}{2}\varphi_3 \tau_z} F(z) e^{-\frac{i}{2}\varphi_3 \tau_z} \end{pmatrix}, \quad (5.206)$$

where m_N is an effective electron mass in the semiconductor region, and $v_{F,S}$ is the Fermi velocity in the superconducting leads. The dependence k_0 on z is sensitive to whether we are in the particle or in the hole sector, that is $k_0 = k_0(\tau_z z)$.

We see that the obtained matrix d has the same structure as the matrix D in (5.132), (5.133) of the three-dot model, i.e. the former is obtained from the latter by the replacements "explicit" $z \rightarrow 0$ and

$$W^i \rightarrow \mathcal{U}_i + \tau_z \frac{k_0}{m_N} \cot k_0 L, \quad (5.207)$$

$$t_d \rightarrow \frac{k_0}{2m_N} \frac{1}{\sin k_0 L}, \quad (5.208)$$

$$\Gamma_S \rightarrow \frac{v_{F,S}}{2}. \quad (5.209)$$

It is instructive to solve a model with three semiconducting segments neglecting both the spin-orbit interaction and the coupling to superconducting leads. Its spectrum is given by the equation

$$0 = \det \begin{pmatrix} -2 \cos k_0 L & 1 & 1 \\ 1 & -2 \cos k_0 L & 1 \\ 1 & 1 & -2 \cos k_0 L \end{pmatrix} \quad (5.210)$$

$$= -8 \cos^3 k_0 L + 6 \cos k_0 L + 2 = -2 \cos 3k_0 L + 2. \quad (5.211)$$

This equation gives the spectrum

$$k_0 = \frac{2\pi n}{3L}, \quad n \in \mathbb{Z} \quad (5.212)$$

of the ring with the circumference $3L$ and periodic boundary conditions.

5.6 Necessary condition for existence of zero-energy solutions

Looking at (5.132) and (5.133) we realize that $D_N + D_S$ can be generally represented as

$$D = \tilde{\omega} - \tau_z h_{eff} - \tilde{\Gamma}_S e^{\frac{i}{2}\hat{\varphi}\tau_z} \tau_x e^{-\frac{i}{2}\hat{\varphi}\tau_z} \quad (5.213)$$

$$= e^{\frac{i}{2}\hat{\varphi}\tau_z} \left(\tilde{\omega} - \tau_z e^{-\frac{i}{2}\hat{\varphi}\tau_z} h_{eff} e^{\frac{i}{2}\hat{\varphi}\tau_z} - \tilde{\Gamma}_S \tau_x \right) e^{-\frac{i}{2}\hat{\varphi}\tau_z}, \quad (5.214)$$

where we introduced $\hat{\varphi} = \text{diag}(\varphi_1, \varphi_2, \varphi_3)$ and

$$\tilde{\omega} = \omega \left(1 + \frac{\Gamma_S}{\sqrt{\Delta_0^2 - \omega^2}} \right), \quad 1 + \frac{\Gamma_S}{\sqrt{\Delta_0^2 - \omega^2}} > 1, \quad (5.215)$$

$$\tilde{\Gamma}_S = \Gamma_S \frac{\Delta_0}{\sqrt{\Delta_0^2 - \omega^2}} > 0, \quad (5.216)$$

assuming $|\omega| < \Delta_0$.

The effective particle Hamiltonian h_{eff} has the TR-symmetry

$$h_{eff} = (i\sigma_y) h_{eff}^* (-i\sigma_y) = (i\sigma_y) h_{eff}^T (-i\sigma_y). \quad (5.217)$$

The bound-state equation $\det D = 0$ can be equivalently rewritten as

$$\begin{aligned} & \det \left[\left(\tilde{\omega} - e^{-\frac{i}{2}\hat{\varphi}} h_{eff} e^{\frac{i}{2}\hat{\varphi}} \right) \left(\tilde{\omega} + e^{\frac{i}{2}\hat{\varphi}} h_{eff} e^{-\frac{i}{2}\hat{\varphi}} \right) - \tilde{\Gamma}_S^2 \right] \\ &= \det \left[(\tilde{\omega} - h_{eff}) U_\varphi (\tilde{\omega} + h_{eff}) U_\varphi^{-1} - \tilde{\Gamma}_S^2 \right] = 0, \end{aligned} \quad (5.218)$$

where $U_\varphi = e^{i\hat{\varphi}}$.

Instead of studying a lower bound on either the ABS energy $|\omega|$ or its counterpart $|\tilde{\omega}|$ (note that by the virtue of (5.215) $\tilde{\omega}$ vanishes if and only if ω vanishes), we set $\omega = 0$ (which implies $\tilde{\omega} = 0$ and $\tilde{\Gamma}_S = \Gamma_S$) and study for which configurations of the phases φ_l the equation

$$\det [h_{eff}U_\varphi h_{eff} + \Gamma_S^2 U_\varphi] = 0 \quad (5.219)$$

may have a solution.

First we note that at $\varphi_l = 0$ for $\forall l \in \{1, 2, 3\}$ this is impossible: The matrix $h_{eff}^2 + \Gamma_S^2$ is positive definite, and therefore all its eigenvalue are strictly greater than zero, yielding $\det [h_{eff}^2 + \Gamma_S^2] > 0$.

Let us now introduce hermitian matrices

$$Q_l = \sum_{\sigma} (h_{eff}|l, \sigma\rangle\langle l, \sigma| h_{eff} + \Gamma_S^2 |l, \sigma\rangle\langle l, \sigma|), \quad (5.220)$$

where $l = 1, 2, 3$ is a site index and σ is a spin index. These matrices are positive semi-definite, since for any state $|\psi\rangle$ with a nonzero norm it holds

$$\langle\psi|Q_l|\psi\rangle = \sum_{\sigma} (|\langle l, \sigma|h_{eff}|\psi\rangle|^2 + \Gamma_S^2 |\langle l, \sigma|\psi\rangle|^2) \geq 0. \quad (5.221)$$

The matrix $\sum_l Q_l = h_{eff}^2 + \Gamma_S^2$ is positive definite though, as it was mentioned above.

Rewriting the equation (5.219) in terms of Q_l yields

$$\det \left[\sum_l Q_l e^{i\varphi_l} \right] = 0. \quad (5.222)$$

Let us study the eigenvalues of the non-hermitian matrix $\sum_l Q_l e^{i\varphi_l}$ and try to establish a lower bound on its absolute values. We set up the eigenvalue equation

$$\sum_l Q_l e^{i\varphi_l} |\chi\rangle = \lambda |\chi\rangle, \quad (5.223)$$

where λ is in general complex-valued. Multiplying it with $\langle\chi|$ (note that it is not a left eigenvector of $\sum_l Q_l e^{i\varphi_l}$!), we obtain

$$\sum_l \tilde{w}_l e^{i\varphi_l} = \lambda \langle\chi|\chi\rangle, \quad (5.224)$$

where

$$\tilde{w}_l = \langle\chi|Q_l|\chi\rangle \geq 0. \quad (5.225)$$

In turn,

$$N = \sum_l \tilde{w}_l = \langle\chi|\sum_l Q_l|\chi\rangle > 0. \quad (5.226)$$

Defining $w_l = \frac{\tilde{w}_l}{N}$ such that $\sum_l w_l = 1$, we obtain

$$\left| \sum_l w_l e^{i\varphi_l} \right| = |\lambda| \frac{\langle\chi|\chi\rangle}{N}. \quad (5.227)$$

If the triangle inscribed into the unit circle with the vertices given by the phases $\{\varphi_l\}$ does not surround the circle's centre, then $|\lambda|$ has a positive lower bound. In this case, there is no zero eigenvalue λ , and the equation (5.219) is not satisfied for any parameters of h_{eff} .

In order to have a zero-energy solution, it is necessary for the triangle to surround the circle's centre.

5.7 Generalization from equilateral to isosceles triangular geometry of contacts

Let us focus on the model formulated in the subsection 5.5.5. Introducing the dimensionless counterpart $\bar{d} = \frac{2}{v_{F,S}}d$ of d defined by (5.205), (5.206), we cast it to the compact form

$$\bar{d} = -\tau_z \bar{h}(\tau_z z) + e^{\frac{i}{2}\tau_z \hat{\varphi}} i F(z) e^{-\frac{i}{2}\tau_z \hat{\varphi}}. \quad (5.228)$$

In this representation we use

$$\bar{h}(z) = \begin{pmatrix} \eta_1 & -\bar{t}_d e^{i\varphi_{SO}\sigma_x} & -\bar{t}_d U^\dagger e^{-i\varphi_{SO}\sigma_x} U \\ -\bar{t}_d e^{-i\varphi_{SO}\sigma_x} & \eta_2 & -\bar{t}_d U e^{i\varphi_{SO}\sigma_x} U^\dagger \\ -\bar{t}_d U^\dagger e^{i\varphi_{SO}\sigma_x} U & -\bar{t}_d U e^{-i\varphi_{SO}\sigma_x} U^\dagger & \eta_3 \end{pmatrix}, \quad (5.229)$$

where $U = e^{i\frac{\pi}{3}\sigma_z}$, and

$$\eta_i = \frac{2}{v_{F,S}}V_i + \frac{k_0}{v_{F,S}m_N} \cot k_0 L, \quad (5.230)$$

$$\bar{t}_d = \frac{k_0}{v_{F,S}m_N} \frac{1}{\sin k_0 L}, \quad (5.231)$$

with the tunnel barrier strength V_i is taken from $\mathcal{U}_i = \tau_z V_i$. The parameters η_i and $\bar{t}_d$ depend on frequency via k_0 . Recall that for $|\omega| < \Delta_0$ it holds $iF(z) = \frac{\omega - \Delta_0 \tau_x}{\sqrt{\Delta_0^2 - \omega^2}}$.

To generalize the model given by (5.229) we assume that the geometry between the contacts 1,2, and 3 is not equilateral, but rather isosceles. In particular, we assume that the distance between the contacts 1 and 2 is L , while the distance from the contact 3 to either of the contacts 1 and 2 is $L' \neq L$. Then we achieve the following generalization of (5.229):

$$\bar{h}(z) = \begin{pmatrix} \eta_1 & -\bar{t}_d e^{i\varphi_{SO}\sigma_x} & -\bar{t}'_d U_\alpha^\dagger e^{-i\varphi'_{SO}\sigma_x} U_\alpha \\ -\bar{t}_d e^{-i\varphi_{SO}\sigma_x} & \eta_2 & -\bar{t}'_d U_\alpha e^{i\varphi'_{SO}\sigma_x} U_\alpha^\dagger \\ -\bar{t}'_d U_\alpha^\dagger e^{i\varphi'_{SO}\sigma_x} U_\alpha & -\bar{t}'_d U_\alpha e^{-i\varphi'_{SO}\sigma_x} U_\alpha^\dagger & \eta_3 \end{pmatrix}, \quad (5.232)$$

Hereby $U_\alpha = e^{\frac{i}{2}(\pi-\alpha)\sigma_z}$, with the acute angle α at the contacts 1 and 2 given by

$$\cos \alpha = \frac{L}{2L'}. \quad (5.233)$$

The new parameters are given by

$$\eta_{1,2} = \frac{2}{v_{F,S}}V_{1,2} + \frac{k_0}{v_{F,S}m_N} (\cot k_0 L + \cot k_0 L'), \quad (5.234)$$

$$\eta_3 = \frac{2}{v_{F,S}}V_3 + \frac{2k_0}{v_{F,S}m_N} \cot k_0 L', \quad (5.235)$$

$$\bar{t}_d = \frac{k_0}{v_{F,S}m_N} \frac{1}{\sin k_0 L}, \quad (5.236)$$

$$\bar{t}'_d = \frac{k_0}{v_{F,S}m_N} \frac{1}{\sin k_0 L'}, \quad (5.237)$$

and $\varphi_{SO} = k_{SO}L$, $\varphi'_{SO} = k_{SO}L'$, $k_{SO} = m_N \alpha_{SO}$.

The bound-state equation for this model reads

$$\det \left[\left(\frac{\omega}{\sqrt{\Delta_0^2 - \omega^2}} - \bar{h}(\omega) \right) e^{i\hat{\varphi}} \left(\frac{\omega}{\sqrt{\Delta_0^2 - \omega^2}} + \bar{h}(-\omega) \right) e^{-i\hat{\varphi}} - \frac{\Delta_0^2}{\Delta_0^2 - \omega^2} \right] = 0. \quad (5.238)$$

For short junctions, there is a standard approximation $\bar{h}(\omega) \approx \bar{h}(-\omega) \approx \bar{h}(0) \equiv \bar{h}$.

5.7.1 Relation to the scattering matrix approach

In the scattering matrix approach the bound state equation has the following form

$$\det(1 - s_A s_N) = 0. \quad (5.239)$$

The normal scattering matrix for the present model has the form

$$s_N = \begin{pmatrix} s & 0 \\ 0 & s^\dagger \end{pmatrix}, \quad s = \frac{\bar{h} + i}{-\bar{h} + i}. \quad (5.240)$$

The matrix s_A encodes the Andreev reflections

$$s_A = -ie^{i\chi} \begin{pmatrix} 0 & e^{i\varphi} \\ e^{-i\varphi} & 0 \end{pmatrix}, \quad \chi = \arctan \frac{\omega}{\sqrt{\Delta_0^2 - \omega^2}}. \quad (5.241)$$

Rewriting (5.239) in the form

$$\det \left(1 + e^{2i\chi} e^{i\varphi} \frac{-\bar{h} + i}{\bar{h} + i} e^{-i\varphi} \frac{\bar{h} + i}{-\bar{h} + i} \right) = 0, \quad (5.242)$$

we aim to show its equivalence with (5.238). Bringing (5.242) to the form

$$\det [e^{-2i\chi} e^{i\varphi} (i + \bar{h}) e^{-i\varphi} (i - \bar{h}) + e^{i\varphi} (i - \bar{h}) e^{-i\varphi} (i + \bar{h})] = 0 \quad (5.243)$$

and collecting the terms of the same order in $\bar{h}$,

$$\det [(e^{-2i\chi} + 1)(-1 - e^{i\varphi} \bar{h} e^{-i\varphi} \bar{h}) - i(e^{-2i\chi} - 1)(\bar{h} - e^{i\varphi} \bar{h} e^{-i\varphi})] = 0, \quad (5.244)$$

$$\det [-1 - e^{i\varphi} \bar{h} e^{-i\varphi} \bar{h} - \tan \chi (\bar{h} - e^{i\varphi} \bar{h} e^{-i\varphi})] = 0, \quad (5.245)$$

we indeed see that the last expression is equivalent to (5.238).

5.8 Microwave drive

We add time-dependent variations $\varphi_m \rightarrow \varphi_m + 2\delta\varphi_m(t)$ to the superconducting phases φ_m (thereby transforming $h_m \rightarrow h_m(t)$) as well as time-dependent potentials to the Hamiltonians $h_m(t) \rightarrow h_m(t) - \delta\dot{\varphi}_m(t)\tau_z$.

The gauge transformations $U_m(t) = e^{i\tau_z\delta\varphi_m(t)}$ remove the $\delta\varphi_m(t)$ dependence from the lead Hamiltonians,

$$\tilde{h}_m = U_m^\dagger(t)[h_m(t) - \delta\dot{\varphi}_m(t)\tau_z]U_m(t) - iU_m^\dagger(t)\partial_t U_m(t) = h_m, \quad (5.246)$$

shifting it into the tunneling terms

$$\tilde{t}'_m(t) = U_m^\dagger(t)\hat{t}'_m, \quad (5.247)$$

$$\tilde{t}'_m{}^\dagger(t) = \hat{t}'_m{}^\dagger U_m(t). \quad (5.248)$$

For sufficiently small drive amplitudes $|\delta\varphi_m(t)| \ll 2\pi$ we may expand

$$\tilde{t}'_m(t) \approx (1 - i\tau_z\delta\varphi_m(t))\hat{t}'_m = \hat{t}'_m - i\tau_z\hat{t}'_m\delta\varphi_m(t), \quad (5.249)$$

$$\tilde{t}'_m{}^\dagger(t) \approx \hat{t}'_m{}^\dagger(1 + i\tau_z\delta\varphi_m(t)) = \hat{t}'_m{}^\dagger + i\hat{t}'_m{}^\dagger\tau_z\delta\varphi_m(t). \quad (5.250)$$

This expansion results in the linear drive terms added to the initial unperturbed Hamiltonian

$$\tilde{H}(t) = H - \frac{\hbar}{e} \sum_{m'} \hat{J}_{m'} \delta\varphi_{m'}(t), \quad (5.251)$$

where $\hat{J}_m$ is the current operator defined in (5.78). In particular, in the Nambu representation it has the form

$$\hat{J}_m = \frac{1}{2}\Psi_{1_m}^\dagger \left(-\frac{ie}{\hbar}\tau_z\hat{t}_m' \right) \Psi_{0_m} + \frac{1}{2}\Psi_{0_m}^\dagger \left(\frac{ie}{\hbar}\hat{t}_m'\tau_z \right) \Psi_{1_m} = \frac{1}{2}\Psi_\alpha^\dagger J_m^{\alpha\beta} \Psi_\beta, \quad (5.252)$$

where the multi-indices α and β contain both spatial and Nambu indices.

The linear current response by the Kubo formula equals

$$\langle \hat{J}_m(t) \rangle = J_m + 2 \sum_{m'} \frac{\partial J_m}{\partial \varphi_{m'}} \delta \varphi_{m'}(t) \quad (5.253)$$

$$+ \frac{i}{e} \int_{-\infty}^t dt' \langle [\hat{J}_m(t-t'), \hat{J}_{m'}]_0 \rangle \delta \varphi_{m'}(t'), \quad (5.254)$$

where $J_m = \langle \hat{J}_m \rangle_0$ is the equilibrium Josephson current, the second term $\propto \delta \varphi_{m'}(t)$ arises due to the correction to the observable (current operator), and the last term is the time nonlocal response due to the variation of the system's density matrix. In other words, the last term is a dissipative contribution occurring by the virtue of transitions between eigenstates. For $\delta \varphi_{m'}(t') = \delta \varphi_{m'} e^{-i(\omega+i0^+)t'}$ it can be written as

$$J_m^{\text{diss}}(t) = \sum_{m'} \chi_{m,m'}^r(\omega) \delta \varphi_{m'}(t), \quad (5.255)$$

where

$$\chi_{m,m'}^r(\omega) = -\frac{\hbar}{e} G_{\hat{J}_m, \hat{J}_{m'}}^r(\omega) \quad (5.256)$$

is the retarded response function given in terms of the retarded Green's functions of the two current operators

$$G_{\hat{J}_m, \hat{J}_{m'}}^r(\omega) = \frac{1}{\hbar} \int dt G_{\hat{J}_m, \hat{J}_{m'}}^r(t, 0) e^{i(\omega+i0^+)t}, \quad (5.257)$$

$$G_{\hat{J}_m, \hat{J}_{m'}}^r(t, 0) = -i\Theta(t) \langle [\hat{J}_m(t), \hat{J}_{m'}(0)] \rangle_0 \quad (5.258)$$

$$= \Theta(t) \left(G_{\hat{J}_m, \hat{J}_{m'}}^>(t, 0) - G_{\hat{J}_m, \hat{J}_{m'}}^<(t, 0) \right), \quad (5.259)$$

$$G_{\hat{J}_m, \hat{J}_{m'}}^>(t, 0) = -i \langle \hat{J}_m(t) \hat{J}_{m'}(0) \rangle_0, \quad (5.260)$$

$$G_{\hat{J}_m, \hat{J}_{m'}}^<(t, 0) = -i \langle \hat{J}_{m'}(0) \hat{J}_m(t) \rangle_0. \quad (5.261)$$

The greater $G^>$ and the lesser $G^<$ Green's functions introduced above are most conveniently computed on the Keldysh contour. In particular, the contour-ordered Green's function of two arbitrary bosonic operators A and B

$$G_{AB}^{\lambda\lambda'}(t, t') = -i \langle T_c A(\lambda t) B(\lambda' t') \rangle, \quad (5.262)$$

where $\lambda, \lambda' = 1, 2$ denote the contour's branches of forward (1) and backward (2) propagation, encodes the time-ordered, anti-time-ordered, lesser and greater functions

$$G_{AB}^t(t, t') = G_{AB}^{11}(t, t') = -i \langle T A(t) B(t') \rangle \quad (5.263)$$

$$= -i\Theta(t-t') \langle A(t) B(t') \rangle - i\Theta(t'-t) \langle B(t') A(t) \rangle, \quad (5.264)$$

$$G_{AB}^{\tilde{t}}(t, t') = G_{AB}^{22}(t, t') = -i \langle \tilde{T} A(t) B(t') \rangle \quad (5.265)$$

$$= -i\Theta(t'-t) \langle A(t) B(t') \rangle - i\Theta(t-t') \langle B(t') A(t) \rangle, \quad (5.266)$$

$$G_{AB}^<(t, t') = G_{AB}^{12}(t, t') = -i \langle B(t') A(t) \rangle, \quad (5.267)$$

$$G_{AB}^>(t, t') = G_{AB}^{21}(t, t') = -i \langle A(t) B(t') \rangle. \quad (5.268)$$

Defining

$$A(t) = \frac{1}{2} \Psi_{\alpha}^{\dagger}(t) J_m^{\alpha\beta} \Psi_{\beta}(t), \quad (5.269)$$

$$B(t') = \frac{1}{2} \Psi_{\alpha'}^{\dagger}(t') J_{m'}^{\alpha'\beta'} \Psi_{\beta'}(t'), \quad (5.270)$$

we apply the Wick's theorem to the function (5.262):

$$G_{\hat{j}_m, \hat{j}_{m'}}^{\lambda\lambda'}(t, t') = -\frac{i}{4} J_m^{\alpha\beta} J_{m'}^{\alpha'\beta'} \langle T_c \Psi_{\alpha}^{\dagger}(\lambda t) \Psi_{\beta}(\lambda t) \Psi_{\alpha'}^{\dagger}(\lambda' t') \Psi_{\beta'}(\lambda' t') \rangle \quad (5.271)$$

$$\begin{aligned} &= -\frac{i}{4} J_m^{\alpha\beta} J_{m'}^{\alpha'\beta'} \\ &\times \left(\langle T_c \Psi_{\alpha}^{\dagger}(\lambda t) \Psi_{\beta}(\lambda t) \rangle \langle T_c \Psi_{\alpha'}^{\dagger}(\lambda' t') \Psi_{\beta'}(\lambda' t') \rangle \right. \\ &- \langle T_c \Psi_{\alpha}^{\dagger}(\lambda t) \Psi_{\alpha'}^{\dagger}(\lambda' t') \rangle \langle T_c \Psi_{\beta}(\lambda t) \Psi_{\beta'}(\lambda' t') \rangle \\ &\left. + \langle T_c \Psi_{\alpha}^{\dagger}(\lambda t) \Psi_{\beta'}(\lambda' t') \rangle \langle T_c \Psi_{\beta}(\lambda t) \Psi_{\alpha'}^{\dagger}(\lambda' t') \rangle \right). \end{aligned} \quad (5.272)$$

Then

$$G_{\hat{j}_m, \hat{j}_{m'}}^r(t, 0) = -\frac{i}{4} J_m^{\alpha\beta} J_{m'}^{\alpha'\beta'} \Theta(t) \quad (5.273)$$

$$\times \left(-\langle \Psi_{\alpha}^{\dagger}(t) \Psi_{\alpha'}^{\dagger}(0) \rangle \langle \Psi_{\beta}(t) \Psi_{\beta'}(0) \rangle + \langle \Psi_{\alpha}^{\dagger}(t) \Psi_{\beta'}(0) \rangle \langle \Psi_{\beta}(t) \Psi_{\alpha'}^{\dagger}(0) \rangle \right. \quad (5.274)$$

$$\left. + \langle \Psi_{\alpha'}^{\dagger}(0) \Psi_{\alpha}^{\dagger}(t) \rangle \langle \Psi_{\beta'}(0) \Psi_{\beta}(t) \rangle - \langle \Psi_{\beta'}(0) \Psi_{\alpha}^{\dagger}(t) \rangle \langle \Psi_{\alpha'}^{\dagger}(0) \Psi_{\beta}(t) \rangle \right). \quad (5.275)$$

Using

$$\Psi_{\beta'}(0) = C_{\beta'\beta''} \Psi_{\beta''}^{\dagger}(0), \quad (5.276)$$

$$\Psi_{\alpha'}^{\dagger}(0) = C_{\alpha''\alpha'} \Psi_{\alpha''}^{\dagger}(0), \quad (5.277)$$

with $C = \tau_y \sigma_y$, as well as

$$C_{\alpha''\alpha'} J_{m'}^{\alpha'\beta'} C_{\beta'\beta''} = -J_{m'}^{\beta''\alpha''}, \quad (5.278)$$

we obtain

$$\begin{aligned} G_{\hat{j}_m, \hat{j}_{m'}}^r(t, 0) &= -\frac{i}{2} J_m^{\alpha\beta} J_{m'}^{\alpha'\beta'} \Theta(t) \\ &\times \left(\langle \Psi_{\alpha}^{\dagger}(t) \Psi_{\beta'}(0) \rangle \langle \Psi_{\beta}(t) \Psi_{\alpha'}^{\dagger}(0) \rangle - \langle \Psi_{\beta'}(0) \Psi_{\alpha}^{\dagger}(t) \rangle \langle \Psi_{\alpha'}^{\dagger}(0) \Psi_{\beta}(t) \rangle \right). \end{aligned} \quad (5.279)$$

In terms of the retarded, advanced, and Keldysh components of the fermionic Green's functions in the Nambu representation

$$G_{\alpha\beta}^r(t, 0) = -i\Theta(t) \langle \{ \Psi_{\alpha}(t), \Psi_{\beta}^{\dagger}(0) \} \rangle, \quad (5.280)$$

$$G_{\alpha\beta}^a(t, 0) = i\Theta(-t) \langle \{ \Psi_{\alpha}(t), \Psi_{\beta}^{\dagger}(0) \} \rangle, \quad (5.281)$$

$$G_{\alpha\beta}^K(t, 0) = -i \langle [\Psi_{\alpha}(t), \Psi_{\beta}^{\dagger}(0)] \rangle, \quad (5.282)$$

we find the expression

$$G_{\hat{j}_m, \hat{j}_{m'}}^r(t, 0) = -\frac{i}{4} J_m^{\alpha\beta} J_{m'}^{\alpha'\beta'} \left(G_{\beta\alpha'}^r(t, 0) G_{\beta'\alpha}^K(0, t) + G_{\beta\alpha'}^K(t, 0) G_{\beta'\alpha}^a(0, t) \right). \quad (5.283)$$

Its Fourier transform equals

$$G_{\hat{j}_m, \hat{j}_{m'}}^r(\omega) = -\frac{i}{4} J_m^{\alpha\beta} J_{m'}^{\alpha'\beta'} \hbar \int \frac{d\omega'}{2\pi} [G_{\beta\alpha'}^r(\omega + \omega') G_{\beta'\alpha}^K(\omega') + G_{\beta\alpha'}^K(\omega + \omega') G_{\beta'\alpha}^a(\omega')] \quad (5.284)$$

$$= -\frac{i}{4} \hbar \int \frac{d\omega'}{2\pi} \text{tr} [J_m G^r(\omega + \omega') J_{m'} G^K(\omega') + J_m G^K(\omega + \omega') J_{m'} G^a(\omega')] . \quad (5.285)$$

In the global equilibrium, the fluctuation-dissipation theorem takes place:

$$G^K(\omega) = (1 - 2f(\omega))(G^r(\omega) - G^a(\omega)). \quad (5.286)$$

Using that

$$\int \frac{d\omega'}{2\pi} \text{tr} [J_m G^r(\omega + \omega') J_{m'} G^r(\omega')] = \int \frac{d\omega'}{2\pi} \text{tr} [J_m G^a(\omega + \omega') J_{m'} G^a(\omega')] = 0 \quad (5.287)$$

for the analyticity reason, we cast

$$G_{\hat{j}_m, \hat{j}_{m'}}^r(\omega) = \frac{i}{2} \hbar \int \frac{d\omega'}{2\pi} f(\omega') \text{tr} [J_m G^r(\omega + \omega') J_{m'} \{G^r(\omega') - G^a(\omega')\}] \quad (5.288)$$

$$+ \frac{i}{2} \hbar \int \frac{d\omega'}{2\pi} f(\omega') \text{tr} [J_m \{G^r(\omega') - G^a(\omega')\} J_{m'} G^a(-\omega + \omega')] . \quad (5.289)$$

Most easily one can compute the "imaginary" part of $G_{\hat{j}_m, \hat{j}_{m'}}^r(\omega)$:

$$\begin{aligned} & \frac{G_{\hat{j}_m, \hat{j}_{m'}}^r(\omega) - G_{\hat{j}_{m'}, \hat{j}_m}^{r*}(\omega)}{2i} \\ &= \frac{\hbar}{4} \int \frac{d\omega'}{2\pi} f(\omega') \text{tr} [J_m \{G^r(\omega + \omega') - G^a(\omega + \omega')\} J_{m'} \{G^r(\omega') - G^a(\omega')\}] \end{aligned} \quad (5.290)$$

$$- \frac{\hbar}{4} \int \frac{d\omega'}{2\pi} f(\omega') \text{tr} [J_m \{G^r(\omega') - G^a(\omega')\} J_{m'} \{G^r(-\omega + \omega') - G^a(-\omega + \omega')\}] . \quad (5.291)$$

How to get transitions from this formula? Choosing $m = m'$, we see that it is an odd function of ω . Therefore it vanishes at $\omega = 0$. Supposing $T = 0$, we get an integral of the type

$$\begin{aligned} & \int_{-\infty}^0 d\omega' \delta(\omega' + E'_A) \\ & \times [\delta(\omega + \omega' - E_A) + \delta(\omega + \omega' + E_A) - \delta(-\omega + \omega' - E_A) - \delta(-\omega + \omega' + E_A)] \\ & = \delta(\omega - E'_A - E_A) + \delta(\omega - E'_A + E_A) - \delta(\omega + E'_A + E_A) - \delta(\omega + E'_A - E_A). \end{aligned} \quad (5.292)$$

5.9 Rest theory

5.9.1 Model Hamiltonian for general terminal case

Model

We consider a d -dimensional local quantum system described by a set of fermion fields $\{\hat{\psi}_{\alpha, \sigma}(\mathbf{x})\}$, where $\mathbf{x} \in \Omega \subset \mathbb{R}^d$ is the position-space quantum number, $\alpha \in \{1, \dots, N_c\}$ is the channel index (N_c -number of channels), $\sigma = \pm$ is the spin projection, and Ω is the volume occupied by the system. As always, these satisfy the following Fermion algebra:

$$\{\hat{\psi}_{\alpha\sigma}(\mathbf{x}), \hat{\psi}_{\alpha'\sigma'}(\mathbf{x}')\} = 0, \quad \{\hat{\psi}_{\alpha\sigma}(\mathbf{x}), \hat{\psi}_{\alpha'\sigma'}^\dagger(\mathbf{x}')\} = \delta_{\alpha, \alpha'} \delta_{\sigma, \sigma'} \delta^{(d)}(\mathbf{x} - \mathbf{x}'). \quad (5.293)$$

We assume that such fields evolve under the following effective single-particle Hamiltonian

$$H_S = \sum_{\alpha, \alpha'=1}^{N_c} \sum_{\sigma, \sigma'=\pm} \int_{\Omega} d^{(d)}\mathbf{x} \hat{\psi}_{\alpha\sigma}^{\dagger}(\mathbf{x}) h_{\alpha\sigma, \alpha'\sigma'}(\mathbf{p}, \mathbf{x}) \hat{\psi}_{\alpha'\sigma'}(\mathbf{x}). \quad (5.294)$$

Here the Fermi fields $\hat{\psi}_{\alpha\sigma}(\mathbf{x})$ respect the open boundary conditions

$$\hat{\psi}_{\alpha\sigma}(\mathbf{x}) \Big|_{\mathbf{x} \in \partial\Omega} = 0, \quad (5.295)$$

at the system's surface². Above $\mathbf{p} = -i\nabla$ stands for the first-quantized momentum operator.

Further, we assume that there are N_e superconducting electrodes attached to N_e points $\mathbf{R}_j \in \Omega$, $j = 1, \dots, N_e$.

In what follows, we model the electrodes as one-dimensional semi-infinite s -wave superconductors, described by the following set of Fermi field operators $\{\chi_{\alpha\sigma,j}(x)\}$ governed by the BCS-type Hamilton operator:

$$\begin{aligned} H_E = & \sum_{j=1}^{N_e} \sum_{\alpha=1}^{N_c} \sum_{\sigma=\pm} \int_0^{\infty} dx \hat{\chi}_{\alpha\sigma,j}^{\dagger}(x) \left[\frac{p^2}{2m_j} - \mu_j \right] \hat{\chi}_{\alpha\sigma,j}(x) \\ & + \sum_{j=1}^{N_e} \sum_{\alpha=1}^{N_c} \int_0^{\infty} dx \left[\Delta_j \hat{\chi}_{\alpha+,j}^{\dagger}(x) \hat{\chi}_{\alpha-,j}^{\dagger}(x) + \Delta_j^* \hat{\chi}_{\alpha-,j}(x) \hat{\chi}_{\alpha+,j}(x) \right], \end{aligned} \quad (5.296)$$

$$\{\hat{\chi}_{\alpha\sigma,j}(\mathbf{x}), \hat{\chi}_{\alpha'\sigma',j'}(\mathbf{x}')\} = 0, \quad \{\hat{\chi}_{\alpha\sigma,j}(\mathbf{x}), \hat{\chi}_{\alpha'\sigma',j'}^{\dagger}(\mathbf{x}')\} = \delta_{\alpha,\alpha'} \delta_{\sigma,\sigma'} \delta_{j,j'} \delta^{(d)}(\mathbf{x} - \mathbf{x}'). \quad (5.297)$$

To describe particle tunneling, we propose the following "point-hopping" Hamiltonian

$$H_T = \sum_{\alpha, \alpha'=1}^{N_c} \sum_{\sigma, \sigma'=\pm} \sum_{j=1}^{N_e} W_{\alpha\sigma, \alpha'\sigma'}^{(j)} \hat{\chi}_{\alpha\sigma,j}^{\dagger}(a) \hat{\psi}_{\alpha'\sigma'}(\tilde{\mathbf{R}}_j) + \text{h.c.}, \quad (5.298)$$

where a is the short-distance cut-off, and

$$\tilde{\mathbf{R}}_j = \begin{cases} \mathbf{R}_j, & \mathbf{R}_j \notin \partial\Omega, \\ \mathbf{R}_j - a\mathbf{n}_j, & \mathbf{R}_j \in \partial\Omega, \end{cases} \quad (5.299)$$

with $\mathbf{n}_j$ being the outward-pointing normal to the surface $\partial\Omega$ at the point $\mathbf{R}_j$.

Here $W_{\alpha\sigma, \alpha'\sigma'}^{(j)}$ is a phenomenological³ numerical constant of dimensions $[\text{Energy}] \times [\text{Length}]^{-d/2-1/2}$, describing the tunneling rate of quasi-particles.

²Depending on the structure of $h_{\alpha\sigma, \alpha'\sigma'}(\mathbf{p}, \mathbf{x})$, we may need to further specify the conditions on higher-order derivatives as well. Due to marginal relevance we shall not dwell on this point too much here.

³Through a first-principles microscopic study one can imagine relating these parameters to the Wannier-function overlap integral of the Coulomb interaction vertex, or similar. For the purposes of our effective model we shall treat these as parameters, and shall see that these in certain way assert the transparencies of the contacts.

It is beneficial to rewrite the Hamiltonian in the so-called extended Nambu basis. One finds

$$H_S = \frac{1}{2} \int_{\Omega} d^{(d)} \mathbf{x} \hat{\Psi}^\dagger(\mathbf{x}) \mathcal{H}_S(\mathbf{p}, \mathbf{x}) \hat{\Psi}(\mathbf{x}), \quad \mathcal{H}_S(\mathbf{p}, \mathbf{x}) = \begin{pmatrix} h(\mathbf{p}, \mathbf{x}) & 0 \\ 0 & -h^*(\mathbf{p}, \mathbf{x}) \end{pmatrix}, \quad (5.300)$$

$$\hat{\Psi}(\mathbf{x}) = \begin{pmatrix} \hat{\psi}(\mathbf{x}) \\ [\hat{\psi}^\dagger(\mathbf{x})]^T \end{pmatrix}, \quad \hat{\psi}(\mathbf{x}) = \sum_{\alpha=1}^{N_c} \sum_{\sigma=\pm} \hat{\psi}_{\alpha\sigma}(\mathbf{x}) |\alpha\sigma\rangle, \quad (5.301)$$

$$h(\mathbf{p}, \mathbf{x}) = \sum_{\alpha, \alpha'=1}^{N_c} \sum_{\sigma, \sigma'=\pm} h_{\alpha\sigma, \alpha'\sigma'}(\mathbf{p}, \mathbf{x}) |\alpha\sigma\rangle \langle \alpha'\sigma'|, \quad (5.302)$$

$$H_E = \frac{1}{2} \sum_{j=1}^{N_e} \int_0^\infty dx \hat{\Upsilon}_j^\dagger(x) \mathcal{H}_j(p, x) \hat{\Upsilon}_j(x), \quad \mathcal{H}_j(p, x) = \begin{pmatrix} \left[\frac{p^2}{2m_j} - \mu_j \right] & i\Delta_j \sigma_y \\ -i\Delta_j^* \sigma_y & -\left[\frac{p^2}{2m_j} - \mu_j \right] \end{pmatrix}, \quad (5.303)$$

$$\hat{\Upsilon}_j(x) = \begin{pmatrix} \hat{\chi}_j(x) \\ [\hat{\chi}_j^\dagger(x)]^T \end{pmatrix}, \quad \hat{\chi}_j(x) = \sum_{\alpha=1}^{N_c} \sum_{\sigma=\pm} \hat{\chi}_{\alpha\sigma,j}(\mathbf{x}) |\alpha\sigma\rangle, \quad \sigma_y = -i \sum_{\alpha=1}^{N_c} (|\alpha+\rangle \langle \alpha-| - \text{h.c.}), \quad (5.304)$$

$$H_T = \frac{1}{2} \sum_{j=1}^{N_e} \hat{\Upsilon}_j^\dagger(a) \mathcal{W}^{(j)} \hat{\Psi}(\tilde{\mathbf{R}}_j) + \text{h.c.}, \quad \mathcal{W}^{(j)} = \begin{pmatrix} W^{(j)} & 0 \\ 0 & -W^{(j)*} \end{pmatrix}, \quad (5.305)$$

$$W^{(j)} = \sum_{\alpha, \alpha'=1}^{N_c} \sum_{\sigma, \sigma'=\pm} W_{\alpha\sigma, \alpha'\sigma'}^{(j)} |\alpha\sigma\rangle \langle \alpha'\sigma'|. \quad (5.306)$$

We note that extended Nambu fields satisfy the following charge conjugation constraint

$$\mathcal{C} [\Psi^\dagger(\mathbf{x})]^T = \Psi(\mathbf{x}), \quad \mathcal{C} [\Upsilon_j^\dagger(x)]^T = \Upsilon_j(x), \quad \mathcal{C} = \begin{pmatrix} 0 & 1_{2N_c} \\ 1_{2N_c} & 0 \end{pmatrix}. \quad (5.307)$$

Furthermore, the following commutation relations are in due

$$\{\Psi_a(\mathbf{x}), \Psi_b(\mathbf{x}')\} = \mathcal{C}_{a,b} \delta^{(d)}(\mathbf{x} - \mathbf{x}'), \quad \{\Psi_a(\mathbf{x}), \Psi_b^\dagger(\mathbf{x}')\} = \delta_{a,b} \delta^{(d)}(\mathbf{x} - \mathbf{x}'), \quad (5.308)$$

$$\{\Upsilon_{a,j}(x), \Upsilon_{b,j'}(x')\} = \mathcal{C}_{a,b} \delta_{j,j'} \delta(x - x'), \quad \{\Upsilon_{a,j}(x), \Upsilon_{b,j'}^\dagger(x')\} = \delta_{a,b} \delta_{j,j'} \delta(x - x'). \quad (5.309)$$

We note that the Hamiltonian possesses chiral symmetry

$$C \mathcal{H}_S(\mathbf{p}, \mathbf{x}) C = -\mathcal{H}_S^*(\mathbf{p}, \mathbf{x}), \quad C \mathcal{H}_j(p, x) C = -\mathcal{H}_j^*(p, x), \quad C \mathcal{W}^{(j)} C = -\mathcal{W}^{(j)*}. \quad (5.310)$$

Now we consider the Heisenberg equations of motion for the Nambu Field operators

$$i \frac{d}{dt} \hat{\Upsilon}_j(x, t) = \mathcal{H}_j(p, x) \hat{\Upsilon}_j(x, t) + \delta(x - a) \mathcal{W}^{(j)} \hat{\Psi}(\tilde{\mathbf{R}}_j, t), \quad (5.311)$$

$$i \frac{d}{dt} \hat{\Psi}(\mathbf{x}, t) = \mathcal{H}_S(\mathbf{p}, \mathbf{x}) \hat{\Psi}(\mathbf{x}, t) + \sum_{j=1}^{N_e} \delta^{(d)}(\mathbf{x} - \tilde{\mathbf{R}}_j) \mathcal{W}^{(j)\dagger} \hat{\Upsilon}_j(a, t). \quad (5.312)$$

We define the following set of Gorkov/Nambu Green's functions

$$[G_{SS}(\mathbf{x}, \mathbf{x}'|t)]_{ab} = -i\Theta(t - t') \langle \{ \hat{\Psi}_a(\mathbf{x}, t), \hat{\Psi}_b^\dagger(\mathbf{x}', 0) \} \rangle, \quad (5.313)$$

$$[G_{j,j'}(x, x'|t)]_{ab} = -i\Theta(t - t') \langle \{ \hat{\Upsilon}_{a,j}(x, t), \hat{\Upsilon}_{b,j'}^\dagger(x', 0) \} \rangle, \quad (5.314)$$

$$[G_{S,j'}(\mathbf{x}, \mathbf{x}'|t)]_{ab} = -i\Theta(t - t') \langle \{ \hat{\Psi}_a(\mathbf{x}, t), \hat{\Upsilon}_{b,j'}^\dagger(x', 0) \} \rangle, \quad (5.315)$$

$$[G_{j,S}(x, \mathbf{x}'|t)]_{ab} = -i\Theta(t - t') \langle \{ \hat{\Upsilon}_{a,j}(x, t), \hat{\Psi}_b^\dagger(\mathbf{x}', 0) \} \rangle. \quad (5.316)$$

These satisfy the following set of Dyson's equations

$$\left(i\frac{d}{dt} - \mathcal{H}_S(\mathbf{p}, \mathbf{x})\right) G_{SS}(\mathbf{x}, \mathbf{x}'|t) = \sum_{j=1}^{N_e} \delta^{(d)}(\mathbf{x} - \tilde{\mathbf{R}}_j) \mathcal{W}^{(j)\dagger} G_{j,S}(a, \mathbf{x}'|t) + \delta(t - t') \delta^{(d)}(\mathbf{x} - \mathbf{x}'), \quad (5.317)$$

$$\left(i\frac{d}{dt} - \mathcal{H}_j(p, \mathbf{x})\right) G_{j,j'}(x, x'|t) = \delta(x - a) \mathcal{W}^{(j)} G_{S,j'}(\tilde{\mathbf{R}}_j, x'|t) + \delta(t - t') \delta(x - x') \delta_{j,j'}, \quad (5.318)$$

$$\left(i\frac{d}{dt} - \mathcal{H}_j(p, \mathbf{x})\right) G_{j,S}(x, \mathbf{x}'|t) = \delta(x - a) \mathcal{W}^{(j)} G_{SS}(\tilde{\mathbf{R}}_j, \mathbf{x}'|t), \quad (5.319)$$

$$\left(i\frac{d}{dt} - \mathcal{H}_S(\mathbf{p}, \mathbf{x})\right) G_{S,j'}(\mathbf{x}, x'|t) = \sum_{j=1}^{N_e} \delta^{(d)}(\mathbf{x} - \tilde{\mathbf{R}}_j) \mathcal{W}^{(j)\dagger} G_{j,j'}(a, x'|t). \quad (5.320)$$

Laplace transforming the equations we find

$$(z - \mathcal{H}_S(\mathbf{p}, \mathbf{x})) G_{SS}(\mathbf{x}, \mathbf{x}') = \sum_{j=1}^{N_e} \delta^{(d)}(\mathbf{x} - \tilde{\mathbf{R}}_j) \mathcal{W}^{(j)\dagger} G_{j,S}(a, \mathbf{x}') + \delta^{(d)}(\mathbf{x} - \mathbf{x}'), \quad (5.321)$$

$$(z - \mathcal{H}_j(p, \mathbf{x})) G_{j,j'}(x, x') = \delta(x - a) \mathcal{W}^{(j)} G_{S,j'}(\tilde{\mathbf{R}}_j, x') + \delta(x - x') \delta_{j,j'}, \quad (5.322)$$

$$(z - \mathcal{H}_j(p, \mathbf{x})) G_{j,S}(x, \mathbf{x}') = \delta(x - a) \mathcal{W}^{(j)} G_{SS}(\tilde{\mathbf{R}}_j, \mathbf{x}'), \quad (5.323)$$

$$(z - \mathcal{H}_S(\mathbf{p}, \mathbf{x})) G_{S,j'}(\mathbf{x}, x') = \sum_{j=1}^{N_e} \delta^{(d)}(\mathbf{x} - \tilde{\mathbf{R}}_j) \mathcal{W}^{(j)\dagger} G_{j,j'}(a, x'). \quad (5.324)$$

Defining the following Green's functions

$$[z - \mathcal{H}_j(p, \mathbf{x})] g_j(x, x') = \delta(x - x'), \quad (5.325)$$

$$(z - \mathcal{H}_S(\mathbf{p}, \mathbf{x})) m(\mathbf{x}, \mathbf{x}') = \delta^{(d)}(\mathbf{x} - \mathbf{x}') \quad (5.326)$$

we find

$$G_{SS}(\mathbf{x}, \mathbf{x}') = m(\mathbf{x}, \mathbf{x}') + \sum_{j=1}^{N_e} m(\mathbf{x}, \tilde{\mathbf{R}}_j) \mathcal{W}^{(j)\dagger} g_j(a, a) \mathcal{W}^{(j)} G_{SS}(\tilde{\mathbf{R}}_j, \mathbf{x}'), \quad (5.327)$$

$$G_{j,j'}(x, x') = \delta_{j,j'} g_j(x, x') + g_j(x, a) \sum_{j''=1}^{N_e} \mathcal{W}^{(j)} m(\tilde{\mathbf{R}}_j, \tilde{\mathbf{R}}_{j''}) \mathcal{W}^{(j'')\dagger} G_{j'',j'}(a, x'), \quad (5.328)$$

$$G_{j,S}(a, \mathbf{x}') = g_j(a, a) \mathcal{W}^{(j)} G_{SS}(\tilde{\mathbf{R}}_j, \mathbf{x}'), \quad (5.329)$$

$$G_{S,j'}(\tilde{\mathbf{R}}_j, x') = \sum_{j''=1}^{N_e} m(\tilde{\mathbf{R}}_j, \tilde{\mathbf{R}}_{j''}) \mathcal{W}^{(j'')\dagger} G_{j'',j'}(a, x'). \quad (5.330)$$

From the first equation we have

$$\sum_{j'=1}^{N_e} D_{j,j'} G_{SS}(\tilde{\mathbf{R}}_{j'}, \mathbf{x}') = m(\tilde{\mathbf{R}}_j, \mathbf{x}'), \quad (5.331)$$

where

$$D_{j,j'} = \delta_{j,j'} - m(\tilde{\mathbf{R}}_j, \tilde{\mathbf{R}}_{j'}) \mathcal{W}^{(j')\dagger} g_{j'}(a, a) \mathcal{W}^{(j')}. \quad (5.332)$$

One hence has

$$G_{SS}(\tilde{\mathbf{R}}_j, \mathbf{x}') = \sum_{j'=1}^{N_e} [D^{-1}]_{j,j'} m(\tilde{\mathbf{R}}_{j'}, \mathbf{x}'), \quad (5.333)$$

and

$$G_{SS}(\mathbf{x}, \mathbf{x}') = m(\mathbf{x}, \mathbf{x}') + \sum_{j,j'=1}^{N_e} m(\mathbf{x}, \tilde{\mathbf{R}}_j) \mathcal{W}^{(j)\dagger} g_j(a, a) \mathcal{W}^{(j)} [D^{-1}]_{j,j'} m(\tilde{\mathbf{R}}_{j'}, \mathbf{x}'). \quad (5.334)$$

From this we infer the spectral density of the local system

$$\rho_S(\omega) = -\frac{1}{\pi} \text{Im} \int_{\Omega} d^{(d)} \mathbf{x} \text{tr} G_{SS}(\mathbf{x}, \mathbf{x}) \quad (5.335)$$

$$= -\frac{1}{\pi} \text{Im} \int_{\Omega} d^{(d)} \mathbf{x} \text{tr} m(\mathbf{x}, \mathbf{x}) \quad (5.336)$$

$$\underbrace{\hspace{10em}}_{=\rho_S^{(0)}(\omega)}$$

$$- \frac{1}{\pi} \text{Im} \int_{\Omega} d^{(d)} \mathbf{x} \text{tr} \mathcal{W}^{(j)\dagger} g_j(a, a) \mathcal{W}^{(j)} [D^{-1}]_{j,j'} \int_{\Omega} d^{(d)} \mathbf{x} m(\tilde{\mathbf{R}}_{j'}, \mathbf{x}) m(\mathbf{x}, \tilde{\mathbf{R}}_j) \quad (5.337)$$

$$= \rho_S^{(0)}(\omega) + \frac{1}{\pi} \text{Im} \text{tr} \mathcal{W}^{(j)\dagger} g_j(a, a) \mathcal{W}^{(j)} [D^{-1}]_{j,j'} \frac{\partial m(\tilde{\mathbf{R}}_{j'}, \tilde{\mathbf{R}}_j)}{\partial \omega}. \quad (5.338)$$

The second equation in matrix form bcomes

$$\hat{G}(x, x') = \hat{g}(x, x') + \hat{g}(x, a) \mathcal{M} \hat{G}(a, x'), \quad (5.339)$$

where

$$[\hat{G}(x, x')]_{j,j'} = G_{j,j'}(x, x'), \quad [\hat{g}(x, x')]_{j,j'} = \delta_{j,j'} g_j(x, x'), \quad \mathcal{M}_{j,j'} = \mathcal{W}^{(j)} m(\tilde{\mathbf{R}}_j, \tilde{\mathbf{R}}_{j'}) \mathcal{W}^{(j')\dagger}. \quad (5.340)$$

It follows that

$$\hat{G}(a, x') = [1 - \hat{g}(a, a) \mathcal{M}]^{-1} \hat{g}(a, x'), \quad (5.341)$$

and

$$\hat{G}(x, x') = \hat{g}(x, x') + \hat{g}(x, a) \mathcal{M} [1 - \hat{g}(a, a) \mathcal{M}]^{-1} \hat{g}(a, x'). \quad (5.342)$$

It follows that

$$\rho_E(\omega) = -\frac{1}{\pi} \text{Im} \int_0^\infty dx \text{tr} \{ \hat{G}(x, x) \} = -\frac{1}{\pi} \text{Im} \underbrace{\int_0^\infty dx \text{tr} \{ \hat{g}(x, x) \}}_{=\rho_E^{(0)}(\omega)} \quad (5.343)$$

$$- \frac{1}{\pi} \text{Im} \text{tr} \left\{ \mathcal{M} [1 - \hat{g}(a, a) \mathcal{M}]^{-1} \int_0^\infty dx \hat{g}(a, x) \hat{g}(x, a) \right\} \quad (5.344)$$

$$= \rho_E^{(0)}(\omega) + \frac{1}{\pi} \text{Im} \text{tr} \left\{ [1 - \hat{g}(a, a) \mathcal{M}]^{-1} \frac{\partial \hat{g}(a, a)}{\partial \omega} \mathcal{M} \right\} \quad (5.345)$$

$$= \rho_E^{(0)}(\omega) + \frac{1}{\pi} \text{Im} \text{tr} [D^{-1}]_{j,j'} m(\tilde{\mathbf{R}}_{j'}, \tilde{\mathbf{R}}_j) \frac{\partial \mathcal{W}^{(j)\dagger} g_j(a, a) \mathcal{W}^{(j)}}{\partial \omega} \quad (5.346)$$

Combining the results together we find the total spectral density

$$\rho(\omega) = \rho_S^{(0)}(\omega) + \rho_E^{(0)}(\omega) - \frac{1}{\pi} \text{Im} \text{tr} [D^{-1}]_{j,j'} \frac{\partial [D]_{j',j}}{\partial \omega} \quad (5.347)$$

$$= \rho_S^{(0)}(\omega) + \rho_E^{(0)}(\omega) - \frac{1}{\pi} \text{Im} \frac{\partial}{\partial \omega} \text{tr} \log [D] \quad (5.348)$$

Equation of motion and Green's function

Heisenberg equations

Now we consider the Heisenberg equations of motion for the Nambu field operators

$$i \frac{d}{dt} \hat{\Upsilon}_j(x, t) = \mathcal{H}_j(p, x) \hat{\Upsilon}_j(x, t) + \delta(x - a) \mathcal{W}^{(j)} \hat{\Psi}(\tilde{\mathbf{R}}_j, t), \quad (5.349)$$

$$i \frac{d}{dt} \hat{\Psi}(\mathbf{x}, t) = \mathcal{H}_{\text{sys}}(\mathbf{p}, \mathbf{x}) \hat{\Psi}(\mathbf{x}, t) + \sum_{j=1}^{N_e} \delta^{(d)}(\mathbf{x} - \tilde{\mathbf{R}}_j) \mathcal{W}^{(j)\dagger} \hat{\Upsilon}_j(a, t). \quad (5.350)$$

(??? derivation???)

Green's functions and their properties

(??? why???) We define the following set of Gorkov/Nambu Green's functions

$$[G_{SS}(\mathbf{x}, \mathbf{x}'|t)]_{ab} := -i\Theta(t - t') \langle \{ \hat{\Psi}_a(\mathbf{x}, t), \hat{\Psi}_b^\dagger(\mathbf{x}', 0) \} \rangle, \quad (5.351)$$

$$[G_{j,j'}(x, x'|t)]_{ab} := -i\Theta(t - t') \langle \{ \hat{\Upsilon}_{a,j}(x, t), \hat{\Upsilon}_{b,j'}^\dagger(x', 0) \} \rangle, \quad (5.352)$$

$$[G_{S,j'}(\mathbf{x}, \mathbf{x}'|t)]_{ab} := -i\Theta(t - t') \langle \{ \hat{\Psi}_a(\mathbf{x}, t), \hat{\Upsilon}_{b,j'}^\dagger(\mathbf{x}', 0) \} \rangle, \quad (5.353)$$

$$[G_{j,S}(x, \mathbf{x}'|t)]_{ab} := -i\Theta(t - t') \langle \{ \hat{\Upsilon}_{a,j}(x, t), \hat{\Psi}_b^\dagger(\mathbf{x}', 0) \} \rangle. \quad (5.354)$$

Let's derive an equation for them.

(I'll add the derivation here. I just don't understand now how to get pure delta-symbol terms)

These satisfy the following set of Dyson's equations

$$\left(i \frac{d}{dt} - \mathcal{H}_{\text{sys}}(\mathbf{p}, \mathbf{x}) \right) G_{SS}(\mathbf{x}, \mathbf{x}'|t) = \sum_{j=1}^{N_e} \delta^{(d)}(\mathbf{x} - \tilde{\mathbf{R}}_j) \mathcal{W}^{(j)\dagger} G_{j,S}(a, \mathbf{x}'|t) + \delta(t - t') \delta^{(d)}(\mathbf{x} - \mathbf{x}'), \quad (5.355)$$

$$\left(i \frac{d}{dt} - \mathcal{H}_j(p, x) \right) G_{j,j'}(x, x'|t) = \delta(x - a) \mathcal{W}^{(j)} G_{S,j'}(\tilde{\mathbf{R}}_j, x'|t) + \delta(t - t') \delta(x - x') \delta_{j,j'}, \quad (5.356)$$

$$\left(i \frac{d}{dt} - \mathcal{H}_j(p, x) \right) G_{j,S}(x, \mathbf{x}'|t) = \delta(x - a) \mathcal{W}^{(j)} G_{SS}(\tilde{\mathbf{R}}_j, \mathbf{x}'|t), \quad (5.357)$$

$$\left(i \frac{d}{dt} - \mathcal{H}_{\text{sys}}(\mathbf{p}, \mathbf{x}) \right) G_{S,j'}(\mathbf{x}, \mathbf{x}'|t) = \sum_{j=1}^{N_e} \delta^{(d)}(\mathbf{x} - \tilde{\mathbf{R}}_j) \mathcal{W}^{(j)\dagger} G_{j,j'}(a, \mathbf{x}'|t). \quad (5.358)$$

By Laplace transforming $\tilde{G}(p) = \int_0^\infty dt \exp(-pt) G(t)$ the equations and remembering that (I'll add obvious formulas of L tr which are used here) we find

$$(z - \mathcal{H}_{\text{sys}}(\mathbf{p}, \mathbf{x})) G_{SS}(\mathbf{x}, \mathbf{x}') = \sum_{j=1}^{N_e} \delta^{(d)}(\mathbf{x} - \tilde{\mathbf{R}}_j) \mathcal{W}^{(j)\dagger} G_{j,S}(a, \mathbf{x}') + \delta^{(d)}(\mathbf{x} - \mathbf{x}'), \quad (5.359)$$

$$(z - \mathcal{H}_j(p, x)) G_{j,j'}(x, x') = \delta(x - a) \mathcal{W}^{(j)} G_{S,j'}(\tilde{\mathbf{R}}_j, x') + \delta(x - x') \delta_{j,j'}, \quad (5.360)$$

$$(z - \mathcal{H}_j(p, x)) G_{j,S}(x, \mathbf{x}') = \delta(x - a) \mathcal{W}^{(j)} G_{SS}(\tilde{\mathbf{R}}_j, \mathbf{x}'), \quad (5.361)$$

$$(z - \mathcal{H}_{\text{sys}}(\mathbf{p}, \mathbf{x})) G_{S,j'}(\mathbf{x}, \mathbf{x}') = \sum_{j=1}^{N_e} \delta^{(d)}(\mathbf{x} - \tilde{\mathbf{R}}_j) \mathcal{W}^{(j)\dagger} G_{j,j'}(a, \mathbf{x}'). \quad (5.362)$$

“Little” Green’s functions g and m and the matrix form by D and $\mathcal{M}$ matrices

We will see that a new definition can simplify all.

Defining the following Green’s functions

$$[z - \mathcal{H}_j(p, x)]g_j(x, x') = \delta(\mathbf{x} - \mathbf{x}'), \quad (5.363)$$

$$(z - \mathcal{H}_{\text{sys}}(\mathbf{p}, \mathbf{x}))m(\mathbf{x}, \mathbf{x}') = \delta^{(d)}(\mathbf{x} - \mathbf{x}') \quad (5.364)$$

we find

$$G_{SS}(\mathbf{x}, \mathbf{x}') = m(\mathbf{x}, \mathbf{x}') + \sum_{j=1}^{N_e} m(\mathbf{x}, \tilde{\mathbf{R}}_j) \mathcal{W}^{(j)\dagger} g_j(a, a) \mathcal{W}^{(j)} G_{SS}(\tilde{\mathbf{R}}_j, \mathbf{x}'), \quad (5.365)$$

$$G_{j,j'}(x, x') = \delta_{j,j'} g_j(x, x') + g_j(x, a) \sum_{j''=1}^{N_e} \mathcal{W}^{(j)} m(\tilde{\mathbf{R}}_j, \tilde{\mathbf{R}}_{j''}) \mathcal{W}^{(j'')\dagger} G_{j'',j'}(a, x'), \quad (5.366)$$

$$G_{j,S}(a, \mathbf{x}') = g_j(a, a) \mathcal{W}^{(j)} G_{SS}(\tilde{\mathbf{R}}_j, \mathbf{x}'), \quad (5.367)$$

$$G_{S,j'}(\tilde{\mathbf{R}}_j, x') = \sum_{j''=1}^{N_e} m(\tilde{\mathbf{R}}_j, \tilde{\mathbf{R}}_{j''}) \mathcal{W}^{(j'')\dagger} G_{j'',j'}(a, x'). \quad (5.368)$$

We can introduce

$$D_{j,j'} := \delta_{j,j'} - m(\tilde{\mathbf{R}}_j, \tilde{\mathbf{R}}_{j'}) \mathcal{W}^{(j')\dagger} g_{j'}(a, a) \mathcal{W}^{(j')}. \quad (5.369)$$

From the first equation we have

$$\sum_{j'=1}^{N_e} D_{j,j'} G_{SS}(\tilde{\mathbf{R}}_{j'}, \mathbf{x}') = m(\tilde{\mathbf{R}}_j, \mathbf{x}'), \quad (5.370)$$

One hence has

$$G_{SS}(\tilde{\mathbf{R}}_j, \mathbf{x}') = \sum_{j'=1}^{N_e} [D^{-1}]_{j,j'} m(\tilde{\mathbf{R}}_{j'}, \mathbf{x}'), \quad (5.371)$$

and

$$G_{SS}(\mathbf{x}, \mathbf{x}') = m(\mathbf{x}, \mathbf{x}') + \sum_{j,j'=1}^{N_e} m(\mathbf{x}, \tilde{\mathbf{R}}_j) \mathcal{W}^{(j)\dagger} g_j(a, a) \mathcal{W}^{(j)} [D^{-1}]_{j,j'} m(\tilde{\mathbf{R}}_{j'}, \mathbf{x}'). \quad (5.372)$$

The second equation in matrix form becomes

$$\hat{G}(x, x') = \hat{g}(x, x') + \hat{g}(x, a) \mathcal{M} \hat{G}(a, x'), \quad (5.373)$$

where

$$[\hat{G}(x, x')]_{j,j'} = G_{j,j'}(x, x'), \quad [\hat{g}(x, x')]_{j,j'} = \delta_{j,j'} g_j(x, x'), \quad \mathcal{M}_{j,j'} = \mathcal{W}^{(j)} m(\tilde{\mathbf{R}}_j, \tilde{\mathbf{R}}_{j'}) \mathcal{W}^{(j')\dagger}. \quad (5.374)$$

It follows that

$$\hat{G}(a, x') = [1 - \hat{g}(a, a) \mathcal{M}]^{-1} \hat{g}(a, x'), \quad (5.375)$$

and

$$\hat{G}(x, x') = \hat{g}(x, x') + \hat{g}(x, a) \mathcal{M} [1 - \hat{g}(a, a) \mathcal{M}]^{-1} \hat{g}(a, x'). \quad (5.376)$$

Spectral density

(I'll write what is this sp dens-ty)

From the first equation, we infer the spectral density of the local system

$$\rho_{\text{sys}}(\omega) = -\frac{1}{\pi} \text{Im} \int_{\Omega} d^{(d)} \mathbf{x} \text{tr} G_{SS}(\mathbf{x}, \mathbf{x}) \quad (5.377)$$

$$= -\frac{1}{\pi} \text{Im} \int_{\Omega} d^{(d)} \mathbf{x} \text{tr} m(\mathbf{x}, \mathbf{x}) \quad (5.378)$$

$= \rho_{\text{sys}}^{(0)}(\omega)$

$$- \frac{1}{\pi} \text{Im} \int_{\Omega} d^{(d)} \mathbf{x} \text{tr} \mathcal{W}^{(j)\dagger} g_j(a, a) \mathcal{W}^{(j)} [D^{-1}]_{j,j'} \int_{\Omega} d^{(d)} \mathbf{x} m(\tilde{\mathbf{R}}_{j'}, \mathbf{x}) m(\mathbf{x}, \tilde{\mathbf{R}}_j) \quad (5.379)$$

$$= \rho_{\text{sys}}^{(0)}(\omega) + \frac{1}{\pi} \text{Im} \text{tr} \mathcal{W}^{(j)\dagger} g_j(a, a) \mathcal{W}^{(j)} [D^{-1}]_{j,j'} \frac{\partial m(\tilde{\mathbf{R}}_{j'}, \tilde{\mathbf{R}}_j)}{\partial \omega}. \quad (5.380)$$

From the second equation, we get

$$\rho_{\text{SC el-d}}(\omega) = -\frac{1}{\pi} \text{Im} \int_0^{\infty} dx \text{tr} \{ \hat{G}(x, x) \} = \quad (5.381)$$

$$= -\frac{1}{\pi} \text{Im} \int_0^{\infty} dx \text{tr} \{ \hat{g}(x, x) \} - \frac{1}{\pi} \text{Im} \text{tr} \left\{ \mathcal{M} [1 - \hat{g}(a, a) \mathcal{M}]^{-1} \int_0^{\infty} dx \hat{g}(a, x) \hat{g}(x, a) \right\}$$

$=: \rho_{\text{SC el-d}}^{(0)}(\omega)$

(5.382)

$$= \rho_{\text{SC el-d}}^{(0)}(\omega) + \frac{1}{\pi} \text{Im} \text{tr} \left\{ [1 - \hat{g}(a, a) \mathcal{M}]^{-1} \frac{\partial \hat{g}(a, a)}{\partial \omega} \mathcal{M} \right\} \quad (5.383)$$

$$= \rho_{\text{SC el-d}}^{(0)}(\omega) + \frac{1}{\pi} \text{Im} \text{tr} [D^{-1}]_{j,j'} m(\tilde{\mathbf{R}}_{j'}, \tilde{\mathbf{R}}_j) \frac{\partial \mathcal{W}^{(j)\dagger} g_j(a, a) \mathcal{W}^{(j)}}{\partial \omega} \quad (5.384)$$

Combining the results together we find the total spectral density

$$\rho(\omega) = \rho_{\text{sys}}^{(0)}(\omega) + \rho_{\text{SC el-d}}^{(0)}(\omega) - \frac{1}{\pi} \text{Im} \text{tr} [D^{-1}]_{j,j'} \frac{\partial [D]_{j',j}}{\partial \omega} \quad (5.385)$$

$$= \rho_{\text{sys}}^{(0)}(\omega) + \rho_{\text{SC el-d}}^{(0)}(\omega) - \frac{1}{\pi} \text{Im} \frac{\partial}{\partial \omega} \text{tr} \log [D] \quad (5.386)$$

(maybe it is possible to write here some trivial example, of how it works?!!! this is the best what can be written here)

5.9.2 Formalism for one SC terminal

Formalism for one SC terminal

We consider a d -dimensional local quantum system described by a set of fermion fields with the following parameters.

- $\{\hat{\psi}_{\alpha,\sigma}(\mathbf{x})\}$, where $d = \{1, 2, 3\}$, is the set of wave functions to describe fermion fields.
- Ω is the volume occupied by the system (it can be a box or a curved tube where our system is located),

- $\mathbf{x} \in \Omega \subset \mathbb{R}^d$ is the position-space coordinate,
- $\sigma = \pm$ is the spin projection.

Each field-operator with, for example, some indices α_i, σ_i is aimed to act on vacuum as $\hat{\psi}_{\alpha_i, \sigma_i}(\mathbf{x}) |0\rangle \equiv \hat{a}^\dagger \psi_{\alpha_i, \sigma_i}(\mathbf{x}) |0\rangle$ and create a particle with a Shrodinger wavefunction $\psi_{\alpha_i, \sigma_i}(\mathbf{x})$. The index σ_i is a typical spin index, and index α is (????) telling us, how many subfermions are combined to form such effective field (???????). Together the set of wave functions with normalized coefficients describe a fermion field (????). This fermionic field can correspond to an electron, or to some effective field that describes some collective excitation in the system.

Components of fermion operator by definition satisfy the following relation:

$$\{\hat{\psi}_{\alpha\sigma}(\mathbf{x}), \hat{\psi}_{\alpha'\sigma'}(\mathbf{x}')\} = 0, \quad \{\hat{\psi}_{\alpha\sigma}(\mathbf{x}), \hat{\psi}_{\alpha'\sigma'}^\dagger(\mathbf{x}')\} = \delta_{\alpha,\alpha'} \delta_{\sigma,\sigma'} \delta^{(d)}(\mathbf{x} - \mathbf{x}'), \quad (5.387)$$

where $\{, \}$ is anticommutator.

Throughout this paper, a contracted notation with many indices will be used. (?????) In this notation the effective single-particle Hamiltonian that acts on fermion, in general acts on each component of a fermion by multiplying its each component by some function $h_{\alpha\sigma, \alpha'\sigma'}(\mathbf{p}, \mathbf{x})$:

$$H_{\text{sys}} := \sum_{\sigma, \sigma' = \pm} \int_{\Omega} d^{(d)}\mathbf{x} \hat{\psi}_{\alpha\sigma}^\dagger(\mathbf{x}) h_{\alpha\sigma, \alpha'\sigma'}(\mathbf{p}, \mathbf{x}) \hat{\psi}_{\alpha'\sigma'}(\mathbf{x}), \quad (5.388)$$

where $\mathbf{p} = -i\nabla$ is a well-known momentum operator. Here the Fermi fields $\hat{\psi}_{\alpha\sigma}(\mathbf{x})$ respect the open boundary conditions $\hat{\psi}_{\alpha\sigma}(\mathbf{x})|_{\mathbf{x} \in \partial\Omega} = 0$, at the system's surface, which corresponds to perfect scattering back at its surface. In general, we could also impose some conditions on higher-order derivatives as well, but then it will be a more specific model.

Further, we assume that there are N_e superconducting electrodes attached to N_e points $\mathbf{R}_j \in \Omega$, $j = 1, \dots, N_e$.

(???? what is the sense of the third index here???) We model the electrodes as one-dimensional semi-infinite s -wave superconductors, described by the following set of Fermi field operators $\{\chi_{\alpha\sigma, j}(\mathbf{x})\}$ governed by the BCS-type Hamilton operator:

$$H_{\text{SC el-d}} = \sum_{\sigma=\pm} \int_0^\infty dx \hat{\chi}_{\alpha\sigma, j}^\dagger(\mathbf{x}) \left[\frac{p^2}{2m_j} - \mu_j \right] \hat{\chi}_{\alpha\sigma, j}(\mathbf{x}) + \int_0^\infty dx \left[\Delta_j \hat{\chi}_{\alpha+, j}^\dagger(\mathbf{x}) \hat{\chi}_{\alpha-, j}^\dagger(\mathbf{x}) + \Delta_j^* \hat{\chi}_{\alpha-, j}(\mathbf{x}) \hat{\chi}_{\alpha+, j}(\mathbf{x}) \right], \quad (5.389)$$

here in terms of the first type, we have shifted the energy to the fermi-energy by introduction of a chemical potential μ (??? more intuition about this???). In the second type terms like in typical BCS theory we have introduced Δ (??? write here def!!!) as an effective field like in typical BCS hamiltonian (?????? I'll write later, I have too little intuition about what is here)

we have the same aticommutator relations for operators.

$$\{\hat{\chi}_{\alpha\sigma, j}(\mathbf{x}), \hat{\chi}_{\alpha'\sigma', j'}(\mathbf{x}')\} = 0, \quad \{\hat{\chi}_{\alpha\sigma, j}(\mathbf{x}), \hat{\chi}_{\alpha'\sigma', j'}^\dagger(\mathbf{x}')\} = \delta_{\alpha,\alpha'} \delta_{\sigma,\sigma'} \delta_{j,j'} \delta^{(d)}(\mathbf{x} - \mathbf{x}').$$

Let us consider a simple 2d model with one coupling to SC electrond.

To describe particle tunneling, we propose the following "point-hopping" Hamiltonian

$$H_T = \sum_{\sigma, \sigma' = \pm} W_{\alpha\sigma, \alpha'\sigma'}^{(j)} \hat{\chi}_{\alpha\sigma, j}^\dagger(a) \hat{\psi}_{\alpha'\sigma'}(\tilde{\mathbf{R}}_j) + \text{h.c.}, \quad (5.390)$$

where $\hat{\chi}$ is a fermionic operator of electrodes, $\hat{\psi}$ is a fermionic operator of superconducting system, and with the following parameters.

- the argument a is the short-distance cut-off; (?????)
- $W_{\alpha\sigma, \alpha'\sigma'}^{(j)}$ is a numerical constant matrix of dimensions $[\text{Energy}] \times [\text{Length}]^{-d/2-1/2}$, describing the tunneling rate of quasi-particles. It can be obtained, for example, from experiments.

The Nambu basis

Now we will introduce some useful formalism.

It is beneficial to rewrite the Hamiltonian in the so-called extended Nambu basis. One finds

$$\hat{\Psi}(\mathbf{x}) = \begin{pmatrix} \hat{\psi}(\mathbf{x}) \\ [\hat{\psi}^\dagger(\mathbf{x})]^T \end{pmatrix}, \quad \hat{\psi}(\mathbf{x}) = \sum_{\sigma=\pm} \hat{\psi}_{\alpha\sigma}(\mathbf{x}) |\alpha\sigma\rangle,$$

$$H_{\text{sys}} = \frac{1}{2} \int_{\Omega} d^{(d)}\mathbf{x} \hat{\Psi}^\dagger(\mathbf{x}) \mathcal{H}_{\text{sys}}(\mathbf{p}, \mathbf{x}) \hat{\Psi}(\mathbf{x}), \quad \mathcal{H}_{\text{sys}}(\mathbf{p}, \mathbf{x}) = \begin{pmatrix} h(\mathbf{p}, \mathbf{x}) & 0 \\ 0 & -h^*(\mathbf{p}, \mathbf{x}) \end{pmatrix}, \quad (5.391)$$

$$h(\mathbf{p}, \mathbf{x}) = \sum_{\sigma, \sigma' = \pm} h_{\alpha\sigma, \alpha'\sigma'}(\mathbf{p}, \mathbf{x}) |\alpha\sigma\rangle \langle \alpha'\sigma'|, \quad (5.392)$$

(????)

Now, let us rewrite superconductive electrodes with the introduction of

$$\hat{\Upsilon}_j(\mathbf{x}) := \begin{pmatrix} \hat{\chi}_j(\mathbf{x}) \\ [\hat{\chi}_j^\dagger(\mathbf{x})]^T \end{pmatrix}, \quad \hat{\chi}_j(\mathbf{x}) = \sum_{\sigma=\pm} \hat{\chi}_{\alpha\sigma, j}(\mathbf{x}) |\alpha\sigma\rangle, \quad \sigma_y = -i(|\alpha+\rangle \langle \alpha-| - \text{h.c.}),$$

Note that in comparison with Ψ now we have an index (???)

$$H_{\text{SC el-d}} = \frac{1}{2} \int_0^\infty dx \hat{\Upsilon}_j^\dagger(\mathbf{x}) \mathcal{H}_j(p, x) \hat{\Upsilon}_j(\mathbf{x}), \quad \mathcal{H}_j(p, x) = \begin{pmatrix} \left[\frac{p^2}{2m_j} - \mu_j \right] & i\Delta_j \sigma_y \\ -i\Delta_j^* \sigma_y & -\left[\frac{p^2}{2m_j} - \mu_j \right] \end{pmatrix}, \quad (5.393)$$

(????)

Finally, the coupling between electrodes and the quantum system will take the form:

$$H_T = \frac{1}{2} \hat{\Upsilon}_j^\dagger(a) \mathcal{W}^{(j)} \hat{\Psi}(\tilde{\mathbf{R}}_j) + \text{h.c.}, \quad \mathcal{W}^{(j)} = \begin{pmatrix} W^{(j)} & 0 \\ 0 & -W^{(j)*} \end{pmatrix}, \quad (5.394)$$

$$W^{(j)} = \sum_{\sigma, \sigma' = \pm} W_{\alpha\sigma, \alpha'\sigma'}^{(j)} |\alpha\sigma\rangle \langle \alpha'\sigma'|. \quad (5.395)$$

(????)

(Here I'll add particular example, how they will look in some simple systems)

Note that extended Nambu fields satisfy the following so-called charge conjugation constraint given by the matrix

$$\mathcal{C} := \begin{pmatrix} 0 & 1_{2N_c} \\ 1_{2N_c} & 0 \end{pmatrix},$$

$$\mathcal{C} [\Psi^\dagger(\mathbf{x})]^T = \Psi(\mathbf{x}), \quad \mathcal{C} [\Upsilon_j^\dagger(\mathbf{x})]^T = \Upsilon_j(\mathbf{x}), \quad (5.396)$$

(??? intuition about it)

Furthermore, the following commutation relations are in due

$$\{\Psi_a(\mathbf{x}), \Psi_b(\mathbf{x}')\} = \mathcal{C}_{a,b} \delta^{(d)}(\mathbf{x} - \mathbf{x}'), \quad \{\Psi_a(\mathbf{x}), \Psi_b^\dagger(\mathbf{x}')\} = \delta_{a,b} \delta^{(d)}(\mathbf{x} - \mathbf{x}'), \quad (5.397)$$

$$\{\Upsilon_{a,j}(\mathbf{x}), \Upsilon_{b,j'}(\mathbf{x}')\} = \mathcal{C}_{a,b} \delta_{j,j'} \delta(\mathbf{x} - \mathbf{x}'), \quad \{\Upsilon_{a,j}(\mathbf{x}), \Upsilon_{b,j'}^\dagger(\mathbf{x}')\} = \delta_{a,b} \delta_{j,j'} \delta(\mathbf{x} - \mathbf{x}'). \quad (5.398)$$

(????)

We note that the Hamiltonian possesses chiral symmetry

$$C\mathcal{H}_{\text{sys}}(\mathbf{p}, \mathbf{x})C = -\mathcal{H}_{\text{sys}}^*(\mathbf{p}, \mathbf{x}), \quad C\mathcal{H}_j(p, x)C = -\mathcal{H}_j^*(p, x), \quad C\mathcal{W}^{(j)}C = -\mathcal{W}^{(j)*}. \quad (5.399)$$

5.9.3 Formalism for several SC terminals

(a reminder that more intuition should be added in all the text below)

We consider a d -dimensional local quantum system described by a set of fermion fields with the following parameters.

- $\{\hat{\psi}_{\alpha,\sigma}(\mathbf{x})\}$, where $d = \{1, 2, 3\}$, is the set of wave functions to describe fermion fields.
- Ω is the volume occupied by the system (it can be a box or a curved tube where our system is located),
- $\mathbf{x} \in \Omega \subset \mathbb{R}^d$ is the position-space coordinate,
- $\alpha \in \{1, \dots, N_c\}$ is the channel index (N_c -number of channels)
- $\sigma = \pm$ is the spin projection.

Each field-operator with, for example, some indices α_i, σ_i is aimed to act on vacuum as $\hat{\psi}_{\alpha_i, \sigma_i}(\mathbf{x}) |0\rangle \equiv \hat{a}^\dagger \psi_{\alpha_i, \sigma_i}(\mathbf{x}) |0\rangle$ and create a particle with a Shrodinger wavefunction $\psi_{\alpha_i, \sigma_i}(\mathbf{x})$. The index σ_i is a typical spin index, and index α is (????) telling us, how many subfermions are combined to form such effective field (???????). Together the set of wave functions with normalized coefficients describe a fermion field (????). This fermionic field can correspond to an electron, or to some effective field that describes some collective excitation in the system.

Components of fermion operator by definition satisfy the following relation:

$$\{\hat{\psi}_{\alpha\sigma}(\mathbf{x}), \hat{\psi}_{\alpha'\sigma'}(\mathbf{x}')\} = 0, \quad \{\hat{\psi}_{\alpha\sigma}(\mathbf{x}), \hat{\psi}_{\alpha'\sigma'}^\dagger(\mathbf{x}')\} = \delta_{\alpha,\alpha'} \delta_{\sigma,\sigma'} \delta^{(d)}(\mathbf{x} - \mathbf{x}'), \quad (5.400)$$

where $\{, \}$ is anticommutator.

Throughout this paper, a contracted notation with many indices will be used. (?????) In this notation the effective single-particle Hamiltonian that acts on fermion, in general acts on each component of a fermion by multiplying its each component by some function $h_{\alpha\sigma,\alpha'\sigma'}(\mathbf{p}, \mathbf{x})$:

$$H_{\text{sys}} := \sum_{\alpha,\alpha'=1}^{N_c} \sum_{\sigma,\sigma'=\pm} \int_{\Omega} d^{(d)}\mathbf{x} \hat{\psi}_{\alpha\sigma}^{\dagger}(\mathbf{x}) h_{\alpha\sigma,\alpha'\sigma'}(\mathbf{p}, \mathbf{x}) \hat{\psi}_{\alpha'\sigma'}(\mathbf{x}), \quad (5.401)$$

where $\mathbf{p} = -i\nabla$ is a well-known momentum operator. Here the Fermi fields $\hat{\psi}_{\alpha\sigma}(\mathbf{x})$ respect the open boundary conditions $\hat{\psi}_{\alpha\sigma}(\mathbf{x})|_{\mathbf{x} \in \partial\Omega} = 0$, at the system's surface, which corresponds to perfect scattering back at its surface. In general, we could also impose some conditions on higher-order derivatives as well, but then it will be a more specific model.

Further, we assume that there are N_e superconducting electrodes attached to N_e points $\mathbf{R}_j \in \Omega$, $j = 1, \dots, N_e$.

(???? what is the sense of the third index here???) We model the electrodes as one-dimensional semi-infinite s -wave superconductors, described by the following set of Fermi field operators $\{\chi_{\alpha\sigma,j}(\mathbf{x})\}$ governed by the BCS-type Hamilton operator:

$$H_{\text{SC el-d}} = \sum_{j=1}^{N_e} \sum_{\alpha=1}^{N_c} \sum_{\sigma=\pm} \int_0^{\infty} dx \hat{\chi}_{\alpha\sigma,j}^{\dagger}(\mathbf{x}) \left[\frac{p^2}{2m_j} - \mu_j \right] \hat{\chi}_{\alpha\sigma,j}(\mathbf{x}) + \sum_{j=1}^{N_e} \sum_{\alpha=1}^{N_c} \int_0^{\infty} dx \left[\Delta_j \hat{\chi}_{\alpha+,j}^{\dagger}(\mathbf{x}) \hat{\chi}_{\alpha-,j}^{\dagger}(\mathbf{x}) + \Delta_j^* \hat{\chi}_{\alpha-,j}(\mathbf{x}) \hat{\chi}_{\alpha+,j}(\mathbf{x}) \right], \quad (5.402)$$

here in terms of the first type, we have shifted the energy to the fermi-energy by introduction of a chemical potential μ (??? more intuition about this???). In the second type terms like in typical BCS theory we have introduced Δ (??? write here def!!!) as an effective field like in typical BCS hamiltonian (?????? I'll write later, I have too little intuition about what is here) we have the same aticommutator relations for operators.

$$\{\hat{\chi}_{\alpha\sigma,j}(\mathbf{x}), \hat{\chi}_{\alpha'\sigma',j'}(\mathbf{x}')\} = 0, \quad \{\hat{\chi}_{\alpha\sigma,j}(\mathbf{x}), \hat{\chi}_{\alpha'\sigma',j'}^{\dagger}(\mathbf{x}')\} = \delta_{\alpha,\alpha'} \delta_{\sigma,\sigma'} \delta_{j,j'} \delta^{(d)}(\mathbf{x} - \mathbf{x}').$$

This can be seen in the illustration:

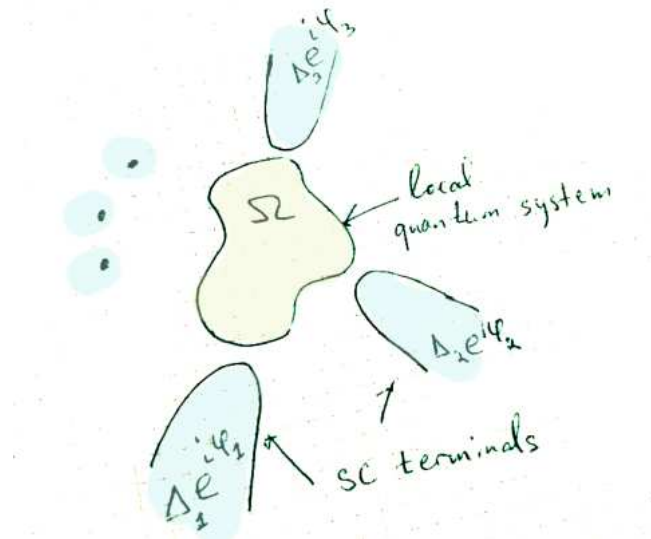


Figure 32: Illustration of the fermionic system that coupled to superconducting terminals

(paragraph about Fig. 32 (??))

To describe particle tunneling, we propose the following "point-hopping" Hamiltonian

$$H_T = \sum_{\alpha, \alpha'=1}^{N_c} \sum_{\sigma, \sigma'=\pm} \sum_{j=1}^{N_e} W_{\alpha\sigma, \alpha'\sigma'}^{(j)} \hat{\chi}_{\alpha\sigma, j}^\dagger(a) \hat{\psi}_{\alpha'\sigma'}(\tilde{\mathbf{R}}_j) + \text{h.c.}, \quad (5.403)$$

where $\hat{\chi}$ is a fermionic operator of electrodes, $\hat{\psi}$ is a fermionic operator of superconducting system, and with the following parameters.

- the argument a is the short-distance cut-off; (????)
- $\tilde{\mathbf{R}}_j$ is (???????)

$$\tilde{\mathbf{R}}_j = \begin{cases} \mathbf{R}_j, & \mathbf{R}_j \notin \partial\Omega, \\ \mathbf{R}_j - a\mathbf{n}_j, & \mathbf{R}_j \in \partial\Omega, \end{cases} \quad (5.404)$$

with $\mathbf{n}_j$ being the outward-pointing normal to the surface $\partial\Omega$ at the point $\mathbf{R}_j$;

- $W_{\alpha\sigma, \alpha'\sigma'}^{(j)}$ is a numerical constant matrix of dimensions $[\text{Energy}] \times [\text{Length}]^{-d/2-1/2}$, describing the tunneling rate of quasi-particles. It can be obtained, for example, from experiments.

Self-energies and boundary points

To get rid of the non-universal short-energy cutoff a , we expand

$$\Sigma_j(z) = \mathcal{W}^{(j)\dagger} g^{(j)}(a, a) \mathcal{W}^{(j)} = a^2 \mathcal{W}^{(j)\dagger} g_{12}^{(j)}(0^+, 0^+) \mathcal{W}^{(j)} \quad (5.405)$$

6 Quasiclassical Green's function approach to mesoscopic superconductivity by Belzig, Wilhelm, Bruder, Schön, Zaikin

Abstract

Recent experiments on mesoscopic normal-metal–superconductor heterostructures resolve properties on length scales and at low temperatures such that the temperature is below the Thouless energy $k_B T \leq E_{\text{Th}}$. We describe the properties of these systems within the framework of quasiclassical many-body techniques. Diffusive and ballistic systems are covered, both in equilibrium and nonequilibrium situations. Thereby we demonstrate the common physical basis of various subtopics.

6.1 Introduction

Quasiclassical techniques [1, 2] have become a widely used tool in the description of mesoscopic superconductivity. In comparison to early work which concentrated mostly on the regime near the superconducting transition temperature, recent work resolved much smaller length scales and covered lower temperatures. At temperatures below the Thouless energy, i.e. in dirty metals of size d for $k_B T \leq E_{\text{Th}} \sim \hbar D/d^2$, novel features have been observed. An important physical property is the proximity effect in a normal metal in contact with a superconductor. It is intimately related to the Andreev reflection [3], a process in which an electron incident from a normal metal tries to enter a superconductor. If its energy lies below the gap, it can do so only if it finds a partner electron with opposite momentum and spin to form a Cooper pair, leaving a retroreflected hole in the normal metal. In this process two charges are transferred into the superconductor. The reflection is phase-coherent since the electron and the reflected

hole combine in a way consistent with the phase of the superconductor. The pair can maintain its phase information in the normal metal over a distance $\sim \hbar v_F/2E$, where $2E$ is the energy difference of the electron and hole. This space dependence, with diverging length scale at low energies, is responsible for several of the mesoscopic effects to be discussed below.

The proximity effect is characterized by the existence of superconducting correlations in the normal metal, expressed by a non-vanishing pair *amplitude*, $F \neq 0$. This is to be distinguished from the pair *potential* Δ , which in the absence of attractive interactions vanishes in the normal metal. Finally, the question whether the normal metal shows a *gap* in the excitation spectrum is a more subtle issue. We will analyze below how the proximity effect reduces the tunneling density of states (DOS) at low energies, and, depending on the geometry, induces a gap or a pseudo-gap.

We will present here a unified view of both equilibrium and nonequilibrium properties of normal-metal–superconductor proximity systems. The basis are the quasiclassical Green’s functions, which will be reviewed in Section 2. Both the Matsubara and the Keldysh formalism are introduced. We discuss the clean and dirty limits, which allow systematic approximations. We finish the section with remarks on the boundary conditions and the strategies of solution. There exist several excellent introductions to this topic [4, 5, 6]. They have been written mostly with the goal of describing massive superconductors at temperatures near the transition temperature. Recent experimental work covered much lower temperatures and resolved properties on sub-micron length scales. We, therefore, present the theory and applications focusing on nano-structured systems at low temperatures.

In the following sections we will discuss several applications of the formalism and compare the results with recent experiments. In the equilibrium case (Section 3) we study the change in the local density of states (LDOS) in the vicinity of a normal metal - superconductor interface. It shows a strong suppression around the Fermi energy, leading to a mini-gap if the normal metal is finite in size [7, 8]. Our results agree with experiments performed by the Saclay group [9]. Next, we analyze the proximity induced magnetic screening in a normal metal in contact with a superconductor. This allows us to describe the experiments of Mota and coworkers [10, 11] on the diamagnetic properties of normal - superconductor heterostructures. For a comparison it is necessary to study the induced superconductivity in a material with intermediate strength of the impurity scattering. We found several new phenomena regarding the non-locality of the current response in this regime [12]. Finally we consider the supercurrent through a normal metal sandwiched between two superconductors, which allows us to interpret experimental results of Courtois *et al.* [13].

In Section 4 we study mesoscopic superconductors and heterostructures in nonequilibrium situations. The quasiparticle distribution function can be disturbed, e.g., by external applied gate voltages. Under mesoscopic conditions, it acquires a specific double-step form, which has been observed by the Saclay group [14]. This nonequilibrium form can be used, in a 4-terminal geometry, to tune a supercurrent and, hence, to build a mesoscopic SNS transistor [15]. Moreover, the SNS-junction can be switched by the gate voltages to become a π -junction. We can explain in this way recent experiments by the Groningen group [16, 17, 18]. A different SN-heterostructure allows us to study the behavior of induced superconducting correlations on a dissipative current. In systems with transparent SN-contacts, the dissipative conductance is enhanced, with a peculiar non-monotonic temperature- and voltage-dependence, usually referred to as “reentrant” conductance [19, 20]. This means, at high temperatures the conductance grows with decreasing T since the proximity effect develops, it reaches a maximum at $k_B T \approx E_{Th}$ where E_{Th} is the Thouless energy, and finally it drops back to the normal state value at $T = 0$. This behavior has been observed in experiments of the Grenoble group [21, 22]. In more complex, multiply-connected structures the proximity conductance can be modulated by an external magnetic flux. This modulation is dominated by states close to the Fermi surface, for

which the spatial range of the proximity effect is large even at fairly high temperatures. This makes it possible to detect the influence of the proximity effect over distances much larger than the temperature dependent coherence length [23, 22, 19]. In systems with tunneling barriers, the excess conductance competes with the conductance suppression at the tunneling interfaces governed by the opening of an induced gap. But even without tunneling barriers, the conductance may be reduced as a consequence of a 4-terminal measurement in the thin film geometry [24, 25, 26].

The purpose of this paper is to demonstrate by the use of quasiclassical many-body techniques the common physical basis of several striking phenomena observed in recent experiments on mesoscopic superconductors and heterostructures. Of course, neither our presentation nor the bibliography can provide a complete review of the field. For a number of further applications of quasiclassical techniques we refer to articles in Refs. [27, 28, 29, 30, 31, 32] as well as the articles of Nazarov, Volkov, and Yip in this volume.

6.2 Quasiclassical formalism

6.2.1 Green's functions

Quantum field theoretical methods in terms of Green's functions are a powerful tool in all many-body problems (see e.g. [33, 34, 35]). Various systematic approximation methods and computation schemes have been formulated for them. In this section we will outline, how they can be used, within a quasiclassical approximation to describe mesoscopic normal metal - superconductor proximity systems.

The starting point for all problems in superconductivity is the Green's function in Nambu space, which combines the particle and hole space [36]. Using the pseudo-spinors as a compact notation, $\hat{\Psi}^\dagger = (\Psi_\uparrow^\dagger, \Psi_\downarrow)$, we can express the time-ordered Green's function as

$$\hat{G} = -i\langle T\hat{\Psi}(1)\hat{\Psi}^\dagger(1') \rangle = \begin{pmatrix} G(1, 1') & F(1, 1') \\ F^\dagger(1, 1') & G^\dagger(1, 1') \end{pmatrix}. \quad (6.1)$$

The Green's function contains an 'anomalous' component, the pair amplitude $F = -i\langle T\Psi_\uparrow\Psi_\downarrow \rangle$, characteristic for superconducting systems. Using the BCS Hamiltonian we can write the Gorkov equation of motion [33] for $\hat{G}$ as ⁴

$$\begin{aligned} (\hat{G}_0^{-1} - \hat{\Delta} - \hat{\Sigma}_{imp})(1, 2) \otimes \hat{G}(2, 1') &= \delta(1, 1') \\ \hat{G}(1, 2) \otimes (\hat{G}_0^{-1} - \hat{\Delta} - \hat{\Sigma}_{imp})(2, 1') &= \delta(1, 1'). \end{aligned} \quad (6.2)$$

Here 1 and 1' represent sets of space and time coordinates, and $\otimes$ includes a convolution over the coordinates. The free Green's function in Nambu space reads

$$\hat{G}_0^{-1}(1, 1') = \delta(1 - 1') \left[\hat{\tau}_3 \frac{\partial}{\partial t_1} + \frac{1}{2m} \hat{\boldsymbol{\partial}}_{\mathbf{r}_1}^2 - e\phi(1) + \mu \right], \quad (6.3)$$

where $\hat{\boldsymbol{\partial}}_{\mathbf{r}} = \nabla_{\mathbf{r}} - ie\mathbf{A}(\mathbf{r})\hat{\tau}_3$ is the gauge-invariant spatial derivative. The part of the electron-phonon self-energy which is responsible for superconductivity defines the pair potential

$$\hat{\Delta}(1, 1') = \delta(1, 1') \begin{pmatrix} 0 & \Delta \\ \Delta^* & 0 \end{pmatrix} \quad \text{with} \quad \Delta(1) = \lambda \lim_{2 \rightarrow 1+} F^\dagger(2, 1). \quad (6.4)$$

Here λ is the strength of the attractive interaction. For a normal metal we have $\lambda = 0$ and hence $\Delta = 0$, but still the proximity effect manifests itself in a non-vanishing pair amplitude $F \neq 0$.

⁴Here and in the following we use units in which $\hbar = k_B = 1$.

To simplify notations, we will from now on use center-of-mass coordinates in space and time and Fourier-transform with respect to the relative coordinates.

Elastic impurity-scattering will be included within the framework of the Born approximation. The impurity self-energy then reads

$$\hat{\Sigma}(|p|, R, E, t)_{\text{imp}} = \frac{\pi}{2\tau} \left\langle \int d\xi \hat{G}(\mathbf{p}, \mathbf{R}, E, t) \right\rangle_{\mathbf{p}_F}, \quad (6.5)$$

where τ is the elastic scattering time, $\xi = p^2/(2m) - \mu$ and $\langle \dots \rangle_{\mathbf{p}_F}$ denotes averaging over the Fermi surface. In contrast, we can neglect inelastic scattering, since we will study mesoscopic samples smaller than the inelastic length. Spin-flip scattering can be taken into account by an expression similar to (6.5), see [4, 5, 6].

The Gorkov equation can in principle be used to study mesoscopic proximity systems. However, these are genuine inhomogeneous systems. Hence, dealing with full double-coordinate Green's functions – while not impossible (see e.g. [37, 38]) – may become very cumbersome, and typically one makes use of quasiclassical approximations in the course of the calculations anyhow. For a systematic and more efficient approach it is an advantage to perform this approximation already on the level of the equations of motion.

6.2.2 The quasiclassical approximation

We will now formulate the equations of motion within the quasiclassical approximation. For simplicity, with the applications to be discussed in mind, we restrict ourselves to equilibrium and stationary nonequilibrium situations, although the quasiclassical approximation can be applied also to time-dependent problems, e.g. relaxation processes and collective modes [29].

The Green's function (6.1) oscillates as a function of the relative coordinate $|\mathbf{r} - \mathbf{r}'|$ on a scale of the Fermi wavelength λ_F . This is much shorter than the characteristic length scales in the typical problems in superconductivity, $\xi_0 = v_F/\Delta$ and $\xi_T = v_F/2\pi T$. Moreover, in those problems, it is important to study the phase of the *two-electron* wave function, which according to the definition (6.1) depends on the center of mass coordinate. For these reasons it is possible, and sufficient for most applications, to integrate out the dependence on the relative coordinate. This has been recognized first by Eilenberger [1] and by Larkin and Ovchinnikov [2]. We can add that the reduction is not allowed for another class of mesoscopic effects, e.g. weak localization and persistent currents, which are controlled by the phase-coherence of the *single-electron* wave function, contained in the relative coordinates of (6.1). But these effects are usually much weaker than those related to superconductivity. On the other hand, the reduction is possible also for problems involving Andreev reflection, since the essential information is again contained in the *difference* of electron and hole wave-vectors close to the Fermi surface.

Gradient expansion

When integrating over the difference variables the convolution $\otimes$ in the Gorkov equation Eq. (6.2) (integration over internal variables) requires some care. It can be expressed, after the Fourier transformation, as a Taylor series

$$(A \otimes B)(\mathbf{p}, \mathbf{r}, E) = \exp \frac{i}{2} \left(\partial_{\mathbf{r}}^A \partial_{\mathbf{p}}^B - \partial_{\mathbf{p}}^A \partial_{\mathbf{r}}^B \right) A(\mathbf{p}, \mathbf{r}, E) B(\mathbf{p}, \mathbf{r}, E), \quad (6.6)$$

where $\mathbf{r}$ refers to center of mass coordinates. In the problems to be discussed we can neglect short-range oscillations, hence we expand this expression up to linear order.

To proceed, we first subtract the Gorkov equation (6.2) from its conjugated form. We observe, that the Green's functions and the self-energies are linear combinations of Pauli matrices *not* including the unit matrix. This simplifies the equation of motion to

$$\left[i\mathbf{p}\hat{\boldsymbol{\sigma}}_r + E\hat{\tau}_3 - i\hat{\Delta} - \hat{\Sigma}, \hat{G} \right] - \left\{ \hat{\boldsymbol{\sigma}}_r(\hat{\Delta} + \hat{\Sigma} + e\phi + \mu), \nabla_p \hat{G} \right\} + \left\{ \nabla_p \hat{\Sigma}, \hat{\boldsymbol{\sigma}}_r \hat{G} \right\} = 0. \quad (6.7)$$

This form is much simpler than the original. We note that it still accounts for particle-hole asymmetry, which is necessary, e.g., for the description of thermoelectric effects [39, 40].

Quasiclassical Green's functions

On the one hand, we want to ignore the information contained in the fast oscillations of the full Green's function $\hat{G}$ as a function of $|\mathbf{r}_1 - \mathbf{r}_2|$, which produces after Fourier transformation a pronounced peak at $|\mathbf{p}| = p_F$. On the other hand, we have to pay attention to the dependence on the transport direction, i.e. on the direction of the velocity, $\mathbf{v}_F$, at the Fermi surface. To make this explicit we write $\hat{G}(\xi, \mathbf{v}_F, r, E)$, where $\xi = \frac{p^2}{2m} - \mu$ depends on the magnitude of the momentum. The quasiclassical Green's function is then defined by

$$\hat{g}(\mathbf{r}, \mathbf{v}_F, E) \equiv \frac{i}{\pi} \int d\xi \hat{G}(\xi, \mathbf{v}_F, \mathbf{r}, E), \quad (6.8)$$

where the integration contour has two parts covering both half planes, see [4] for details. From Eq. (6.7) we get, from now on setting $p = p_F$, the Eilenberger equation of motion for quasiclassical Green's functions,

$$-\left[\mathbf{v}_F \hat{\boldsymbol{\sigma}}, \hat{g}(\mathbf{r}, \mathbf{v}_F, E) \right] = \left[-iE\hat{\tau}_3 + \hat{\Delta} + \frac{1}{2\tau} \langle \hat{g}(\mathbf{r}, \mathbf{v}_F, E) \rangle_{\mathbf{v}_F}, \hat{g}(\mathbf{r}, \mathbf{v}_F, E) \right]. \quad (6.9)$$

The elements of $\hat{g}$ are

$$\hat{g} = \begin{pmatrix} g & f \\ f^\dagger & -g \end{pmatrix}, \quad (6.10)$$

which is still a linear combination of three Pauli matrices $\hat{\tau}_{1/2/3}$ with $f^\dagger$ being the time-reversed counterpart of f . This symmetry can be used for convenient parameterizations of the Eilenberger equation (see Section 6.2.6).

As the right-hand side of Eq. (6.9) vanishes (in contrast to the Gorkov equation (6.2)), it only defines the Green's function up to a multiplicative constant. The constant can be fixed by the following argument: As $\hat{g}$ is a linear combination of the Pauli matrices, the square of the Green's function is proportional to the unit matrix $\hat{g}\hat{g} = c\hat{1}$. From Eq. (6.9) we see that the proportionality constant c is space-independent. If we now consider a system containing a sufficiently large superconductor, we can identify a region “deep inside the superconductor”. Here $\hat{g}$ equals its bulk value, which can be calculated by performing the steps used in the quasiclassical approximation for the known solution [33] of the homogeneous Gorkov equation Eq. (6.2). This procedure yields $c = 1$, i.e. the Green's functions are normalized

$$\hat{g}\hat{g} = \hat{1}, \quad \text{i.e.} \quad g^2 + ff^\dagger = 1. \quad (6.11)$$

A more general, mathematical derivation of the normalization condition can be found in [39]

6.2.3 Kinetics and time-ordering

From the knowledge of the retarded and advanced Green's functions alone, we can calculate energy-dependent quantities like the density of states

$$N(\mathbf{r}, E) = N_0 \text{Re} \left\{ \langle g^R(\mathbf{r}, \mathbf{v}_F, E) \rangle \right\}. \quad (6.12)$$

This, and similar expressions including the off-diagonal Green's function will be denoted in the following as *spectral quantities*. In thermal equilibrium these quantities also determine all properties of our system. Equivalently, the thermal Green's function may be calculated in Matsubara imaginary-time technique. Then, the spectral quantities are found by analytic continuation.

In general, we need in addition information on how the quasiparticle states are occupied, i.e. about distribution functions. To evaluate those under nonequilibrium conditions we will use the Keldysh technique. Both techniques are described in numerous references, so we will not rederive them but rather explain the basic idea and provide a practical guideline for using them in various situations.

Matsubara technique

The Matsubara Green's functions technique [41] has been developed to describe many-body systems in equilibrium at finite temperature [33, 34]. In thermal equilibrium, the eigenvalues of physical observables do not depend on (real) time. In order to calculate thermal expectation values, we trace over all states using the Boltzmann weight $\exp(-H/T)$, which can be viewed as analytic continuation of the time-evolution operator $\exp(iHt)$ to imaginary direction $\tau = it$. It is sufficient to know this operator (and with it the Green's function) only in the interval $0 < \tau < 1/T$. By Fourier transformation and exploiting the Fermionic symmetry one sees that all the necessary information is contained in the Green's functions defined for a discrete set of energies $E = i\omega_n$, proportional to the Matsubara frequencies $\omega_n = (2n + 1)\pi T$ with integer values of n .

For those frequencies, in a bulk superconductor, the Green's functions read

$$g_\omega = \omega/\Omega \quad , \quad f_\omega = \Delta/\Omega \quad \text{with } \Omega = \sqrt{|\Delta|^2 + \omega^2}. \quad (6.13)$$

This form will serve in the following frequently as a boundary condition. The Green's functions in imaginary times show usually no singular or oscillatory structure and behave rather smooth and monotonic. This, together with the fact that g_ω is real in the absence of fields or phase gradients, simplifies the numerical calculations in thermal equilibrium.

Expectation values of physical quantities can be expressed via Green's functions. To calculate thermal averages we rotate onto the imaginary time axis and perform the quasiclassical approximation. For instance, the result for the supercurrent density is

$$\mathbf{j}(\mathbf{r}) = -2ie\pi N_0 T \sum_{\omega} \text{Tr} \langle \mathbf{v}_F \hat{\tau}_3 \hat{g}_\omega(v_F, r) \rangle. \quad (6.14)$$

In general, it depends non-locally on the applied fields. Similarly, the self-consistency equation for the pair potential can be obtained from the corresponding expression (6.4) for Gorkov's Green's functions [33]. The integrations leading to the quasiclassical Green's function can be performed, and the self-consistency relation reads

$$\Delta = \lambda N_0 2\pi T \sum_{\omega} \langle f_\omega \rangle. \quad (6.15)$$

The Keldysh technique

The Keldysh Green's function technique [42] allows describing many-body systems outside equilibrium. Pedagogical reviews are given in articles [6] and books [43]. It has proven particularly useful in the description of nonequilibrium superconductors [4, 5, 29]. Again, we will not rederive the formalism, rather we summarize the concepts and some major results. We

will later restrict ourselves to the dirty limit, although the Keldysh technique can be applied to ballistic systems as well [44].

By means of the Keldysh technique it is possible to describe the real-time evolution of systems in nonequilibrium and at finite temperature. We assume that the system is initially, at $t = -\infty$, in a thermal equilibrium state. The time evolution of the system, for instance of its density matrix, is described by a forward and a backward propagator. Keldysh showed that this real-time evolution, as well as an evolution along the imaginary time axis to account for thermal averaging, can be combined to a propagation along a single contour in the complex time plane, with a forward and backward branch running parallel to the real axis and a vertical part between $t = -\infty - i/T$ and $t = -\infty$ [42]. Green's functions are defined on this contour with time-ordering along the Keldysh contour. Depending on whether the time arguments are on the forward or backward part, the Green's functions reduce to different analytic parts. For instance, the time-ordered Green's functions or the Kadanoff functions $G^{(< >)}$ can be obtained [35]. From those one obtains by appropriate linear combinations the retarded and advanced Green's functions $\hat{G}^{R(A)}$ as well as the Keldysh Green's function G^K . They are related by

$$\hat{G}^>(1, 1') = i\langle \Psi^\dagger(1')\Psi(1) \rangle \quad , \quad \hat{G}^<(1, 1') = -i\langle \Psi(1)\Psi^\dagger(1') \rangle \quad , \quad (6.16)$$

and

$$\begin{aligned} \hat{G}^R(1, 1') &= \theta(t_1 - t'_1) [\hat{G}^<(1, 1') - \hat{G}^>(1, 1')] \\ \hat{G}^A(1, 1') &= -\theta(t'_1 - t_1) [\hat{G}^<(1, 1') - \hat{G}^>(1, 1')] \\ \hat{G}^K(1, 1') &= \hat{G}^>(1, 1') + \hat{G}^<(1, 1'). \end{aligned} \quad (6.17)$$

To compactify notations, these Green's functions can be written in matrix form in 'Keldysh space',

$$\check{G} = \begin{pmatrix} \hat{G}^R & \hat{G}^K \\ 0 & \hat{G}^A \end{pmatrix}. \quad (6.18)$$

If we describe superconductors, each entry of this matrix is still a 2×2 matrix in Nambu space.

The retarded and advanced Green's functions $G^{R(A)}$ determine what we will call spectral, i.e. energy-dependent properties of the system. Their Eilenberger equations have just the standard form, viz., Eq. (6.9). Usually they can be obtained by analytical continuation of the Matsubara Green's function if we set $\omega \rightarrow -iE \pm 0$. On the other hand, G^K is needed to account for properties of the system which depend on the nonequilibrium distribution function. We will later relate it to this physical quantity and indicate ways to calculate it. Expectation values of physical quantities can again be related to the Green's functions, now defined on the Keldysh contour [4, 5, 6, 35]. For instance the electrical current becomes

$$\mathbf{j}(\mathbf{r}) = -eN_0 \int_{-\infty}^{\infty} dE \text{Tr} \langle \mathbf{v}_F \hat{\tau}_3 \hat{g}^K(\mathbf{v}_F, \mathbf{r}, E) \rangle. \quad (6.19)$$

6.2.4 The dirty limit

Usadel equations in Matsubara technique

Frequently the superconducting material has strong impurity scattering and is described by the so-called 'dirty limit'. The requirement is that the elastic scattering self-energy dominates all other terms in the Eilenberger equation. In this limit the electron motion is diffusive and the

Green's functions are nearly isotropic. I.e. one can expand the Green's functions in spherical harmonics

$$\hat{g}_\omega(\mathbf{r}, \mathbf{v}_F) = \hat{G}_\omega(\mathbf{r}) + \mathbf{v}_F \hat{\mathbf{g}}_\omega(\mathbf{r}) . \quad (6.20)$$

Following the prevailing convention we denote the angular average of the quasiclassical Green's function again by a capital letter. In the expansion we assumed that $\mathbf{v}_F \hat{\mathbf{g}} \ll \hat{G}$. From the normalization condition (6.11) it follows that $\hat{G}_\omega^2(\mathbf{r}) = 1$ and $\{\hat{G}_\omega(\mathbf{r}), \hat{\mathbf{g}}_\omega(\mathbf{r})\} = 0$. Angular averaging of (6.9) yields

$$-\frac{1}{3}v_F^2 [\hat{\boldsymbol{\partial}}\mathbf{r}, \hat{\mathbf{g}}_\omega(\mathbf{r})] = [\omega\hat{\tau}_3 + \hat{\Delta}(\mathbf{r}), \hat{G}_\omega(\mathbf{r})] , \quad (6.21)$$

while averaging of (6.9) after multiplication by $\mathbf{v}_F$, yields

$$\hat{\mathbf{g}}_\omega(\mathbf{r}) = -\tau\hat{G}_\omega(\mathbf{r}) [\hat{\boldsymbol{\partial}}\mathbf{r}, \hat{G}_\omega(\mathbf{r})] . \quad (6.22)$$

Here the condition $\hat{g}(\mathbf{r})/\tau \gg \omega\hat{\tau}_3 + \hat{\Delta}$ has been used. Inserting (6.22) into (6.21) leads to the Usadel equation [45],

$$D [\hat{\boldsymbol{\partial}}\mathbf{r}, \hat{G}_\omega(\mathbf{r}) [\hat{\boldsymbol{\partial}}\mathbf{r}, \hat{G}_\omega(\mathbf{r})]] = [\omega\hat{\tau}_3 + \hat{\Delta}(\mathbf{r}), \hat{G}_\omega(\mathbf{r})] , \quad (6.23)$$

where $D = v_F^2\tau/3$ is the diffusion constant. This equation is much simpler than the original Eilenberger equations and has been widely applied to describe properties of mesoscopic proximity systems.

The current becomes in the dirty limit

$$\mathbf{j}(\mathbf{r}) = -i\frac{\pi\sigma_N}{2e}T \sum_\omega \text{Tr} \hat{\tau}_3 \hat{G}_\omega(\mathbf{r}) [\hat{\boldsymbol{\partial}}\mathbf{r}, \hat{G}_\omega(\mathbf{r})] , \quad (6.24)$$

where $\sigma_N = 2e^2N_0D$ is the normal-state conductivity. Note, that a additional restrictions on the mean free path are necessary for this local relation to hold. We will discuss this in Section 6.3.2. E.g. for a bulk superconductor with a real order parameter it is given by

$$\mathbf{j}(\mathbf{r}) = -\pi\sigma_N T \sum_\omega \frac{\Delta^2}{\Omega^2} \mathbf{A}(\mathbf{r}) , \quad (6.25)$$

which demonstrates an important property of the dirty limit, namely a local supercurrent-field relation.

The Usadel equation for Keldysh Green's functions

The reduction to the dirty limit can be performed, similar to the procedure outlined above, also for the Keldysh Green's functions. The Usadel equation in Keldysh $\times$ Nambu space thus reads

$$D [\hat{\boldsymbol{\partial}}\mathbf{r}, \check{G}(E, \mathbf{r}) [\hat{\boldsymbol{\partial}}\mathbf{r}, \check{G}(E, \mathbf{r})]] = [-iE\check{\tau}_3 + \check{\Delta}, \check{G}(E, \mathbf{r})] , \quad (6.26)$$

where

$$\check{\tau}_3 = \begin{pmatrix} \hat{\tau}_3 & 0 \\ 0 & \hat{\tau}_3 \end{pmatrix} \quad \check{\Delta} = \begin{pmatrix} \hat{\Delta} & 0 \\ 0 & \hat{\Delta} \end{pmatrix} . \quad (6.27)$$

Although the Keldysh technique works also for time-dependent situations [4, 29], we restrict ourselves here to stationary nonequilibrium problems. We further consider only structures in which all superconducting reservoirs are at voltage $V = 0$, thus avoiding effects related to the

nonequilibrium Josephson effect. The normalization condition still holds in the form $\check{G}\check{G} = \check{1}$. In terms of the components of (6.18) it implies

$$\hat{G}^R \hat{G}^R = \hat{G}^A \hat{G}^A = \hat{1}, \quad \text{and} \quad \hat{G}^R \hat{G}^K + \hat{G}^K \hat{G}^A = 0. \quad (6.28)$$

As a consequence of the second relation G^K can be parameterized as

$$\hat{G}^K = \hat{G}^R \hat{h} - \hat{h} \hat{G}^A. \quad (6.29)$$

From the Keldysh component (i.e. upper right) of the Keldysh-Usadel equation (6.26) we obtain a kinetic equation of motion for the distribution matrix $\hat{h}$

$$\begin{aligned} D \left[\nabla^2 \hat{h} + (\hat{G}^R \nabla \hat{G}^R) \nabla \hat{h} - \nabla \hat{h} (\hat{G}^A \nabla \hat{G}^A) - \nabla (\hat{G}^R (\nabla \hat{h}) \hat{G}^A) \right] \\ - \left(\hat{G}^R [\hat{\Delta}, \hat{h}] - [\hat{\Delta}, \hat{h}] \hat{G}^A \right) + iE \left(\hat{G}^R [\hat{h}, \hat{\tau}_3] - [\hat{h}, \hat{\tau}_3] \hat{G}^A \right) = 0. \end{aligned} \quad (6.30)$$

Here, we made use of the fact that $G^{R(A)}$ satisfy the respective components of the Usadel equation. The kinetic equation has only two independent entries. Since, furthermore, it is a linear equation, we can assume $\hat{h}$ to be diagonal. Returning to the definitions, we can relate it to the distribution functions for electrons and holes

$$\hat{h} = \begin{pmatrix} 1 - 2f_{el} & 0 \\ 0 & 2f_h - 1 \end{pmatrix}, \quad (6.31)$$

where the energy is measured from the chemical potential of the superconductor. In thermal equilibrium, for instance in the reservoirs at voltage V_R , it can be expressed by Fermi functions

$$\hat{h}_{eq} = \begin{pmatrix} 1 - 2f(E) & 0 \\ 0 & 2f(-E) - 1 \end{pmatrix} = \begin{pmatrix} \tanh\left(\frac{E+eV_R}{2T}\right) & 0 \\ 0 & \tanh\left(\frac{E-eV_R}{2T}\right) \end{pmatrix}. \quad (6.32)$$

For practical calculations, it is convenient to split the distribution matrix into odd and even components with respect to the Fermi surface $\hat{h} = f_L + f_T \hat{\tau}_3$. The notation $f_{L/T}$ refers to ‘longitudinal’ and ‘transverse’ changes of the order parameter associated with the respective nonequilibrium distributions [46, 47]. Here it is sufficient to mention that deviations of f_L from equilibrium are related to an effective temperature change, whereas f_T is related to an effective chemical potential shift. The equilibrium forms in the reservoirs are

$$f_{L(T)} = \frac{1}{2} \left[\tanh\left(\frac{E + V_R}{2T}\right) + (-) \tanh\left(\frac{E - V_R}{2T}\right) \right]. \quad (6.33)$$

We can recover the usual electron distribution function in the end of the calculation as $2f(E) = 1 - f_L(E) - f_T(E)$.

The kinetic equations for the two components reduce to two coupled diffusion equations

$$\nabla(\mathcal{D}_T \nabla f_T) + 2\text{Im}\{j_E\} \nabla f_L = 2 \frac{\mathcal{R}}{D} f_T \quad (6.34)$$

$$\nabla(\mathcal{D}_L \nabla f_L) + 2\text{Im}\{j_E\} \nabla f_T = 0 \quad (6.35)$$

with energy dependent spectral quantities

$$2\text{Im}\{j_E\} = \frac{1}{4} \text{Tr} \left[\hat{\tau}_3 \left(\hat{G}^R \hat{\nabla} \hat{G}^R - \hat{G}^A \hat{\nabla} \hat{G}^A \right) \right], \quad (6.36)$$

$$\mathcal{D}_T = \frac{1}{4} \text{Tr} \left(1 - \hat{G}^R \hat{\tau}_3 \hat{G}^A \hat{\tau}_3 \right), \quad (6.37)$$

$$\mathcal{D}_L = \frac{1}{4} \text{Tr} \left(1 - \hat{G}^R \hat{G}^A \right), \quad (6.38)$$

$$\mathcal{R} = \frac{1}{4} \text{Tr} \left[\hat{\Delta} \left(\hat{G}^R + \hat{G}^A \right) \right]. \quad (6.39)$$

Note, that the diffusion constant is generalized to energy dependent spectral diffusion coefficients, $\mathcal{D}_{T/L}(E)$. They are even functions of energy, whereas the spectral supercurrent, j_E , is an odd function of the energy. Recently, it has been argued, that in some particular situations, the diffusion equation should be supplemented by another term usually neglected[48].

The right-hand side of eq. (6.34) implies, that f_T relaxes to 0 inside the superconductor on a scale of ξ_0 , i.e. it adjusts the chemical potential to that of the superconductor. In contrast the right-hand-side of Eq. (6.35) vanishes. This component f_L relaxes only due to inelastic processes, which are neglected here, on a longer length scale [46].

Physical quantities and Keldysh Green's functions

We now list some results for physical quantities expressed by Keldysh-Usadel Green's functions. The self-consistency equation for the order parameter reads in analogy to (6.15)

$$\Delta = \frac{\lambda}{4i} \int_{-\infty}^{\infty} dE f_L(F^R - F^A), \quad (6.40)$$

while the current becomes

$$j = \frac{\sigma_N}{2} \int dE \text{Tr} \left[\hat{\tau}_3 (\hat{G}^R \nabla \hat{G}^K + \hat{G}^K \nabla \hat{G}^A) \right]. \quad (6.41)$$

We can now use the parameterization of the Keldysh Green's function by distribution functions, Eq. (6.29), and obtain terms proportional to the two distribution functions and their respective derivatives. By studying the parity of the terms in energy we note that some terms vanish, and we end up with the supercurrent

$$j_s = \frac{\sigma_N}{2} \int_{-\infty}^{\infty} dE f_L(E) \text{Im}\{j_E\} \quad (6.42)$$

and the dissipative current component

$$j_n = \frac{\sigma_N}{2} \int_{-\infty}^{\infty} dE \mathcal{D}_T(E) \nabla f_T(E). \quad (6.43)$$

The kinetic equation (6.34) guarantees the conservation of the total current, $j = j_s + j_n$, in the normal metal. The other kinetic equation (6.35) can be identified as describing the convective part of the thermal current.

6.2.5 Boundary Conditions

In the quasiclassical approximation the information on length scales of the order of the Fermi wavelength have been integrated out. Consequently one can no longer account directly for the effect of potential barriers or interfaces on this level. It turns out, however, from a study of the full theory, that potential barriers and interfaces can be accounted for by effective boundary conditions for the quasiclassical Green's functions. A derivation of boundary conditions valid at arbitrary transmission of the interface has been given by Zaitsev [49]. These boundary conditions couple the classically transmitted and reflected trajectories. The corresponding wave vectors $\mathbf{k}_{1,2}^\pm$ are shown in Fig. 33. The in- and outgoing Green's functions $\hat{g}_{1,2}^\pm$ on both sides (1 and 2) on these trajectories are combined as indicated in Fig. 33. They fulfill the boundary conditions

$$\begin{aligned} \hat{g}_1^a &= \hat{g}_2^a, \\ \hat{g}^a [R(\hat{g}_s^+)^2 + (g_s^-)^2] &= T \hat{g}_s^- \hat{g}_s^+, \end{aligned} \quad (6.44)$$

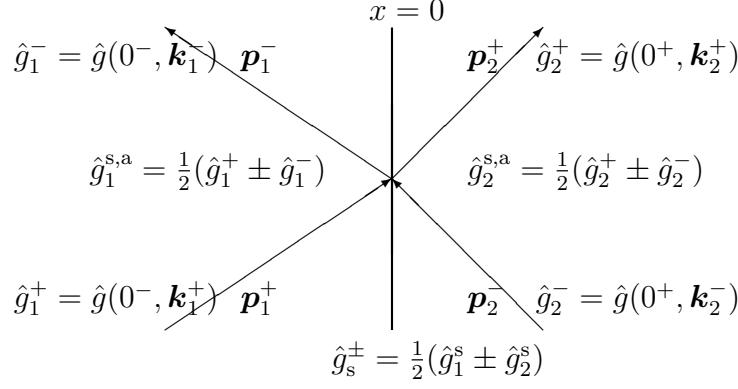


Figure 33: Trajectories at an interface. The parallel momenta are conserved, but for different Fermi velocities on both sides the perpendicular momenta differ. The Green's functions on the trajectories involved are indicated by $\hat{g}_{1,2}^{+,-}$.

where $R = 1 - T$ is the momentum-dependent reflection coefficient. The functions $\hat{g}_{1,2}^a$ and $\hat{g}_s^{+,-}$ are defined in Fig. 33. Note, that current conservation through the boundary is ensured by the first of these equations. The boundary conditions are also valid for Keldysh-Green's function, provided the Nambu matrices are replaced by Keldysh matrices.

The boundary conditions can be simplified in the dirty limit for low transparencies of the interfaces. As was shown by Kuprianov and Lukichev [50] they read in this limit

$$\begin{aligned} p_{F1}^2 l_1 \hat{G}_1 \frac{d}{dx} \hat{G}_1 &= p_{F2}^2 l_2 \hat{G}_2 \frac{d}{dx} \hat{G}_2, \\ l_2 \hat{G}_2 \frac{d}{dx} \hat{G}_2 &= t [\hat{G}_2, \hat{G}_1]. \end{aligned} \quad (6.45)$$

Here $t = \langle p_{F2x} T / p_{F2} R \rangle$ is a (small) parameter related to the transparency, and p_{F2x} is the projection of the Fermi momentum onto the normal of the interface. Recently Lambert *et al.* [51] showed that the second of these conditions constitutes the first term of an expansion in the parameter t , and they calculated the second term in this expansion.

6.2.6 Parameterizations

For the further handling of the quasiclassical equations, both numerically and analytically, two parameterizations have turned out to be especially useful: the Riccati- and the θ -parameterization.

Riccati parameterization

We will make use of the Riccati parameterization mostly in equilibrium problems for general strength of the impurity scattering. In this case we write the Matsubara Green's functions as [52]

$$\hat{g}_\omega(\mathbf{r}, \mathbf{v}_F) = \frac{1}{1 + aa^\dagger} \begin{pmatrix} 1 - aa^\dagger & 2a \\ 2a^\dagger & -1 + aa^\dagger \end{pmatrix}, \quad (6.46)$$

where the functions $a_\omega(\mathbf{r}, \mathbf{v}_F)$ and $a_\omega^\dagger(\mathbf{r}, \mathbf{v}_F)$ obey the equations

$$\begin{aligned} -\mathbf{v}_F \nabla a &= 2\tilde{\omega}a + \tilde{\Delta}^* a^2 - \tilde{\Delta}, \\ \mathbf{v}_F \nabla a^\dagger &= 2\tilde{\omega}a^\dagger + \tilde{\Delta} a^{\dagger 2} - \tilde{\Delta}^*. \end{aligned} \quad (6.47)$$

Here

$$\tilde{\omega} = \omega + \langle g \rangle / 2\tau \quad \text{and} \quad \tilde{\Delta} = \Delta + \langle f \rangle / 2\tau \quad (6.48)$$

are the renormalized energy and pair potential. These equations are of the Riccati type. The functions a and $a^\dagger$ appearing in this parameterization are directly related to the coefficients of the Andreev amplitudes. This can be seen, by noting that (6.47) in the clean limit directly follow from the Andreev equations [3] identifying $a = u/v$ [53]. For numerical purposes these equations are well suited, since they are two *uncoupled stable* differential equations (one for each direction of integration). This should be contrasted to the original Eilenberger equations, which are coupled and unstable differential equations. These equations also provide a basis to treat the linear response to an external field [12, 54].

θ parameterization

The dirty-limit equations, which do not depend on v_F any more, allow a simpler parameterization

$$\begin{aligned} \hat{G}^R(E, \mathbf{r}) &= \begin{pmatrix} \cosh(\theta) & \sinh(\theta) \exp(i\chi) \\ -\sinh(\theta) \exp(-i\chi) & -\cosh(\theta) \end{pmatrix} \\ \hat{G}^A(E, \mathbf{r}) &= \begin{pmatrix} -\cosh(\bar{\theta}) & \sinh(\bar{\theta}) \exp(i\bar{\chi}) \\ -\sinh(\bar{\theta}) \exp(-i\bar{\chi}) & \cosh(\bar{\theta}) \end{pmatrix}. \end{aligned} \quad (6.49)$$

Here, $\theta(E, \mathbf{r})$ and $\chi(E, \mathbf{r})$ are complex functions the bar denotes complex conjugation⁵. The Usadel equations in the normal metal are now written as

$$\begin{aligned} D\partial_x^2 \theta &= -2iE \sinh \theta + \frac{D}{2} (\partial_x \chi)^2 \sinh 2\theta \\ \partial_x j_E &= 0 \quad , \quad j_E = 2 \sinh^2 \theta \partial_x \chi. \end{aligned} \quad (6.50)$$

In imaginary time, it is more convenient to parameterize [55]

$$\hat{g}_\omega(\mathbf{r}) = \begin{pmatrix} \cos(\theta) & \sin(\theta) \exp(i\chi) \\ \sin(\theta) \exp(-i\chi) & -\cos(\theta) \end{pmatrix}, \quad (6.51)$$

where $\theta(\omega, \mathbf{r})$ and $\chi(\omega, \mathbf{r})$ are now real functions. At the interface to a superconducting reservoir with pair amplitude $\Delta_0 e^{i\phi_0}$ (which vanishes in normal conductors), which is connected to the probe by good metallic contacts, the boundary conditions read

$$\begin{aligned} \theta &= \theta_S = \text{Arctanh}(\Delta/E), \\ \chi &= \phi_0. \end{aligned} \quad (6.52)$$

For a normal reservoir, we set $\theta = 0$. In the presence of tunneling barriers, the boundary condition gets modified. For instance, if we consider a normal layer with resistance per length $\tilde{R}_N$ coupled via a barrier with resistance R_t to a superconductor, the condition on θ following from Eq. (6.45) reads

$$r_t \partial_x \theta = \sinh(\theta - \theta_S). \quad (6.53)$$

The direction of the derivative points away from the superconductor described by θ_S , and $r_t = R_t / \tilde{R}_N$ is the ratio of the resistances. For the kinetic equation Eq. (6.35), the boundary condition is

$$r_t D\partial_x f_{T/L} = M_{T/L} (f_{T0/L0} - f_{T/L}), \quad (6.54)$$

⁵It is also possible to parameterize using sin- and cos-functions. We prefer the present convention, which requires to choose the gauge of Δ such that for a real pair potential, $\Delta = |\Delta|$, we have $\hat{\Delta} \propto \tau_2$. Compared to our previous definition in eq. (6.4) we have to shift the phase of the pair potential, ϕ , by a gauge transformation of the order parameter phase by $\pi/2$.

where $f_{T0/L0}$ are the components of the reservoir distribution function (6.33). Using the Green's functions G_S^R and F_S^R of the superconductor we defined the spectral functions

$$M_{T/L} = \text{Re}(G_S^R)\text{Re}(G^R) \pm \text{Im}(F_S^R)\text{Im}(F^{\dagger R}) . \quad (6.55)$$

6.2.7 Solution strategies

In some cases we can derive approximate analytical solutions, for instance linearize the equation (see e.g. [56]). In general, however, we have to rely on numerical solutions. To find the solution for a given transport problem, we can proceed as follows:

1. Start with a given $\Delta(\mathbf{r})$.
2. Solve the retarded Usadel equation.
3. Take the solutions to calculate $N(E)$, j_E , and $\mathcal{D}_{L/T}$.
4. Solve the kinetic equations.
5. Calculate a new $\Delta(\mathbf{r})$, and iterate until self-consistency is achieved.

It is worth noting that considerable simplifications arise if one is only interested in equilibrium properties. Then it is possible to determine the self-consistent problem in the Matsubara representation, which is numerically much more stable than the real-time version. (In the latter one encounters frequently the cancelation of large terms.) Even if one is interested in transport properties it is of advantage to proceed in two steps. First, the pair potential, which does not depend on energy, is determined in the Matsubara representation. Then, after self-consistency has been achieved, one proceeds to solve the real-time/real-energy quasiclassical equations to obtain spectral quantities and the transport properties.

6.3 Spectral quantities and equilibrium properties

6.3.1 Density of states

As a first application we study of the local density of states in spatially inhomogeneous dirty systems. Examples are a normal metal in contact with a superconductor, inhomogeneous superconductors or boundaries of superconductors with unconventional symmetry. Depending on the problem, the pair potential $\Delta(\mathbf{r})$ has to be determined self-consistently. As mentioned above, for the ease of the numerical procedure this is best performed in imaginary frequencies. The next task is to solve the Usadel equation for real energies for a given pair potential. Once we have determined the matrix Green's function $\hat{G}^R(E, \mathbf{r})$, the local density of states is given by

$$N(E, \mathbf{r}) = N_0 \text{Re} \left[\hat{G}^R(E, \mathbf{r}) \right] = N_0 \text{Re} [\cosh(\theta(E, \mathbf{r}))] , \quad (6.56)$$

and similar for other spectral functions.

In this chapter we will calculate the local density of states $N(E, \mathbf{r})$ for geometries in which a normal metal and a superconductor are in contact[8]. First we consider the simplest case of an infinite normal metal in contact with an infinite superconductor. In this system the spatially resolved density of states has recently been measured with the help of tunnel junctions

contacting the normal metal in different positions [9]. As a second example, we consider a finite thickness normal metal in contact with a superconductor. In this case, there is a gap in the excitation spectrum of the normal metal. The form of the density of states in this system leads to rather unusual current-voltage characteristics of SNINS tunnel junctions. This has recently been measured in proximity point contacts [57], in which the ‘insulator’ is an atomic point contact with variable transparency [58]. The final system examined is an overlap junction, where the normal metal and the superconductor have an finite overlap. Interestingly, in this system a local gap-like feature in the overlap region exists. In all examples in this section the pair potential is calculated self-consistently.

Infinite proximity system

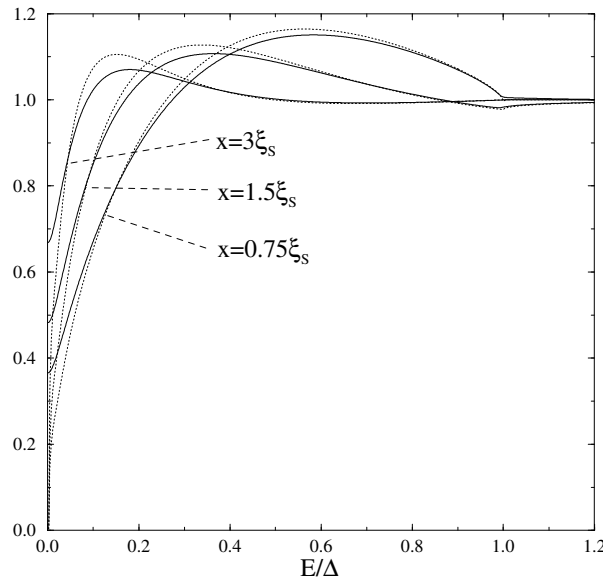


Figure 34: Density of states on the normal side of an N-S boundary for two spin-flip scattering rates: $\Gamma_{sf} = 0$ (dotted lines) and $\Gamma_{sf} = 0.015\Delta$ (solid lines).

We consider a quasi one-dimensional geometry, i. e., an infinite normal metal is in contact with a superconductor of the same cross section. The solution of the Usadel equations yields the DOS on the normal side at different distances from the NS-boundary as shown in Fig. 34. The peak height and position change with distance. In the absence of pair-breaking effects the DOS vanishes at the Fermi level for all distances (dotted curves). Inclusion of a pair-breaking mechanism, $\Gamma_{sf} \neq 0$, (solid curves) regularizes the DOS at the Fermi level, and also suppresses the peak height. The curves are in qualitative agreement with experimental data presented in Ref. [9]. The self-consistent calculation presented here leads to a slightly better fit than the curves derived in Ref. [9], in which a constant pair potential was assumed. In particular, the low-energy behavior of the experimental curves is reproduced better.

Minigap in a finite normal metal

Next we consider a normal layer with finite thickness $d_N \simeq \xi_S$ in contact with a semi-infinite superconductor occupying the half space $x < 0$. This system displays what is called a ‘minigap’ in the density of states of the normal metal. This gap has been first noted by McMillan [59] within a tunneling model which ignores the spatial dependence of the pair amplitude. We consider here the opposite limit, assuming perfect transparency of the interface but accounting for the spatial dependence of the Green’s functions.

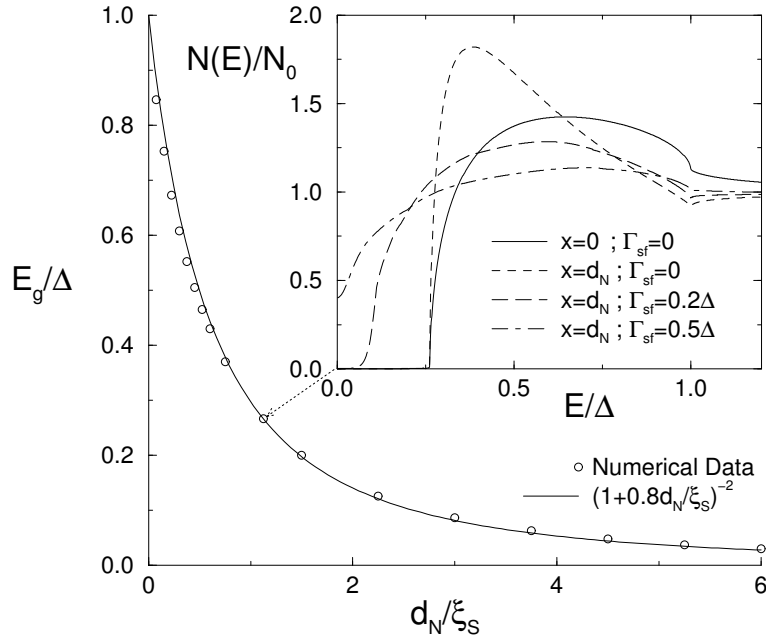


Figure 35: Minigap E_g as a function of the normal-layer thickness. Inset: local DOS of an N-layer of thickness $d_N = 1.1\xi_S$ in proximity with an bulk superconductor .

The boundary conditions at the interface to the superconductor remains unchanged as compared to what has been discussed above. But at the open side, $x = d_N$, the condition is $d\theta(E, x)/dx = 0$, which is the same as if the normal metal were bounded by an insulator. In this case the DOS in the normal layer develops a minigap at the Fermi energy. It is smaller than the superconducting gap, and decreases as the thickness of the normal layer is increased. Results obtained from the self-consistent treatment are shown in Fig. 35. Details of the shape of the DOS depend on the position within in the N-layer [8]. However, the magnitude of the minigap is space-independent, as displayed by the inset of Fig. 35. The magnitude of the gap is expected to be related to the Thouless energy $E_{Th} \sim D/d_N^2$, which becomes a relevant energy scale in mesoscopic proximity system. This relation has to be modified in the limit $d_N \rightarrow 0$. Indeed, as shown in Fig. 35, a relation of the form $E_g \sim (\text{const } \xi_S + d_N)^{-2}$ fits numerically quite well. The sum of the lengths may be interpreted as an effective thickness of the N-layer since the quasiparticle states penetrate into the superconductor to distances of the order of ξ_S . The effect of spin-flip scattering in the normal metal on the minigap structure is also shown in the inset of Fig. 35. The minigap is suppressed as Γ_{sf} is increased until it vanishes at $\Gamma_{sf} \approx 0.4\Delta$. For $\Gamma_{sf} = 0$ our results for the structure of the DOS agree with previous findings [7, 60]. A minigap was also found in heterostructures with low transparency barriers [61] and in a two-dimensional electron gas in contact to a superconductor [62]. A minigap in the DOS has also been found in a quantum dot coupled to a superconductor, if the shape of the dot is chosen such that it shows chaotic behavior in the classical limit [63]. This systems are frequently described by random matrix theory. In contrast, dot structures which are integrable in the classical limit have a non-vanishing DOS at low energy. This behavior is what one finds in clean normal metal coupled to a superconductor [64].

Overlap junction

Finally we consider a geometry, in which the normal metal and the superconductor have a finite overlap region in x -direction (see Fig. 36) which is longer than ξ_S . This geometry tries to mimic the experimental realization of junctions between different metals. A measurement

of the density of states along such a junction would give additional information on the details of the proximity effect. In principle the problem is two-dimensional. But if the total thickness $d_S + d_N$ of the two layers is less than the coherence length, we can neglect all derivatives in the transverse direction and simply average the Usadel equation over the width. This result for the self-consistently determined pair potential is shown in the lower graph of Fig. 36. In the middle of the junction the pair potential saturates at $0.5\Delta_S$.

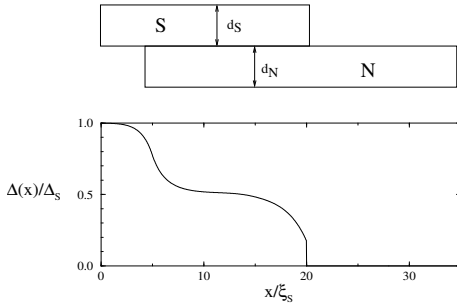


Figure 36: Geometry and pair potential of an overlap junction for $d_S/(d_S + d_N) = 0.7$. The pair potential saturates around $0.5\Delta_S$ in the middle of the junction.

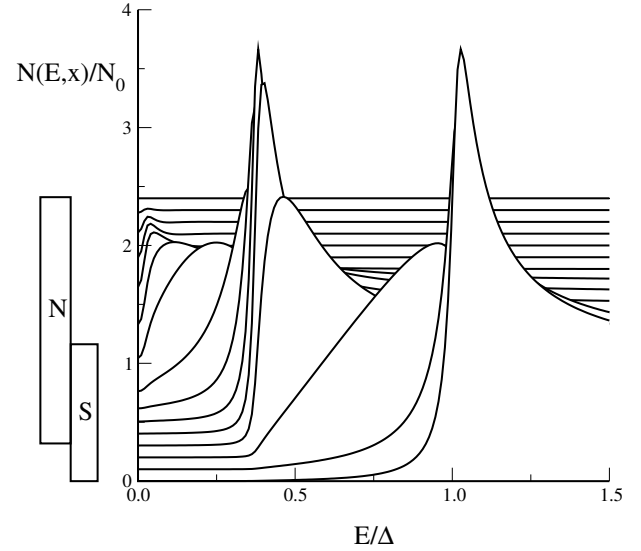


Figure 37: Density of states along an overlap junction. The curves for different positions are offset by 0.1 on the vertical axis. The distance between two adjacent LDOS is $2.5\xi_S$ and the position is indicated by the bar on the left hand side. One clearly sees the BCS-like density of states with reduced gap in the middle of the overlap junction.

The density of states along the overlap junction develops rather interesting structures as shown in Fig. 37. The DOS evolves gradually from the BCS-behavior in the lowest curve to the normal metal behavior in the topmost curve. In the middle of the junction a BCS-like peak in the density of states appears at an intermediate energy due to the aforementioned plateau in the pair potential.

6.3.2 Magnetic response of a proximity structure

The induced superconductivity enables a normal metal to produce magnetic screening [65, 66] reminiscent of the usual Meissner effect in superconductors. However, it depends on mesoscopic scales, like the thermal coherence length $\xi_T = v_F/2\pi T$, and the thickness of the normal metal d_N [67, 68, 69, 10, 70, 71, 72, 73, 74, 11, 12]. In addition, a length related to the penetration depth of superconductors, the London length $\lambda_N^2 \equiv 4\pi e^2 n_e/m$ plays an important role.

In the following we develop a linear response relation between the current and the vector potential in terms of the zero-field Green's functions in the presence of arbitrary impurity concentration. We will show, that this formula reproduces previous results for the magnetic screening obtained in the clean and dirty limit. Finally we show that for intermediate strength impurity scattering qualitatively different results arise, consistent with the observations in recent experiments.

The nonlinear response of the normal metal under proximity shows interesting properties too [65, 73, 74]. At low temperatures, the magnetic response exhibits a first order transition

both in the clean and in the dirty limit. At a certain critical field, called the breakdown field, the induced magnetization is suddenly reduced, leading to a less diamagnetic state, or even vanishing diamagnetism in the clean case. Both, the breakdown field and the size of the jump depend on temperature. All these features have been measured experimentally [10]. The qualitative agreement of the temperature dependence of the breakdown field with the clean limit theory [74] showed that the samples are close to the ballistic regime.

Linear response formalism

We consider the magnetic screening in linear response of a system composed of a superconductor occupying the half-space $x < 0$, in contact with a normal layer of finite thickness occupying the region $0 \leq x \leq d_N$, to a magnetic field $\mathbf{B}(\mathbf{r}) = B(x)\hat{\mathbf{z}}$ in z -direction. We chose a gauge such that the vector potential is $\mathbf{A}(\mathbf{r}) = A(x)\hat{\mathbf{y}}$. For the analysis it is convenient to expand the Green's functions in the Riccati parameterization

$$a_\omega(v_x, v_y, x) = a_\omega^0(v_x, x) + ia_\omega^1(v_x, v_y, x) \quad (6.57)$$

and a similar expansion for $a^\dagger$. Here a^0 and a^1 are the first terms in an expansion in $\mathbf{A}$, and we have explicitly indicated the dependence on the two relevant components of $\mathbf{v}_F$. Accordingly we expand the diagonal and off-diagonal Green's functions g_ω and f_ω . In first order the Eilenberger equation reads

$$-\frac{v_x}{2} \frac{\partial}{\partial x} a_\omega^1(v_x, v_y, x) = [\tilde{\omega}(x) + \tilde{\Delta}(x)a_\omega^0(v_x, x)] a_\omega^1(v_x, v_y, x) + ev_y A(x) a_\omega^0(v_x, x). \quad (6.58)$$

The symmetry relation $a_\omega^1(v_x, v_y, x) = -a_\omega^1(v_x, -v_y, x)$, which is a consequence of the gauge chosen, implies that the impurity self-energies contained in $\tilde{\omega}$ and $\tilde{\Delta}$, see Eq. (6.48), depend on a^0 only. Thus, we can integrate the linear inhomogeneous relation (6.58) directly. The solution still depends on the boundary conditions. Assuming perfect transparent N-S boundaries and specular reflection at the boundary to the vacuum we obtain for the current

$$\begin{aligned} j_y(x) &= -\frac{e^2 p_F^2}{\pi} T \sum_{\omega > 0} \int_{-\infty}^{d_N} dx' \int_0^{v_F} dv_x \frac{v_F^2 - v_x^2}{v_F^2 v_x} [1 + g_\omega^0(v_x, x)][1 - g_\omega^0(v_x, x')] \\ &\times \left[\Theta(x - x') m_\omega(v_x, x, x') + \Theta(x' - x) m_\omega(-v_x, x, x') + m_\omega(-v_x, x, d_N) m_\omega(v_x, d_N, x') \right] A(x') \\ &\equiv - \int_{-\infty}^{d_N} K(x, x') A(x') dx', \end{aligned} \quad (6.59)$$

where $g_\omega^0(v_x, x)$ is related to $a_\omega^0(v_x, x)$ by (6.46), and

$$m_\omega(v_x, x, x') = \exp \left(\frac{2}{v_x} \int_x^{x'} \frac{\tilde{\Delta}(x'')}{f_\omega^{\dagger 0}(v_x, x'')} dx'' \right). \quad (6.60)$$

The relation (6.59) is characterized by the interplay between the range of the kernel determined by the decay of $m_\omega(v_x, x, x')$ and the amplitude of the prefactor $[1 + g_\omega^0(x)][1 - g_\omega^0(x')]$. This determines the non-locality of the current-vector potential relation.

To exemplify the usefulness of this result we reproduce the current response of a half-infinite superconductor. Setting $d_N = 0$, the solution of the Eilenberger equation (6.9) takes the simple

form $g_0 = \omega/\Omega$, $f_0 = f_0^\dagger = \Delta/\Omega$, where $\Omega = (\Delta^2 + \omega^2)^{1/2}$. Inserting this in (6.59) we obtain the linear-response kernel

$$K_S(x, x') = \frac{e^2 p_F^2 T}{\pi} \sum_{\omega > 0} \frac{\Delta^2}{\Omega^2} \int_0^{v_F} du \frac{1 - u^2/v_F^2}{u} \left[e^{-(2\Omega + \frac{1}{\tau}) \frac{|x-x'|}{u}} + e^{(2\Omega + \frac{1}{\tau}) \frac{x+x'}{u}} \right]. \quad (6.61)$$

This result describes the current response of an arbitrary superconductor, as first derived by Gorkov[33], which here additionally includes the effect of the boundary. For fields varying rapidly spatially we arrive at a non-local current-field relation of the Pippard-type[75], while for slowly varying fields the kernel can be integrated out in Eq. (6.59), producing the local London result[76].

The solution (6.59) is valid rather generally. We now concentrate on a structure with a normal metal of thickness d_N in contact with a semi-infinite superconductor. For simplicity we neglect the penetration of the field into the superconductor by choosing $A(x=0) = 0$. This approximation leads to corrections to the induced magnetization of the normal metal of order λ_S/d_N only. As the second boundary condition we put $dA(x)/dx|_{x=d_N} = H$ with the applied magnetic field H . The spatial integration is now restricted to $[0, d_N]$ and m takes the form

$$m_\omega(v_x, x, x') = \exp \left(\frac{1}{v_x \tau} \int_x^{x'} \frac{\langle f_\omega^0(x'') \rangle}{f_\omega^{\dagger 0}(v_x, x'')} dx'' \right). \quad (6.62)$$

Here, $\langle f_\omega^0(x) \rangle = \int_{-1}^1 du f_\omega^0(v_F u, x)$ and $f_\omega^0(v_x, x)$ is related to $a_\omega^0(v_x, x)$ by (6.46).

The magnetic response of the normal-metal layer is expressed by the magnetic susceptibility

$$\chi = -\frac{1}{4\pi} \left[1 - \frac{A(d_N)}{H d_N} \right]. \quad (6.63)$$

It's calculation requires a self-consistent solution of Eq. (6.59) and the Maxwell equation, $d^2 A(x)/dx^2 = -4\pi j_y(x)$. We will now present results for the clean and dirty limits, as well as the intermediate regime.

Clean limit

In the clean limit, $\tau \rightarrow \infty$, and consequently $m = 1$. Furthermore, $g_\omega^0(v_x)$ is constant in space, and the current takes the form [67]

$$j_{\text{clean}} = -\frac{1}{4\pi \lambda^2(T) d_N} \int_0^{d_N} A(x) dx, \quad (6.64)$$

where

$$\lambda^2(T) = \begin{cases} \lambda_N^2 & \text{for } T = 0 \\ \frac{\lambda_N^2 T}{12 T_A} \exp\left(2 \frac{T}{T_A}\right) & \text{for } T \gg T_A \end{cases} \quad (6.65)$$

Here we defined $\lambda_N^2 \equiv 4\pi e^2 n_e/m$ and the Andreev temperature $T_A \equiv v_F/(2\pi d_N)$. The extremely nonlocal form of the current-field relation (6.64) leads to an overscreening for $\lambda(T) < d_N$. This means, the magnetic induction $B(x)$ changes its sign inside the normal metal, reaching for $\lambda_N \ll d_N$ at the interface to the superconductor a value as large as $-H/2$, which points in *opposite* direction of the external field. The temperature is given by the Andreev temperature T_A , which plays the same role as the Thouless energy in the diffusive case.

For the susceptibility we find

$$\chi = -\frac{1}{4\pi} \frac{3}{4 + 12\lambda^2(T)/d_N^2}. \quad (6.66)$$

In the limit $\lambda(T) \ll d_N$ the susceptibility is $3/4$ of $-1/4\pi$, thus the screening is not perfectly diamagnetic. For $\lambda(T) \gg d_N$ screening is exponentially suppressed. Both lengths coincide, $\lambda(T) = d_N$, at a crossover temperature $T \sim 2T_A \log(d_N/\lambda_N)$, which can be considerably larger than T_A itself. Interestingly, it is exactly this temperature below which the first order transition in the nonlinear response occurs [74]. The result for the linear susceptibility is displayed in Fig. 38 as the thin solid curve.

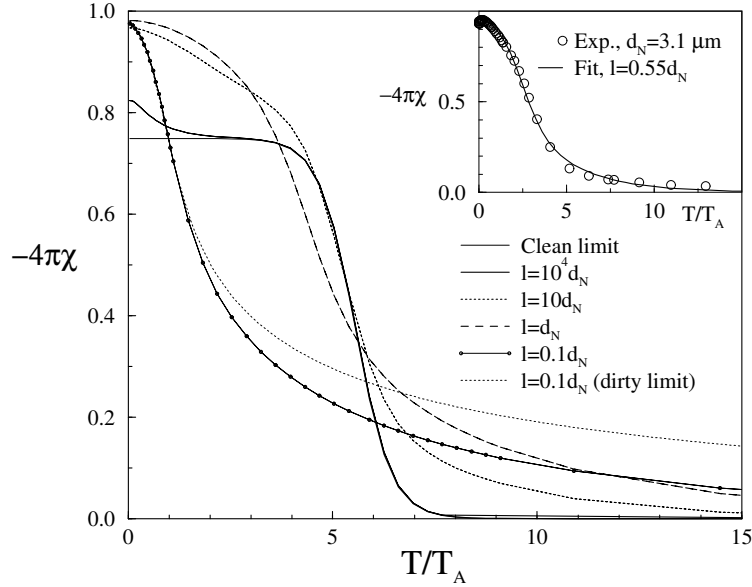


Figure 38: The susceptibility of the normal-metal layer for $\lambda_N = 0.003d_N$. The clean limit is indicated by a thin line reaching 0.75 for $T \rightarrow 0$. Shorter mean free paths, l , (even as large as $10^4 d_N$) lead to an enhanced screening at low temperatures. For shorter mean free paths, in the diffusive regime, a completely different temperature dependence is found. The dirty limit theory including a local current response deviates strongly from the correct result at higher temperatures. Inset: Comparison of experiment and theory. The thickness of the normal metal in the experiment was $3.1\mu\text{m}$, corresponding to $T_A = 540\text{mK}$. The London length $\lambda_N = 22\text{nm}$ was taken from the literature. To fit these samples a mean free path $l = 0.55d_N = 1.7\mu\text{m}$ had to be used. This is in rough agreement with the result $l = 4\mu\text{m}$ of a transport measurement.

Dirty limit

Next we assume that the mean free path, l , is the smallest length scale (except for the Fermi wave length). In the absence of fields the Green's functions vary on a scale $\xi_N = (D/2\pi T)^{1/2}$, and the dirty limit is realized if $l \ll \xi_N$. However in the situation with a space-dependent field, as seen from the kernel (6.59), a stronger requirement is needed. The dirty-limit form of the current-field relation (6.24) is found only if vector potential changes slowly on the scale of the mean free path. Assuming this to be the case we obtain the susceptibility shown in Fig. 38 as the thin dotted curve. The typical energy scale, below which the screening saturates is given by the Thouless energy $E_{\text{Th}} = D/d_N^2$ in analogy to the Andreev temperature T_A in the clean limit.

In order to check whether the requirement of a slowly varying vector potential is satisfied, we define a local penetration depth $\lambda(x, T)$ as

$$\frac{1}{\lambda^2(x, T)} = \frac{4\pi\tau}{\lambda_N^2} T \sum_{\omega>0} F_\omega^2(x). \quad (6.67)$$

In order to evaluate it we can proceed with the approximate form $F_\omega \sim \exp[-x(2\omega/D)^{1/2}]$. As a result we find for the local penetration depth

$$\lambda(x, T) \approx \begin{cases} \lambda_N x/l & \text{if } \xi_N(T) \gg d_N \\ \lambda_N \xi_N(T) e^{x/\xi_N(T)}/l & \text{if } \xi_N(T) \ll d_N. \end{cases} \quad (6.68)$$

For the local relation to be valid we need $l < \lambda(x, T)$ in the spatial region in which the screening occurs. For $T \ll E_{\text{Th}}$ this means $l \ll \lambda(d_N)$, leading to the requirement $l^2 \ll \lambda_N d_N$. The penetration depth $\lambda(d_N)$ is then given by $\lambda_N d_N/l$. It is interesting to note, that this means that full screening $\lambda(d_N) \ll d_N$ is only achieved for $l \gg \lambda_N$. For $T \gg E_{\text{Th}}$ screening takes place at $x \approx \xi_N$ and we have $l^2 \ll \lambda_N \xi_N(T)$. The low and high temperature conditions for the local approximation are different. Consequently it is possible that the dirty limit theory can be applied at low temperatures, whereas it fails to describe the correct behavior at higher temperatures.

A comparison of the dirty-limit result and the more general result for a fairly short mean free path $l = 0.1d_N$, taking into account the nonlocal screening, is shown in Fig. 38. We clearly observe a difference between the two results for higher temperatures. From the above discussion it is clear that the temperature scale at which the deviation sets in is related to the Thouless energy E_{Th} , as demonstrated by the result in Fig. 38.

Arbitrary impurity concentration

Finally we allow for arbitrary values of the mean free path. A qualitative understanding may be gained from looking at the current vector potential relation in the limit $l \gg d_N$. In the limit $T \ll T_A$ the zeroth-order Green's functions are given by the clean limit expressions. We approximate the kernel (6.59) by

$$K(x, x') = \frac{1}{8\pi\lambda^2(T)d_N} \left[e^{-|x-x'|/l} + e^{-(2d_N-x-x')/l} \right]. \quad (6.69)$$

Since $l \gg d_N$, the exponentials may be expanded to first order. As a result, we obtain two contributions to the current $j(x) = j_{\text{clean}} + j_{\text{imp}}(x)$. The first contribution j_{clean} is given by the clean limit expression Eq. (6.64). The second is an impurity induced contribution

$$j_{\text{imp}}(x) = \frac{-1}{8\pi\lambda^2(T)d_N} \int_0^{d_N} \frac{|x-x'| + 2d_N - x - x'}{l} A(x') dx'. \quad (6.70)$$

This relations demonstrates when deviations from the clean limit become important. It is clear that the impurities cannot be neglected, if $j_{\text{imp}}(x)$ is comparable to j_{clean} . We estimate this by calculating the two contributions to the current using the clean-limit vector potential. Comparing the two contributions, we find that impurities can be neglected, if

$$\lambda_{\text{eff}}^3(T) \equiv \lambda^2(T)l \gg d_N^3. \quad (6.71)$$

This equation defines a new length scale, the effective penetration depth λ_{eff} . For the clean limit to be valid at $T = 0$ the condition $\lambda_{\text{eff}}(0) > d_N$ has to be fulfilled, since in this case the screening takes place on the geometrical scale d_N . In the case $\lambda_{\text{eff}}(0) \ll d_N$ the field is

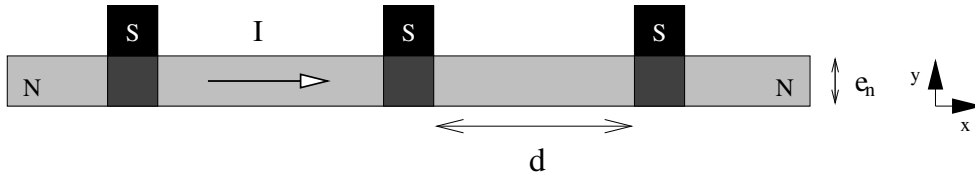


Figure 39: A proximity wire as used in the experiment [13]. The thickness is assumed to be $e_N \ll \xi_0$ and d is the distance between the strips

screened on a scale λ_{eff} and the susceptibility is strongly enhanced in comparison to the clean limit. Nevertheless, the clean-limit behavior may reappear at higher temperatures, since $\lambda_{\text{eff}}(T)$ grows with temperature.

For $T \gg T_A$ the deviations from the clean limit can be calculated as perturbation in the impurity scattering. The correction to g_ω^0 leads to a finite superfluid density close to the superconductor via the factor $1 - g_\omega^0(x') \approx (\xi_T/2l) \exp(-2x/\xi_T)$ in the kernel. The range of the kernel is modified by the correction to $f_\omega^{\dagger 0}$, which is non-negligible in this case. Since $f_\omega^{\dagger 0} \approx (\xi_T/2l) \exp(-x/\xi_T)$ we find for $T \gg T_A$

$$m_\omega(v_x, x, x') = \exp \left(\frac{1}{l} \int_x^{x'} dx'' \frac{f_\omega^0(x'')}{f_\omega^{\dagger 0}(x'')} \right) \approx \exp \left(2 \frac{x' - x}{\xi_T} \right). \quad (6.72)$$

In deriving this equation the v_x -dependence was neglected by putting $v_x = v_F$. The range of the kernel is now given by ξ_T , which is strongly temperature-dependent. For the current we find

$$j(x) \approx \frac{1}{\lambda_{\text{eff}}^3(0)} \int_0^{d_N} dx' e^{-2(x' + |x - x'|)/\xi_T} A(x'), \quad (6.73)$$

again showing the importance of the length scale λ_{eff} . In the limit $\lambda_{\text{eff}}(0) \gg \xi_T$ the field cannot be screened on the scale ξ_T , leading to a vanishing susceptibility. However, if $\lambda_{\text{eff}}(0) \ll \xi_T$ the field can be screened on a length scale smaller than ξ_T and the susceptibility will be finite.

For arbitrary mean free paths we have solved the screening problem numerically. The susceptibility as a function of temperature is shown in Fig. 38. The different curves refer to the clean limit and to mean free paths $l/d_N = 10^4, 10, 1, 0.1$. In this plot we see all the features discussed above. For large mean free paths the screening is enhanced in comparison with the clean limit at low and high temperatures, whereas the clean limit is restored in an intermediate temperature range. For small mean free paths ($l = 0.1d_N$ in the graph) screening is suppressed also in the intermediate temperature range.

A comparison to experimental data from Ref. [11] is shown in the inset. Using the mean free path as the only fit parameter they can be reproduced well by the theory. Furthermore the fitted mean free path is in rough agreement with an independent experimental estimate obtained from a resistance measurement of the same sample above the critical temperature.

6.3.3 Equilibrium supercurrent in a diffusive SNS junction

The supercurrent through an SNS junction had been studied in the past mostly on the level of the Ginzburg-Landau theory [77], which is valid near T_c when the order parameter is small. Recent experiments entered the mesoscopic regime at low temperatures, where, e.g. the coherence length $\xi_N = \sqrt{D/2\pi T}$ is comparable to the sample size d . Here a more general approach is needed.

Strong deviations from the Ginzburg-Landau behavior have shown in the current through a very thin (of thickness e_N much smaller than ξ_0), diffusive normal wire on top of which an array

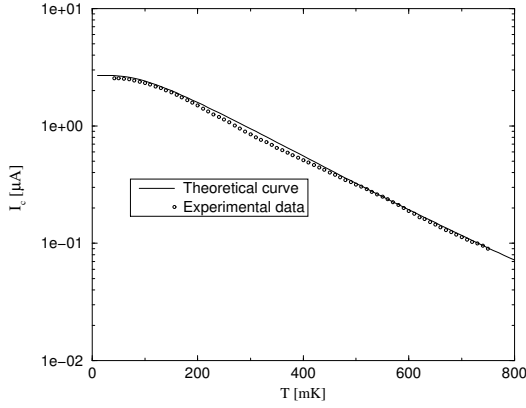


Figure 40: Critical current of the proximity wire compared with the experiment [13]: $d = 800\text{nm}$, resistance of the N-part between two strips $R_N = 0.66\Omega$

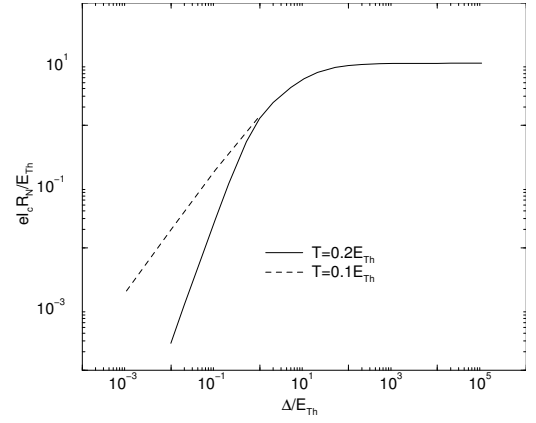


Figure 41: The critical current at low temperature for different values of the gap, normalized to the Thouless Energy. $\Delta \gg E_{\text{Th}}$ corresponds to long and $\Delta \ll E_{\text{Th}}$ to short junctions.

of superconducting strips had been deposited [13, 22], see Fig. 39. If there is a good metallic contact between the superconductors and the wire, we can assume the Green's functions under the strips to coincide [78] with the values in the strips. These, in turn, can be assumed to be reservoirs, so the Green's functions are given by their bulk values (6.13). The structure is thus equivalent to a chain of SNS junctions without barriers between the normal metal and the superconductor.

We are interested in the equilibrium supercurrent. Hence, it is sufficient to use the Matsubara technique in the parameterization (6.49). We thus have to solve the Usadel equation for the Green's function with the boundary conditions (6.52) assuming $\phi_0 = \pm\phi/2$ at $x = \pm d/2$. The general structure of the solution is such that θ decays over a scale of $\xi_\omega = \sqrt{D/2\omega}$ from the boundaries. If $\xi_\omega \ll d$, the overlap between the induced θ from the two superconductors is exponentially small. Then we can approximate the pair amplitude by a superposition of the solutions describing a semi-infinite normal part in the absence of phase gradients, i.e.

$$\begin{aligned} F_\omega(x) &= e^{i\varphi/2} \sin \theta_\omega^+(x) + e^{-i\varphi/2} \sin \theta_\omega^-(x), \\ \tan(\theta_\omega^\pm/4) &= \exp\left(\pm x \sqrt{\frac{2\omega}{D}} + A_\omega^\pm\right). \end{aligned} \quad (6.74)$$

For $T \gg E_{\text{Th}} = D/d^2$ this approximate solution is sufficient to calculate the current. From the sum (6.14) we find $I = I_c \sin \phi$, where the critical current is

$$I_c = \frac{64\pi T}{eR_N} \sum_\omega \frac{d}{\xi_\omega} \frac{\Delta^2}{[\omega + \Omega + \sqrt{2(\Omega^2 + \omega\Omega)}]^2} \exp\left(-\frac{d}{\xi_\omega}\right). \quad (6.75)$$

At high temperatures the sum over Matsubara frequencies is dominated by the first term. If, furthermore, $T \ll \Delta$, we arrive at [79, 80, 55]

$$I_c = \frac{64\pi}{3 + 2\sqrt{2}} \frac{T}{eR_N} \frac{d}{\xi_N} \exp\left(-\frac{d}{\xi_N}\right). \quad (6.76)$$

The exponent reflects the fact, that the weakest part of the structure, i.e. the overlap region in the middle, defines the bottleneck for the supercurrent.

As becomes apparent from the semi-logarithmic plot in Fig. 40 the critical current is approximately described by the numerical fit, $I_c \propto e^{-T/T^*}$, with $T^* = 12E_{\text{Th}}/\pi$. This result arises because the convex exponential factor in (6.76), typical for diffusive metals, competes with the concave pre-exponent $T^{3/2}$. This surprising result resolves a puzzle with the interpretation of the experiments of Ref. [13]. The observed linear T -dependence of the exponent is what one would expect in a ballistic system [77, 80], whereas the experiments were performed with diffusive metals. The present result agrees well with the experimental observation [13].

At lower temperatures, all Matsubara frequencies contribute to (6.14) and we have to solve the Usadel equations numerically. The results are shown in Fig. 40. As $T < E_{\text{Th}}$, the critical current saturates at a finite value $I_c(T = 0)$. The numerical results at very low T (Fig. 41) indicate that for a long junction, $d \gg \xi_0$, the spatial decay of F limits the current to $I_c R_N \propto E_{\text{Th}}$. In contrast, for a short junction the gap is the relevant cutoff and $I_c R_N \propto \Delta$ [81]. This agrees qualitatively with estimates derived from the low-energy expansion of the Usadel equation [79]. The numerical results for the critical current under experimental conditions [13] explain the experiment quantitatively well (see Fig. 40).

6.4 Applications to nonequilibrium transport problems

6.4.1 Supercurrent under nonequilibrium conditions: The SNS Transistor

In this section we will describe nonequilibrium transport problems. For this purpose the Keldysh formalism outlined before is particularly appropriate. As a first example we investigate how the supercurrent through an SNS structure is influenced by a nonequilibrium current in the normal part [15]. Specifically we study the set-up which recently has been investigated experimentally by the Groningen group [16, 17, 18]. The device is shown in Fig. 42. Its center part is a normal conductor connecting massive reservoirs at voltages $\pm V/2$. The distribution function in this part is obtained as a solution of the kinetic equation, which for a diffusive, mesoscopic normal metal reads

$$\mathcal{D}\partial_y^2 f = 0. \quad (6.77)$$

Its solution

$$f(E, y) = \left(\frac{1}{2} - \frac{y}{L}\right) f^{\text{eq}}(E + eV/2) + \left(\frac{1}{2} + \frac{y}{L}\right) f^{\text{eq}}(E - eV/2) \quad (6.78)$$

has two temperature-rounded steps at the electrochemical potentials of both reservoirs. The step heights depend on the position along the wire; in this way the distribution function interpolates linearly between the boundary conditions at $y = \pm L/2$. This functional dependence had been verified in the experiments of Pothier *et al.* [14]. At the center of the normal conductor superconducting electrodes are attached, through which a supercurrent is flowing. By choosing a symmetric geometry we can assure that the effective chemical potential of the distribution function at the site of the superconducting contacts is zero. This guarantees that no net current is flowing between the normal reservoirs and the superconducting leads. On the other hand, the nonequilibrium distribution function influences the properties of the SNS contact, and can be used to tune the supercurrent, which creates the possibility to use the device as a transistor [16]. At low temperatures even a sign reversal of the critical current, i.e. a transition to a so-called π -junction has been predicted [15] and observed [17, 18].

To describe the system we have to evaluate the spectral functions as well as the distribution function. The former can be obtained from Eq. (6.50). The low-energy ($E \ll E_{\text{Th}}$) solution,

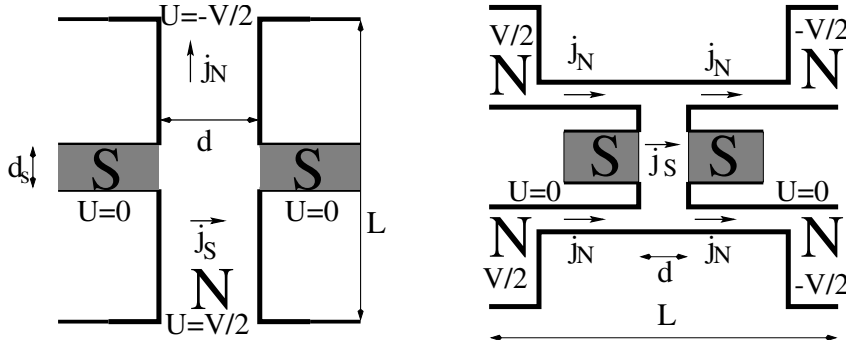


Figure 42: Different realizations of the SNS transistors. The supercurrent is tuned by (a) a perpendicular or (b) a parallel normal current flow.

calculated perturbatively, reads

$$\begin{aligned}\theta &\simeq -i\pi/2 + (E/E_{\text{Th}}) a(\phi) + O(E^3) \\ \chi &\simeq \phi x/d - (E/E_{\text{Th}})^2 b(\phi) + O(E^3),\end{aligned}\quad (6.79)$$

with real-valued functions a and b . This indicates that the spectral supercurrent $\text{Im}\{j_E\}$, vanishes in this energy range. At higher energies the problem can be analyzed numerically. However, more insight into the problem is obtained from an approximate analytic solution. For the calculation of the spectral functions between the superconducting contacts we neglect the widening of the geometry in the normal metal. In this case we can use the solution (6.74) obtained in the previous section. The approximation introduces errors of order one in numerical coefficients, equivalent to a change in an effective area. The result for the spectral supercurrent $\text{Im}\{j_E\}$, shown in Fig. 43, displays a proximity-induced mini-gap [8, 82] of size $E_g \simeq 3.2E_{\text{Th}}$ at $\phi = 0$. This gap decreases with increasing ϕ , when the induced pair-amplitudes from both sides start to interfere destructively, and vanishes at $\phi = \pi$ [83]. At energies directly above the gap, $\text{Im}\{j_E\}$ increases sharply, but rapidly decreases at higher E . At large energies, it changes sign and eventually oscillates around zero with exponentially decaying amplitude.

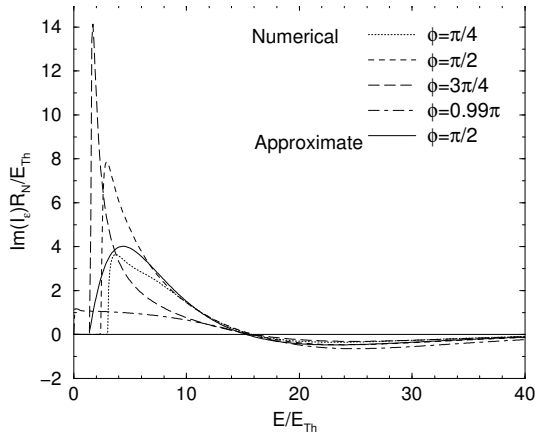


Figure 43: The spectral current $\text{Im}(j_E)$ as a function of energy for different values of the phase difference ϕ . At higher energies we find oscillations as expected from eq. (6.80).

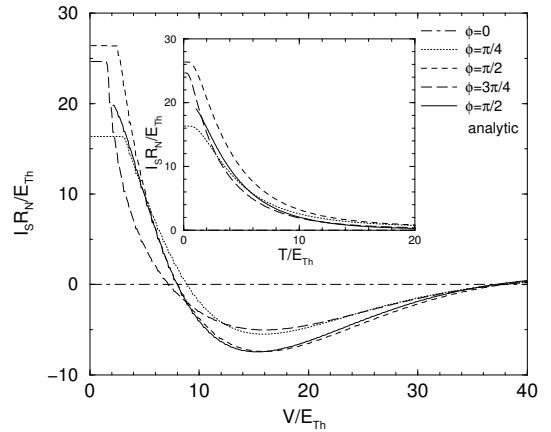


Figure 44: The supercurrent as function of control voltage (at $T = 0$) and temperature (at $V = 0$) for various values of ϕ .

To evaluate the physical current (6.42) we combine the spectral supercurrent and the quasi-particle distribution function f between the reservoirs at electrochemical potentials $\pm eV/2$

derived above. Although no net current is flowing between the normal metal and the superconducting electrodes, the coupling might influence the energy dependence of the distribution. We will now show that this is not the case: Compared to the distance between the reservoirs L the proximity effect is confined to a relatively narrow region d_S . For a quantitative analysis we have to distinguish between the two components $f_{L/T}$, governed by the kinetic equations (6.34) and (6.35). For the supercurrent (6.42) we need to determine only f_L . Because of the symmetry of the setup, the component f_T vanishes at the site of the superconducting leads.

As the voltage is applied between the reservoirs, the gradients of $f_{L/T}$ point in y -direction, i.e. perpendicular to j_E which points in x -direction. As a consequence, the scalar products $\nabla f_{L/T} \cdot j_E$ vanish and (6.34) and (6.35) decouple. Moreover, f_L satisfies the *same* boundary conditions (6.33) at the two reservoirs at $\pm V/2$. Altogether, (6.35) implies that f_L has a constant value given by (6.33) throughout the wire. It is easy to check, using $2f = 1 - f_L - f_T$ with $f_T = 0$ and (6.33), that the total distribution combines to the double-step form described before.

At first glance, it may look surprising that f_L has such a strong nonequilibrium form in the normal metal next to the two superconductors, which are assumed to be in equilibrium [56]. It turns out to be a consequence of the symmetry properties of Andreev reflection. As the SN-interfaces are good metallic contacts, the boundary conditions eq. (6.45) guarantee, that at the interface $\mathcal{D}_L$ coincides with the value in the superconductor. Thus at $|E| < \Delta$ we find from eq. (6.39) that $\mathcal{D}_L = 0$. This implies, that the SN-interfaces decouple f_L in the normal metal from its value in the superconductor. A well known physical consequence of this property is, that below the gap, when quasiparticles are transmitted by the Andreev reflection, no heat is transferred [3]. The operation of the SNS-transistor without tunneling barriers is based on this very idea [16].

Our treatment of the 2D system is approximate. A more complete solution would also take the bending of supercurrent lines into account, which leads to an effective renormalization of the cross section area. This problem has been solved by Volkov [56] for a related system with small d_S and tunneling barriers at the SN-interface. An analysis of his solutions confirms the aforementioned argument, that the effective area of the normal region between the superconductors is larger than $d_N d$. This renormalization is reduced at higher temperature and vanishes at $T \gg E_{Th}$.

The form of the distribution function implies that in the energy window $-V/2 < E < V/2$ current-carrying-states are depopulated and the supercurrent (6.42) is strongly suppressed. For $V < E_g$ this has no effect because of the gap. At higher voltage the supercurrent decays rapidly with increasing voltage (cf. Fig. 44). This is exactly what has been observed in the experiments [16]. Furthermore, at still larger voltage $eV \geq 10E_{Th}$ the dominant contribution to the integral in (6.42) outside this energy window is *negative* (see Fig. 43,44), so the total supercurrent changes sign. We thus find a transition to a π -junction [84], controlled by nonequilibrium effects. This prediction has been verified in recent experiments [17, 18]. The effect is rather pronounced, the critical current of the π -junction is approximately 30% of I_c at $T = eV = 0$.

For high voltages or temperatures $eV, T \gg E_g$ we find a closed expression for the current (6.42). The integral is dominated by the poles of f_L , given by (6.33), at imaginary frequencies $E = i\omega_n \pm V/2$. Also j_E can be found by analytic continuation of the approximation (6.74). We thus find $I_S = I_c \sin \phi$, with critical current

$$I_c R_d = \frac{64\pi}{3 + 2\sqrt{2}} e^{-\sqrt{\Omega_V/E_{Th}}} \cos\left(\frac{eV}{2\sqrt{\Omega_V E_{Th}}} + \varphi_0\right) \begin{cases} eV & \text{for } eV \gg E_g; \quad T \ll E_g \\ T \left(\frac{(2\pi T)^2 + e^2 V^2}{E_{Th}^2}\right)^{\frac{1}{4}} & \text{for } T \gg E_g \end{cases} ; \quad (6.80)$$

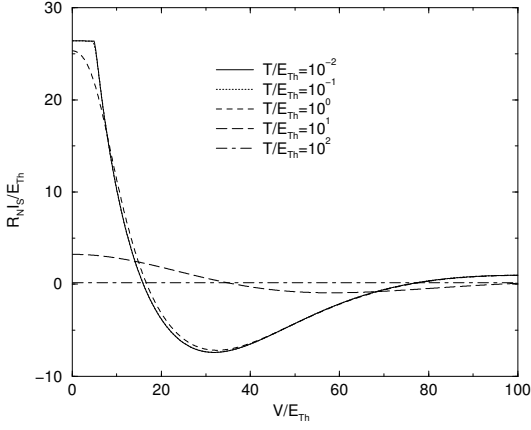


Figure 45: The supercurrent as a function of the control voltage for various values of T at $\phi = \pi/2$.

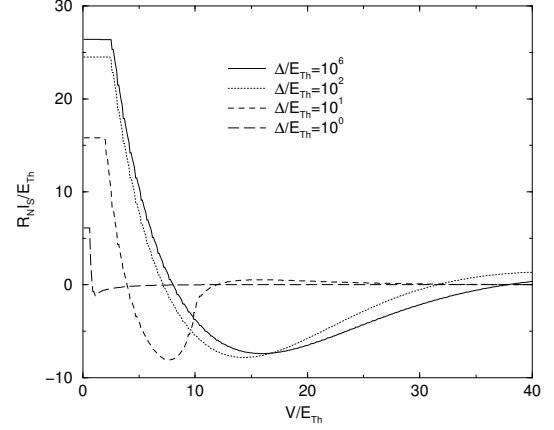


Figure 46: The supercurrent as a function of the control voltage (at $T = 0$) for various values of Δ/E_{Th} at $\phi = \pi/2$.

Here $\Omega_V = \pi T + \sqrt{(\pi T)^2 + (eV/2)^2}$ and

$$\varphi_0 = \begin{cases} \pi/2 & \text{for } eV \gg E_g, T \ll E_g \\ \frac{1}{2} \tan^{-1} \left(\frac{eV}{2\pi T} \right) & \text{for } T \gg E_g \end{cases} \quad (6.81)$$

This result indicates that with increasing temperature the onset of the π -junction behavior is weakened and shifted to higher voltages. Our numerical results, shown in Fig. 45, demonstrate that it is still pronounced as long as $T \leq E_{Th}$.

Above we have discussed the limit of $\Delta \gg E_{Th}$. For smaller values of Δ the minigap E_g is limited by Δ rather than E_{Th} , see Fig. 46. However, the qualitative behavior, including the transition to the π -junction regime persists.

6.4.2 Dissipative current in a mesoscopic wire

Perhaps the simplest system displaying the influence of the proximity effect on transport properties is shown in Fig. 47: A mesoscopic wire connects a superconducting and a normal reservoir. At the interfaces additional barriers (indicated by r in Fig. 47) may exist. In this section we assume no such barriers to be present. We assume the superconductor is assumed to be at voltage $V = 0$. As the dissipative current is a genuine nonequilibrium quantity, we again will use the Keldysh technique. In addition to the previously used boundary condition at the superconductor, we impose $\theta = 0$ at the normal reservoir. Hence, the spectral supercurrent j_E vanishes at the normal lead. As the retarded Usadel equation (6.50) conserves this spectral current, this implies that $j_E = 0$ everywhere and hence $\partial_x \chi = 0$. This means that no ‘supercurrent’ flows in this system. On the other hand, the ‘normal’ current j_n is influenced by superconducting correlations, and we have to study their behavior in an electric field.

Approximate solutions of (6.50) can be obtained by the same methods as described above. We will now concentrate on the case $E_{Th} \ll \Delta$. For $E \ll E_{Th}$, iterating (6.50) from the $E = 0$ -solution yields

$$\theta = -\frac{8}{\pi^2} \frac{E}{E_{Th}} \left[\bar{x} - \sin \left(\frac{\bar{x}\pi}{2} \right) \right] - i \frac{\pi}{2} \bar{x} \quad (6.82)$$

where $\bar{x} = x/d$. For $E \gg E_{Th}$ we have

$$\tanh \left(\frac{\theta}{4} \right) = -\tanh \left(\frac{i\pi}{8} \right) \exp[k(x-d)] \quad (6.83)$$

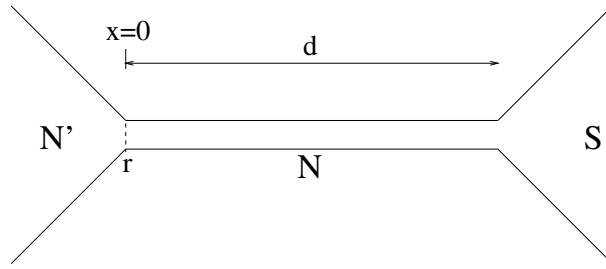


Figure 47: A thin normal wire between normal and superconducting reservoirs.

with $k = \sqrt{-2iE/\mathcal{D}}$. For the calculation of the dissipative current we only need the distribution function f_T . As $j_E = 0$, the kinetic equations decouple and (6.34) reduces to a simple spectral diffusion equation. Relating f_T to the potential and $\mathcal{D}_T$ to the conductivity, we note that our expressions (6.43) and (6.34) are analogous to Ohm's law and the continuity equation on a spectral level, respectively. It is now straightforward to integrate (6.34) and to calculate the current (6.43), from which we obtain the differential conductance

$$G = \int_{-\infty}^{\infty} \frac{dE}{2T \cosh^2[(E - V)/2T]} g(E). \quad (6.84)$$

Here the spectral conductance $g(E)$ is defined by

$$g(E) = G_N d \left(\int_0^d \frac{dx}{\mathcal{D}_T(E, x)} \right)^{-1}. \quad (6.85)$$

Again this expression is the spectral form of Ohm's law. Using the approximate solutions (6.82) and (6.83), we calculate $\mathcal{D}_T$ and $g(E)$, and arrive, e.g. at $V = 0$, at

$$\frac{G}{G_N} = 1 + \begin{cases} a T^2/E_d^2 & \text{for } T \ll E_{Th} \\ b \sqrt{E_{Th}/T} & \text{for } T \gg E_{Th} \end{cases} \quad (6.86)$$

with $a \approx 0.049$ and $b \approx 0.42$. Results are displayed in Fig. 48. The conductance approaches the normal-state value for $T \gg E_{Th}$, because in this limit the proximity effect is confined to a short portion of the wire of length ξ_N . However, also at $T \rightarrow 0$ the correction to the conductance vanishes. This has been first noted by Artemenko *et al.* [85] for short contacts, and later in longer wires by Nazarov and Stoof [20, 86, 19]. The theoretical prediction was confirmed in experiments [22, 21]. To the best of our knowledge, no simple physical explanation of this surprising coincidence has been provided yet. The fact that the conductance is unchanged does *not* imply that the proximity effect is absent. Even at $E = 0$ the solution (6.82) yields $F^R \neq 0$, and the density of states $\text{Re}(G^R)$ displays a proximity-induced gap-structure, as shown in Fig. 49.

At $E = 0$, the local spectral diffusion coefficient $\mathcal{D}_T$ coincides with the normal-state value everywhere, see Eq. (6.82). Returning to the definition (6.37) and using the parameterization (6.49) we find here [19]

$$\mathcal{D}_T = (\text{Re}G^R)^2 + (\text{Im}F^R)^2 = \cosh^2(\text{Re}(\theta)) \geq 1. \quad (6.87)$$

This means, that it is a combination of the density of states and the Maki-Thompson paraconductivity term [47, 87]. Both terms compete. At finite E , the Maki-Thompson contribution is stronger than the suppression of the DOS [19]. At $E = 0$ both contributions, although each of them varies in space, add up to 1.

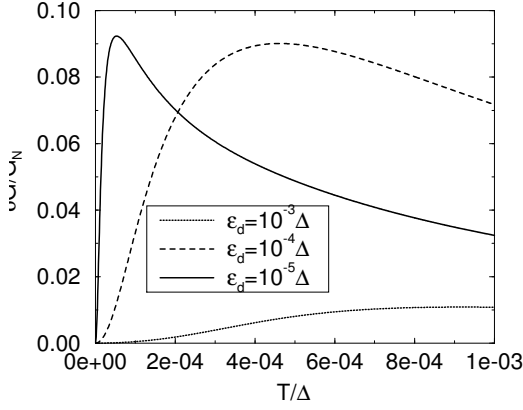


Figure 48: Linear conductance of a proximity wire of length d with transparent interfaces. The energy is measured in units of the Thouless energy $E_{\text{Th}} = D/d^2$ associated with the length d of the wire.

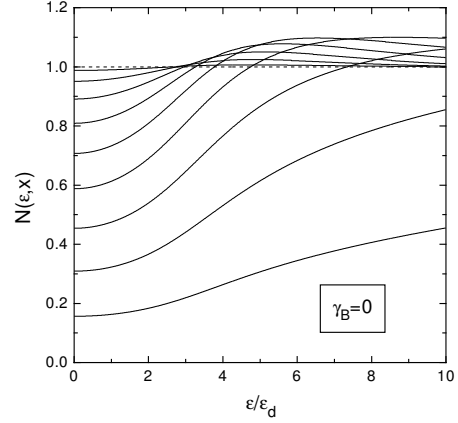


Figure 49: Local density of states in a proximity wire with transparent interfaces for $\Delta \gg E_{\text{Th}}$ at distances $x/d = 0, 0.1, \dots, 1.0$ (top to bottom at low E) from the normal reservoir.

In order to study the space-dependence more thoroughly, we look at the local DOS in our model. Our results (49) indicate, that due to the induced superconductivity, a pseudo-gap in the DOS opens, i.e. the DOS is in general suppressed but does not vanish at low energies. If we move away from the superconductor, the pseudo-gap is weakened and the DOS approaches its normal-state form when we are close to the normal reservoir.

6.4.3 Supercurrent and interference effects on the dissipative current

We will extend our study of the dissipative current to a situation where also a supercurrent can flow and therefore interference of the induced correlations starts to play a role. This type of effects was first studied experimentally [25] in cross-shaped diffusive metal structures with different superconducting leads closed by a SQUID loop, later also in analogous semiconducting structures [88], as well as in metallic structures with an Aharonov-Bohm loop [23, 89]. Usually, these structures are referred to as “Andreev interferometers”. The common key observation is, that the conductance depends periodically on the magnetic flux penetrating the loop with a period of a superconducting flux quantum $\phi_0 = h/2e$. Moreover, it turned out that these oscillations could be detected even if $\xi_N \ll d$ where d is a typical geometrical sample dimension. I.e., the proximity effect had a longer range than expected from simpler arguments.

These effects have been extensively studied theoretically [31, 20, 86, 19, 90, 91]. We want to outline the main concepts in these systems, considering as an example the experiments of Ref. [23]. The layout of the system is shown in Fig. 50. Applying the arguments of the previous chapter, we recognize that the supercurrent is confined to the loop. Instead of dealing with the vector potential in the derivatives in (6.21), we remove it by a gauge transformation $\chi = \chi_0 - 2e \int_0^R dr' \mathbf{A}(r')$ where the integral runs around the loop. This implies, that the vector potential produces an artificial phase difference of $\phi = \Phi/\Phi_0$ (modulo 2π) at one point in the loop, which we choose to be point B . This procedure already guarantees the $h/2e$ -periodicity of the results.

We know from the previous sections the form of the Usadel equations in the single branches.

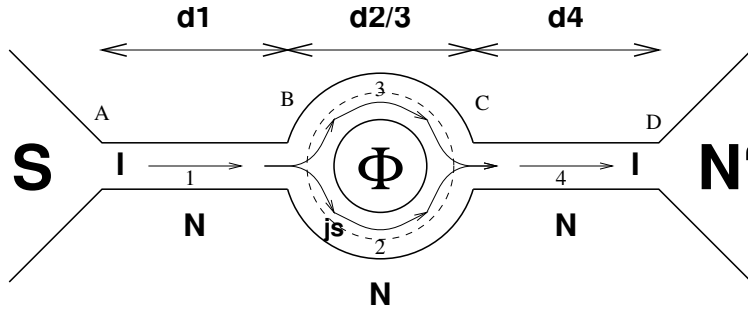


Figure 50: A proximity wire containing a loop. The cross-section areas of the respective components are A_i . j^S is the field-induced supercurrent in the ring part.

In the present problem we also have to account for the nodes. To do so we apply the idea of Zaitsev [92]: By regarding the Green's functions in the connecting wires as cross-sectional averages of the full 3D Green's function, which have well-defined derivatives in space, we will conserve the “spectral flow”. This means

$$\check{G}_n = \check{G}_m \quad (6.88)$$

for all branches n and m and at nodal points we have

$$\sum_n A_n \check{G} \frac{\partial}{\partial x_n} \check{G} = 0, \quad (6.89)$$

where n labels the branches, and the derivatives point in the direction of branch n away from the node. This approach has been used by Nazarov [93, 94] to derive a simple circuit theory for $T = V = 0$. Here, we will assume that the branches of the loop have *half* of the cross-section area of the other branches.

Only very few analytical results have been derived for this system [19, 90]. Here we will focus on numerical results. As shown in Fig. 51, the spectral conductance shows a non-monotonic

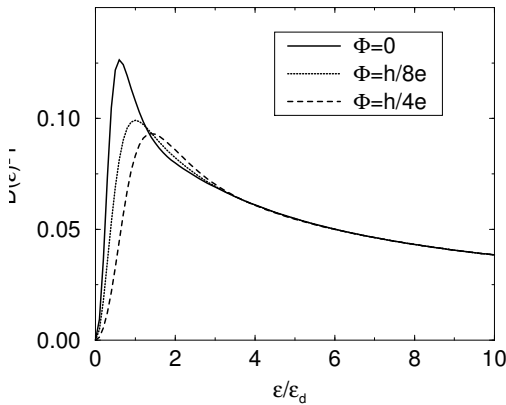


Figure 51: Flux-dependent spectral conductance of the loop structure, assuming $A_{1/4} = 2A_{2/3}$ all $d_i = d$ and $\Delta \gg E_d = D/d^2$.

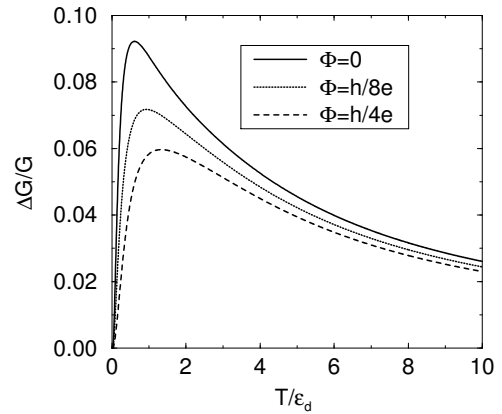


Figure 52: Linear-response conductance of the loop as described in the caption of Fig. 51.

shape as in the previous section. Provided that the induced correlations persist from the superconducting electrode to the loop, the low energy results, for $E \ll E_{Th}$, depend on the flux, whereas the conductance at $E \gg E_{Th}$ is flux-independent, see Fig. 51. This agrees with the

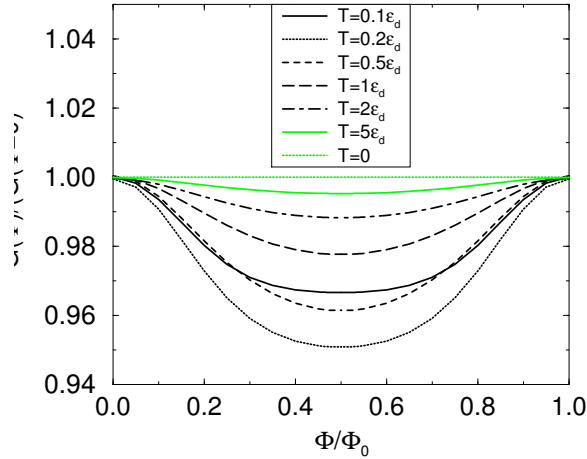


Figure 53: Full period of normalized conductance oscillations in the loop structure at $V = 0$.

intuition on the range of the proximity effect developed in the previous sections. In the loop, the vector potential causes the correlations to interfere destructively, so the induced conductance correction is suppressed by the presence of the loop. At $E \gg E_{\text{Th}}$, the correlations have already decayed before they reach the loop and are therefore not influenced by the flux. On the other hand the temperature dependent conductance also depends notably on the flux at $T > E_{\text{Th}}$, see Fig. 52, so we indeed recover the observed long range proximity effect there [23, 24, 25]. These seemingly paradoxical observation can be resolved by examining (6.84). In order to calculate the conductance Fig. 52 from the spectral conductance Fig. 51, we have to convolute the latter with the derivative of the Fermi function. This distribution function is peaked around $E = 0$ and has the width T , i.e. the window of size E_{Th} around the Fermi surface where the spectral conductance is actually flux-dependent with a relative weight of E_{Th}/T . Thus, the conductance depends sensitively on the contribution of this low-energy range, where the proximity effect has a long range [23], which is only limited by the phase-breaking length l_ϕ .

We note that the intuition on ξ_N developed by regarding thermodynamic quantities as the supercurrent in an SNS junction (section 6.3.3) fails when describing nonequilibrium quantities. On the other hand, we have shown that the same microscopic mechanism (characterized by the same Green's functions) may lead to the exponential temperature dependence of equilibrium quantities as well as to the “long range” tail with power-law temperature dependence of nonequilibrium phenomena.

6.4.4 Transport through tunneling barriers

So far, we only considered situations, where the current flow does not have to overcome tunneling barriers. We found, that the local spectral diffusion coefficient $\mathcal{D}_T = \cosh^2(\text{Re}\theta) \geq 1$ and thereby the global conductance is always *enhanced* in comparison to the normal state. However, as we explained before (section 6.3.1), the tunneling conductance is proportional to the density of states, i.e. at low energies, $E < E_{\text{Th}}$, it is suppressed in comparison to the normal state as a manifestation of the induced superconducting gap [82, 8, 9]. In general, these two mechanisms compete. Moreover, due to the direct coupling to the normal reservoir, correlations immediately feel the presence of normal electrons and only a pseudo-gap develops, which turns into a real minigap when the reservoir is decoupled from the system via a tunneling barrier [19, 95].

Following the lines of Ref. [96] with boundary condition (6.54), we can calculate the distribution functions and from there the current and conductance. The result has the form (6.84), however, the spectral conductance now accounts for the series of the tunneling resistance and

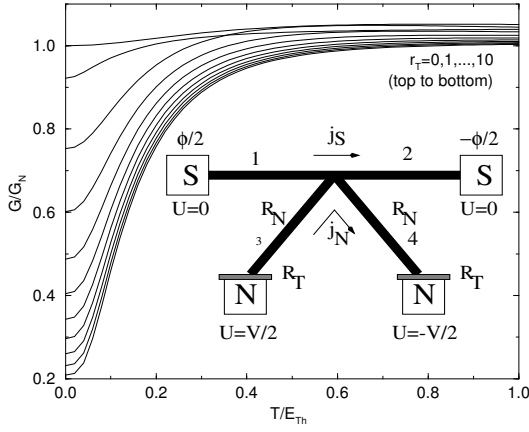


Figure 54: Linear conductance as a function of temperature at different tunneling resistances within the diffusion/tunneling crossover at $\phi = 0$. At $\phi = \pi$ we have $G = G_N$. The inset shows the structure used for calculations. All arms 1–4 are assumed to have the same lengths d and cross sections A .

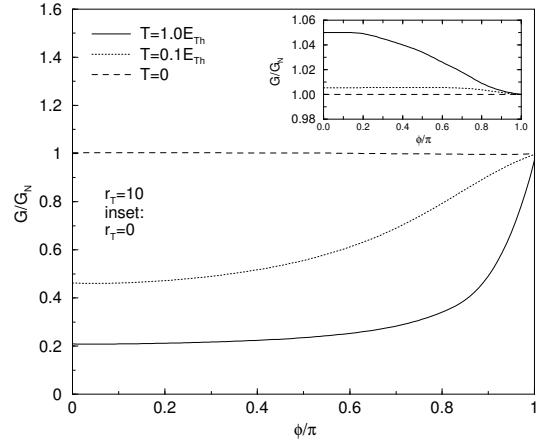


Figure 55: Linear conductance of the interferometer in Fig. 54 as a function of the phase in the tunneling-dominated regime; inset: diffusion-dominated regime.

the diffusive metal

$$g(E) = G_N \frac{1 + r_T}{M(E) + r_T/\nu_T(E)}. \quad (6.90)$$

Here $r_T = R_T/R_N$, where R_N is the normal state resistance of the wire, and $\nu_T = \text{Re}\{G^R\}$ is the reduced DOS at the interface. This result implies that already a rather low barrier of size $\sim R_N$ gives rise to a tunneling-like structure. In typical experiments, this is of order $1 \dots 10\Omega$. In Fig. 54 we show the conductance of a structure relevant for experiments [95] demonstrating that at low T for $r_T > 1$ the conductance is reduced compared to the normal state value.

As we showed in section 6.4.3, a magnetic flux through an Andreev interferometer causes interference effects which influence the conductance. This interference turns out to be destructive. For the system without barriers studied in section 6.4.3, this manifests itself in the fact that the conductance oscillations start from a zero-flux maximum. If the interferometer contains tunneling barriers, we expect that the interference effects inside the interferometer are qualitatively not affected, however the conductance eq. (6.90) is dominated by tunneling into the induced gap. The destructive interference will weaken induced superconductivity, therefore lift the gap and consequently increase the conductance. Thus in this system we also expect magneto-conductance oscillations as described in section 6.4.3, however, they will depart from a zero-flux minimum. By combining the theory of the interferometer developed in section 6.4.3 with the conductance formula (6.90), we can study these structures and confirm this picture [95, 97]. In Fig. 55 we show a half-period of the conductance oscillations at different temperatures for a tunneling-dominated system in comparison to the same system probed through metallic contacts which have a different sign of the flux-dependence. This effect has recently also been observed in experiments [95].

6.4.5 Four-Point conductance

In the previous sections we argued, that the proximity effect increases the conductivity of a normal metal wire compared to the normal state. If the current is probed through tunneling

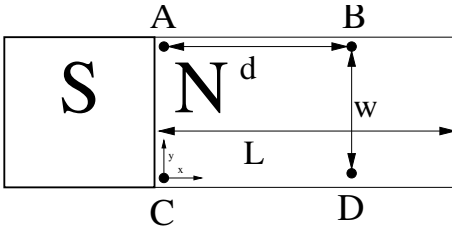


Figure 56: A quasi-2D proximity film. The contacts A and B are used as voltage and C and D as current probes. An alternative setup: A and C are voltage probes, while the current flows through B and D.

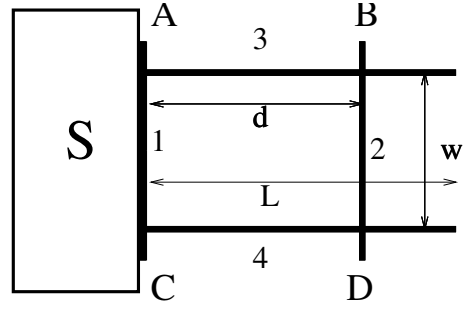


Figure 57: A qualitatively equivalent circuit with the probe configuration as in Fig. 56.

barriers, this increase competes with the opening of a superconducting gap which can cause the total conductance to decrease. However, a decrease in the conductance was also observed in experiments [24, 25] where the current does not have to pass through tunneling barriers. This effect *cannot* be explained by the non-monotonic conductance described in section 6.4.2, because despite its T-dependence, the correction to the normal state conductance is still non-negative.

In the experiments[24, 25], the current was probed by a particular four-probe configuration on a sample of appreciable width. Hence the current flow cannot be described by simple quasi-1D models as above. We will now develop a quasiclassical description of such measurements and show how the inhomogenous conductivity of the sample leads to a decrease of the four-point conductance.

Let us consider a normal film connected to a superconductor with four point- like contacts, which lead into measuring reservoirs after a distance much larger than l_ϕ , see Fig. 56. In this configuration, the probes do not influence the spectral properties, so the solution of the retarded Usadel equation is analogous to the quasi-one dimensional case with an infinite tunneling barrier separating the normal reservoir from the sample

$$\partial_x \theta(x = d) = 0 . \quad (6.91)$$

At large energies, θ is small and we reproduce (6.83). At low energies, a perturbative expansion gives

$$\theta_E(x) = \frac{E}{E_L} \frac{x}{L} \left(2 - \frac{x}{L} \right) - i\pi/2 . \quad (6.92)$$

The kinetic equation is now two-dimensional where the current contacts are represented by source terms

$$\nabla(\mathcal{D}_T(r)\nabla f_T) = I_E(\delta(r - r_C) - \delta(r - r_D)) , \quad (6.93)$$

A “no current flow” condition at the normal metal edges yields

$$\partial_n f_T = 0 . \quad (6.94)$$

A formal solution of Eq. (6.34) reads

$$f_T(E, r) = I_E(\mathcal{G}_E(r, r_C) - \mathcal{G}_E(r, r_D)) , \quad (6.95)$$

where $\mathcal{G}_E = (\mathcal{D}_T \nabla + \mathcal{D}_T \nabla^2)^{-1}$ is the Green’s function of the operator (6.34). Making use of (6.33, 6.94) and (6.95), and integrating I_E over energy we obtain the total current I and arrive at the expression for the differential four-point-conductance $G = dI/dV$:

$$G(V, T) = \int_0^\infty \frac{g(E)}{2T \cosh^2((E - V)/2T)} dE , \quad (6.96)$$

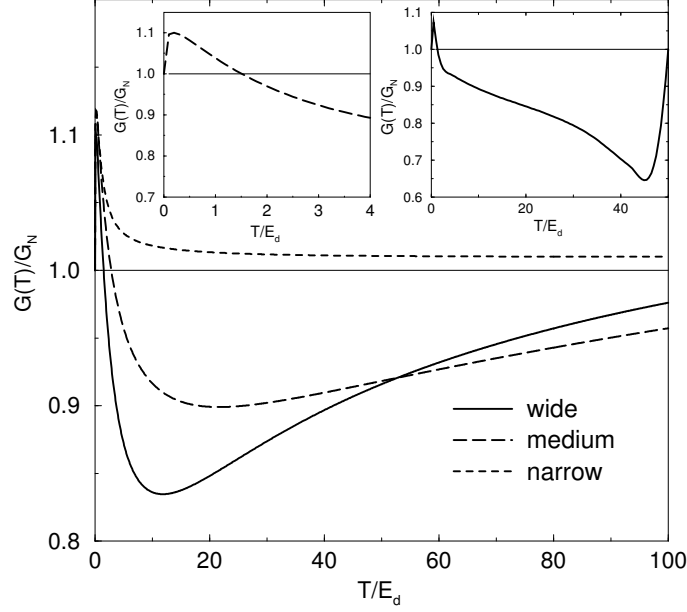


Figure 58: The linear conductance $G(T)$ for films of different widths $w/d = 0.05, 0.5, 1.0$ calculated for $d = L$ and $T_c = 5.7 \cdot 10^5 E_{Th}$. Left inset: The same curve at $w = 0.5d$. The T -axis is zoomed to demonstrate the presence of a usual 1D-type non-monotonic behavior at $T \sim E_{Th}$. Right inset: $G(T)$ for a wide film and $T_c = 50 E_{Th}$. The amplitude of the negative conductance peak is increased due to the effect of a superconducting gap $\Delta(T)$.

where

$$g(E) = G_N \frac{\mathcal{G}_0^{BC} - \mathcal{G}_0^{BD} - \mathcal{G}_0^{AC} + \mathcal{G}_0^{AD}}{\mathcal{G}_E^{BC} - \mathcal{G}_E^{BD} - \mathcal{G}_E^{AC} + \mathcal{G}_E^{AD}} \quad (6.97)$$

is the spectral conductance. We introduced the notation $\mathcal{G}^{ij} = \mathcal{G}(r_i, r_j)$ and $\mathcal{G}_0$ is the Green's function of (6.34) in the normal state ($M_E(r) = \sigma_N$). The spectral conductance (17.19) calculated numerically from Eqs. (6.37), (6.93) and (6.94) is presented in Fig. 58.

For narrow films the well-known results of quasi-1D calculations [20, 19] are qualitatively reproduced: the linear conductance $G(T)$ exceeds G_N at all temperatures, showing a non-monotonic feature at $T \leq E_d$ (for simplicity we put $L = d$ here and below). For broader films $G(T)$ decreases below the normal-state value at high temperatures and reaches the minimum at $T \sim 10 E_{Th}$. At lower T the conductance grows with decreasing T , becomes larger than G_N and then decreases again down to $G(T = 0) = G_N$ similarly to the 1D case (see the left inset in Fig. 58). The behavior of $g(E) \equiv G(E, T = 0)$ as a function of energy (voltage) is qualitatively identical to that of $G(T)$, the negative peak at $E \sim 10 E_{Th}$ turns out to be even somewhat deeper.

In order to develop a semi-quantitative understanding of these results, let us consider a wire network representing the interconnections between the probes as in fig. 57. A similar equivalent-circuit model was previously used for a qualitative description of inhomogeneous superconducting films [98]. Exploiting the analogy between f_T and the electrical potential in a conventional circuit (see above and [26]), Kirchhoff's laws for the spectral conductances can be derived [19, 92]. For the present circuit Sig. 57, we find (c.f. [98])

$$g_{Net} = g_3 g_4 \sum_{i=1}^4 g_i^{-1}, \quad (6.98)$$

where the g_i are the spectral conductances [96, 19] of the wires 1–4 calculated analogous to

previous sections. At $T \gg E_{\text{Th}}$ only the wire 1 directly attached to a superconductor acquires superconducting properties, whereas the proximity effect in the wires 2, 3 and 4 is suppressed. Thus only g_1 increases, and $g_{2,3,4}$ remain unaffected. According to Eq. (6.98) g_{Net} decreases below G_{N} . At $T \approx E_{\text{Th}}$ the proximity induced superconducting correlation penetrates into all four wires, i.e. $g_{2,3,4}$ increase, leading to the increase of g_{Net} above G_{N} . Close to Δ , the conductance correction has a peak corresponding to the peak in the DOS. If close to T_c , Δ becomes close to E_{Th} , the conductance of interconnection 1 is peaked and the whole effect becomes even stronger, in agreement to our numerical result.

A semi-quantitative analysis [26] shows, that the crossover energy between excess and deficit conductance is of the order of $\max(D/d^2, D/W^2)$. A careful analysis of the experiments [24, 25] shows that this explains the different signs of the conductance for different measurements [26].

6.5 Conclusion

We outlined the concepts of quasiclassical Green's functions. They provide a unified view of ballistic and diffusive systems in equilibrium and nonequilibrium situations. We applied the theory to describe mesoscopic proximity structures.

7 Boundary Green functions of topological insulators and superconductors by Peng, Bao

Abstract

Topological insulators and superconductors are characterized by their gapless boundary modes. In this paper, we develop a recursive approach to the boundary Green function which encodes this nontrivial boundary physics. Our approach describes the various topologically trivial and nontrivial phases as fixed points of a recursion and provides direct access to the phase diagram, the localization properties of the edge modes, as well as topological indices. We illustrate our approach in the context of various familiar models such as the Su-Schrieffer-Heeger model, the Kitaev chain, and a model for a Chern insulator. We also show that the method provides an intuitive approach to understand recently introduced topological phases which exhibit gapless corner states.

7.1 Conclusion

The boundary Green function is a natural quantity to characterize topological insulators and superconductors as it directly encodes the boundary modes and hence the topological phase diagram as well as the topological indices. In this paper, we have developed a systematic approach to compute boundary Green functions of topological insulators and superconductors by relying on a recursive approach. We show for several familiar models that our approach can be applied in a rather straight-forward manner.

Though simpler, our recursive approach has some similarities with a real-space renormalization group approach. For the noninteracting fermion lattice problems discussed in this paper, we find that the recursion relation can be solved explicitly in terms of the transfer matrix which greatly facilitates analytical calculations. When attaching leads to the sample, we can relate the boundary Green function to the reflection matrix. In view of earlier work on deriving the periodic table of topological phases of free fermions from the reflection matrix [9], this provides us with explicit expressions for the topological indices in terms of the boundary Green functions.

As an interesting application, we show that for insulators with trivial bulk topology, the boundary Green function defines a boundary Hamiltonian in a natural manner. Remarkably, even if the bulk is nontopological, this boundary Hamiltonian can still be topologically nontrivial which, say in 2D, is reflected in gapless corner states. We show this explicitly for a model which was recently proposed. This not only provides an intuitive approach to these recently introduced topological systems, but also has some potential to serve as a starting point for a more systematic investigation of these higher-order topological insulators.

While we restricted ourselves to clean and noninteracting models, the recursive approach should be extendable to include disorder and possibly interactions. For instance, in the presence of disorder, the recursion for the boundary Green function remains valid. While for a fixed disorder configuration, the boundary Green function would no longer approach a fixed-point Green function in the thermodynamic limit, the distribution of boundary Green functions over the disorder ensemble should still have such a fixed point.

7.2 Introduction

There is currently considerable interest in topological phases of matter and topological quantum phase transitions. The best studied and simplest variety of topological phases are the independent-fermion phases [1, 2, 3] collected and classified in the periodic table and its extensions [4, 5, 6, 7, 8, 9, 10]. These topological phases can be characterized in terms of bulk topological indices and, by the bulk-boundary correspondence, exhibit gapless modes localized at the sample boundaries.

Here, we present a simple and intuitive recursive approach which provides access to these topological phases and their transitions. Our approach focuses on the boundary Green function [42, 12] which encodes the presence or absence of gapless surface modes in an immediate manner. Focusing on tight-binding models, we imagine that the bulk system is constructed by iteratively adding boundary sites (in one dimension) or boundary layers (in higher dimensions). We can then exploit that the boundary Green function can be computed recursively and, by virtue of the bulk gap, becomes invariant under this recursion in the thermodynamic limit. Thus, the boundary Green function can be obtained as fixed points of the recursion.

For a typical model which exhibits topologically trivial and nontrivial phases, the recursion will generate several fixed-point boundary Green functions. The physical boundary Green function and hence the phase diagram can be identified from a stability analysis. The various phases correspond to the stable fixed points of the recursion and the approach to the fixed point contains information about the spatial localization of the boundary mode away from the boundary as well as the bulk gap. This has some similarities with a real space renormalization group approach.

The fixed-point boundary Green function provides access to various observables of interest such as the tunneling density of states. Most importantly, it immediately contains information on the topological invariants by virtue of the fact that it encodes the full boundary spectrum. We can relate our approach more systematically to the general classifications of free-fermion topological phases by establishing a connection to their classification in terms of scattering matrices [9]. This can be done by expressing the reflection matrix describing reflection from a gapped phase in terms of its boundary Green function. In this way, we immediately obtain explicit expressions for the topological invariants in terms of the boundary Green function.

Our approach can be applied to a wide variety of models. Here, we illustrate the method on a number of popular models. For simple cases such as the Su-Schrieffer-Heeger model [13], the recursion is conveniently implemented in a direct manner. To make the approach more systematic and to simplify its implementation, we find it helpful to introduce an approach based on the transfer matrix [14, 15, 16, 17]. This is illustrated in one dimension for the Kitaev

chain [18] and in two dimensions for a model of a Chern insulator [3].

Very recently, it was pointed out [19] that some 2D and 3D models exhibit topologically protected excitations at corners. These models have boundary Green functions which are regular, reflecting the fact that the boundary is gapped. Nevertheless, it is natural to interpret the inverse of the boundary Green function at the Fermi energy as an effective boundary Hamiltonian which can by itself be topologically trivial or nontrivial. We show for a model discussed in Ref. [19] that this effective boundary Hamiltonian is indeed topologically nontrivial when the model exhibits topological corner states. Our boundary Green function method may thus open the path to a more systematic analysis of such higher-order topological phases.

The paper is organized as follows. In Sec. 7.3, we introduce the principal idea of this work and illustrate it with the Su-Schrieffer-Heeger model. We also show how to relate the boundary Green function to the reflection matrix which connects our approach with the systematic classification of free-fermion systems. In Sec. 7.4, we show how the transfer matrix can be employed for computing the fixed-point boundary Green function. This is used in Sec. 7.5 to compute the boundary Green function for the Kitaev chain and a model of a Chern insulator. For both models, we show how the topological invariants can be extracted from the fixed-point boundary Green function. A 2D model with topological corner states is investigated in Sec. 7.6. We conclude in Sec. 7.1. Some aspects are relegated to appendices.

7.3 Basic considerations

7.3.1 Boundary Green function

Consider a d -dimensional tight-binding Hamiltonian, which is translationally invariant along $d - 1$ directions. Assuming periodic boundary conditions in these directions, we can define a $(d-1)$ -dimensional crystal momentum $\mathbf{k}_\perp$ to label the Hamiltonian. In the remaining dimension, we apply open boundary conditions and retain a real-space representation. By choosing a sufficiently large unit cell, a Hamiltonian with a finite hopping range can always be brought into a form which couples only neighboring unit cells. Thus, without loss of generality, we can consider a family of quasi-1D Hamiltonians which are labelled by the transverse momentum $\mathbf{k}_\perp$ and take the form

$$H(\mathbf{k}_\perp) = \sum_{n=1}^N \left(\psi_n^\dagger h_n \psi_n + \psi_{n+1}^\dagger V_n \psi_n + \psi_n^\dagger V_n^\dagger \psi_{n+1} \right). \quad (7.1)$$

Here, $n = 1, 2, \dots, N$ labels the unit cells along the direction with open boundary conditions, and ψ_n ($\psi_n^\dagger$) is a column (row) vector of electron annihilation (creation) operators in the n th unit cell. These vectors have M entries, reflecting the dimension of the local Hilbert space associated with each unit cell and a given $\mathbf{k}_\perp$. The intra- and inter-unit-cell couplings are described by $M \times M$ matrices h_n and V_n , respectively. Note that we suppressed the $\mathbf{k}_\perp$ dependence of ψ_n , h_n , and V_n for notational simplicity. The structure of this quasi-1D tight-binding Hamiltonian is illustrated in Fig. 59.

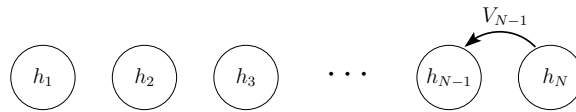


Figure 59: Schematic representation of the quasi-one-dimensional Hamiltonian with N unit cells.

More explicitly, the corresponding first-quantized Hamiltonian can be written as an $NM \times$

NM block tridiagonal Hamiltonian matrix $\mathbf{H}(\mathbf{k}_\perp)$ which takes the form

$$\mathbf{H} = \begin{pmatrix} h_1 & V_1^\dagger & & & \\ V_1 & h_2 & V_2^\dagger & & \\ & \ddots & \ddots & \ddots & \\ & & V_{N-2} & h_{N-1} & V_{N-1}^\dagger \\ & & & V_{N-1} & h_N \end{pmatrix}. \quad (7.2)$$

Both the spectrum and the eigenfunctions of the system are encoded in the Green function (or resolvent operator) $\mathbf{G}(\omega)$ of this Hamiltonian, which is defined through

$$(\omega \mathbb{I} - \mathbf{H})\mathbf{G}(\omega) = \mathbb{I}. \quad (7.3)$$

Here, $\mathbb{I}$ is the identity matrix of dimension NM .

The low-energy excitations of topological insulators and superconductors are localized near their boundary. This motivates us to consider the boundary Green functions $G_1 = \mathbf{G}_{11}$ and $G_N = \mathbf{G}_{NN}$, which correspond to the $1, 1$ and N, N blocks of the full Green function matrix $\mathbf{G}$. Thus, these boundary Green functions are $M \times M$ matrices.

By virtue of the tridiagonal structure of the first-quantized Hamiltonian (7.2), we can compute the boundary Green function recursively, extending the system unit cell by unit cell in the direction with open boundary conditions. In this way, we can relate the boundary Green function G_N of a chain with N unit cells to the boundary Green function G_{N-1} of a chain with $N - 1$ unit cells. Indeed, the two boundary Green functions are simply related by the Dyson equation

$$(g_N^{-1} - V_{N-1}G_{N-1}V_{N-1}^\dagger)G_N = \mathbb{I}. \quad (7.4)$$

Here, $g_N^{-1}(\omega) = \mathbb{I}\omega - h_N$ is the bare Green function of the N th unit cell. Similar recursions were used in the context of topological phases in Refs. [20, 21].

When considering insulators and superconductors with a gapped bulk, be they topological or not, we expect that the boundary Green function becomes independent of the number of layers in the limit of large N . Hence, the boundary Green function should approach a fixed point G which satisfies the fixed-point equation

$$(g^{-1} - VGV^\dagger)G = \mathbb{I}. \quad (7.5)$$

Here, we have assumed that the system is uniform along the chain, so that we can drop the indices of g and V . This closed equation for the boundary Green function (a matrix quadratic equation) is the central starting point of our considerations.

7.3.2 The Su-Schrieffer-Heeger (SSH) model: An illustrative example

It should be evident that the boundary Green function encodes much information of interest in the context of topological insulators and superconductors. Most importantly, it knows about gapless end, edge, or surface states, and can thus be used to derive their properties (including their extension into the bulk) as well as the topological phase diagram of the model. It is useful to illustrate this statement in the context of the Su-Schrieffer-Heeger model. This model is well known to exhibit topologically distinct phases and is sufficiently simple so that we can implement the recursive approach in a rather straight-forward manner, without relying on the more powerful methods developed in the next sections.

The Su-Schrieffer-Heeger (SSH) model describes a 1D chain of spinless fermions with one orbital per site and alternating hopping strengths between neighbouring sites. For a chain with $2N$ sites and open boundary conditions, the Hamiltonian takes the form

$$H = \sum_n \left(-t_1 c_{2n}^\dagger c_{2n-1} - t_2 c_{2n+1}^\dagger c_{2n} + \text{h.c.} \right), \quad (7.6)$$

where c_j ($c_j^\dagger$) annihilates (creates) a fermion at site j , and t_1 and t_2 are the hopping amplitudes (taken to be real for simplicity). While the model is gapless for $|t_1| = |t_2|$, the dimerization opens a gap around $E = 0$ when the hopping amplitudes differ in magnitude. The gapless point separates a topological phase for $|t_1/t_2| < 1$ from a trivial phase for $|t_1/t_2| > 1$. In the topological phase, there is a midgap state localized at each end of the chain in the thermodynamic limit.

For the SSH model, we have $g_N^{-1} = \omega$ and

$$V_N = \begin{cases} t_1 & \text{odd } N \\ t_2 & \text{even } N \end{cases} \quad (7.7)$$

from Eq. (7.6). In view of the dimerization of the hopping amplitudes, it is convenient to iterate the recursion Eq. (7.4) and to obtain a recursion for chains with an even number of sites,

$$G_{2N} = [\omega - t_1^2(\omega - t_2^2 G_{2N-2})^{-1}]^{-1}. \quad (7.8)$$

To analyze this recursion, we write it as

$$G_{2N} - G_{2N-2} = \beta(G_{2N-2}) \quad (7.9)$$

with the β -function

$$\beta(x) = [\omega - t_1^2(\omega - t_2^2 x)^{-1}]^{-1} - x. \quad (7.10)$$

The fixed-point boundary Green function G follows from the zeroes of the β -function. Solving the corresponding quadratic equation, we find the two solutions

$$G_{\text{triv}} = \frac{\omega}{t_2^2 - t_1^2} + O(\omega^3), \quad (7.11)$$

$$G_{\text{top}} = \frac{t_2^2 - t_1^2}{t_2^2 \omega} + O(\omega), \quad (7.12)$$

where we have restricted attention to the limit of small ω .

These two fixed-point boundary Green functions correspond to the topological and nontopological phases of the model. In the topological phase, the midgap state makes the boundary Green function singular for $\omega \rightarrow 0$. In contrast, there is no such singularity in the trivial phase where the end remains gapped and the boundary Green function is regular. These statements are an immediate consequence of the Lehmann representation for the Green function.

The phase diagram of the model emerges when studying the stability of the fixed-point Green functions under the recursion (7.8). A fixed point is stable when $\beta'(x) < 0$. A simple calculation yields

$$\beta'(G_{\text{top}}) = \frac{t_1^2}{t_2^2} - 1 \quad (7.13)$$

$$\beta'(G_{\text{triv}}) = \frac{t_2^2}{t_1^2} - 1 \quad (7.14)$$

for $\omega = 0$. This implies that the singular boundary Green function G_{top} corresponding to the topological phase is stable for $|t_1/t_2| < 1$, while the regular boundary Green function G_{triv} corresponding to the trivial phase is stable for $|t_1/t_2| > 1$.

The fixed-point analysis also provides information about the gap, the correlation length of the model, and their behavior near the topological critical point $t_1 = t_2$. The correlation length of the phase with fixed-point Green function G is given by $\xi = 4/|\beta'(G)|$. (Notice that we define ξ as the decay length of a wavefunction; the diagonal elements of the Green function decay on half of this scale.) Thus, we find the result

$$\xi = \frac{4[\max(t_1, t_2)]^2}{|t_1^2 - t_2^2|}, \quad (7.15)$$

which is written in a way that applies to both phases. The correlation length diverges at the topological critical point $|t_1| = |t_2|$. It is also interesting to note that this result is consistent with the quasiparticle weight Z of the fixed-point Green function. In the topological phase, one expects $G_{\text{top}} \simeq Z/\omega$ for small ω and $Z = 4/\xi$ reflecting the fact that the midgap state is rapidly oscillating and decays exponentially into the bulk on the scale of the correlation length.

To extract the spectral gap Δ , we note that it is related to the correlation length through $\Delta = \hbar v_F/\xi$ (assuming that the gap is small compared to the bandwidth, i.e., $|t_1 - t_2| \ll t_1, t_2$). Here, the Fermi wavelength should be taken at the center of the band with $t = t_1 = t_2$, i.e., $v_F = 2t/\hbar$. This yields $\Delta = |t_1 - t_2|$ in agreement with a direct evaluation of the spectrum.

7.3.3 Topological invariants

The boundary Green function contains direct information on the topological nature of the phases through the gapless end, edge, or surface states described by the boundary Green function. As for the SSH model above, this can be read off directly from the boundary Green function at energies ω below the bulk gap. In 1D, the topological phase is signaled by midgap end states and hence by a contribution to the Green function which is singular for $\omega \rightarrow 0$. In higher dimensions, there are edge or surface modes with a linear dispersion at low energies. We will discuss this in more detail below.

A systematic connection to the periodic table of topological phases can be established by relating the boundary Green function to the reflection matrix. Ref. [9] derived the periodic table based on the reflection matrix r_N and provided explicit expressions for the topological invariants in terms of the r_N . With the relation between the reflection matrix and the boundary Green function, the topological invariants can alternatively be written in terms of the latter.

Start with a tight-binding model of the kind discussed above and attach a lead to one end with M propagating channels for each $\mathbf{k}_\perp$. In the spirit of the tight-binding model, the channels are only coupled to the last site of the chain, with the matrix W denoting the coupling between the lead channels and the adjacent site of the system. By the gapped nature of the system, the incoming wave is fully reflected. Based on the Mahaux-Weidenmüller formula, the corresponding scattering – or reflection – matrix r_N can then be written as [22]

$$r_N = \mathbb{I} - 2iW \frac{1}{G_N^{-1} + iW^\dagger W} W^\dagger. \quad (7.16)$$

Note that the Green function G_N appearing in this expression is indeed the boundary Green function since the coupling is to the last site of the chain only. Some rearranging then leads to the relation [23]

$$r_N = \frac{\mathbb{I} - iWG_N W^\dagger}{\mathbb{I} + iWG_N W^\dagger}. \quad (7.17)$$

We include an alternative and more detailed derivation of this relation in App. 7.7. Moreover, we couch the scattering-matrix approach to the periodic table and the topological invariants in terms of the boundary Green function in App. 7.8.

Before concluding this section, we briefly consider systems with a trivial bulk topological invariant and thus without a gapless boundary mode. In this case, the boundary Green function evaluated at $\omega = 0$ is generally invertible. This can be used to define an effective boundary Hamiltonian

$$H_{\text{bound}} = -[G(\omega = 0)]^{-1}. \quad (7.18)$$

Even if the original model does not have a nontrivial bulk topological invariant, this boundary Hamiltonian can still be topologically nontrivial, leading to a second order topological phase. We use this observation to extend our approach to describe higher-order topological phases in Sec. 7.6.2.

7.4 Boundary Green function and transfer matrix

In this section, we introduce a more efficient method to compute the fixed-point boundary Green function which comes in handy when discussing models which are more complicated than the SSH model. The method is based on the transfer matrix [14, 15] and effectively provides an explicit construction of the fixed-point boundary Green function.

We start by reviewing the transfer matrix based on the Schrödinger equation for the Hamiltonian in Eq. (7.1),

$$g_n^{-1}(\omega)\psi(n) - V_{n-1}\psi(n-1) - V_n^\dagger\psi(n+1) = 0. \quad (7.19)$$

Here, $\Psi(n)$ is an M -component wave function. The site index n takes on values $n = 1, \dots, N$, and we impose the open boundary conditions $\psi_j(0) = 0 = \psi_j(N+1)$.

As usual for tight-binding Hamiltonians, we can relate the two-component quantities $\Psi(n) = [\psi(n+1)^T, \psi(n)^T]^T$ via $\Psi(n) = M_n(\omega)\Psi(n-1)$, where the matrix $M_n(\omega)$ takes the form

$$M_n(\omega) = \begin{pmatrix} (V_n^\dagger)^{-1}g_n^{-1}(\omega) & -(V_n^\dagger)^{-1}V_{n-1} \\ \mathbb{I} & 0 \end{pmatrix}. \quad (7.20)$$

Here and in the following, we assume V_n to be invertible. Notice that V_0 and V_N do not appear in the Hamiltonian and can be chosen as any invertible matrix in order to define M_0 and M_N . The matrix $M_n(\omega)$ is sometimes called a transfer matrix. Here, we reserve this terminology for the matrix

$$\mathcal{M}_N = M_N M_{N-1} \dots M_1, \quad (7.21)$$

connecting the two ends of the chain.

If ω_0 is an eigenenergy of the Schrödinger equation (7.19), the transfer matrix $\mathcal{M}_N(\omega_0)$ connects the eigenfunctions $\psi(n)$ at sites $n = 1$ and $n = N$,

$$\begin{pmatrix} 0 \\ \psi(N) \end{pmatrix} = \mathcal{M}_N(\omega_0) \begin{pmatrix} \psi(1) \\ 0 \end{pmatrix}. \quad (7.22)$$

If the state corresponding to the eigenenergy is k -fold degenerate, $\Psi_p(n)$ with $p = 1, \dots, k$, then

$$\det \mathcal{M}_{N,11}(\omega_0) = 0, \quad (7.23)$$

and

$$\ker \mathcal{M}_{N,11}(\omega_0) = \text{span}\{\psi_1(1), \dots, \psi_k(1)\}. \quad (7.24)$$

Here, the subscript 11 denotes the blocks under the partition of $\mathcal{M}_N$ into a 2×2 block matrix. One immediate consequence of this is that $k \leq M$, namely the degeneracy is bounded by the number of internal degrees of freedom per unit cell.

We now show how the boundary Green function G_N can be expressed in terms of the transfer matrix $\mathcal{M}_N$, starting directly from the recursion in Eq. (7.4). Defining $X_n = G_n V_n^\dagger$, we can rewrite the recursion as

$$(V_N^\dagger)^{-1} g_N^{-1} X_N - (V_N^\dagger)^{-1} V_{N-1} X_{N-1} X_N = \mathbb{I}, \quad (7.25)$$

or equivalently

$$\begin{pmatrix} \mathbb{I} \\ X_N \end{pmatrix} = M_N \begin{pmatrix} \mathbb{I} \\ X_{N-1} \end{pmatrix} X_N, \quad (7.26)$$

where M_N is defined in Eq. (7.20).

This relation can be applied iteratively to obtain

$$\begin{pmatrix} \mathbb{I} \\ X_N \end{pmatrix} = \mathcal{M}_N \begin{pmatrix} \mathbb{I} \\ 0 \end{pmatrix} X_1 \dots X_N. \quad (7.27)$$

The first row of this equation implies that

$$X_1 \dots X_N = \mathcal{M}_{N,11}^{-1}. \quad (7.28)$$

Then, the second row yields

$$X_N = \mathcal{M}_{N,21} X_1 \dots X_N = \mathcal{M}_{N,21} \mathcal{M}_{N,11}^{-1}. \quad (7.29)$$

Hence, we obtain the desired expression

$$G_N = \mathcal{M}_{N,21} \mathcal{M}_{N,11}^{-1} (V_N^\dagger)^{-1} \quad (7.30)$$

for the boundary Green function in terms of the transfer matrix. This formula was also derived in the broader context of block Toeplitz matrices in Ref. [24].

It is useful to exploit the symplectic structure of the transfer matrix. Notice that the matrix M_n has the property

$$M_n^\dagger \Omega_n M_n = \Omega_{n-1} \quad (7.31)$$

in terms of the anti-hermitian matrix

$$\Omega_n = \begin{pmatrix} 0 & -V_n \\ V_n^\dagger & 0 \end{pmatrix}. \quad (7.32)$$

As a consequence, the transfer matrix $\mathcal{M}_N$ obeys the relation

$$\mathcal{M}_N^\dagger \Omega_N \mathcal{M}_N = \Omega_0. \quad (7.33)$$

Remember that the choice of matrices V_N and V_0 at the boundary is arbitrary as long as they are invertible. If we choose $V_0 = V_N$, then the transfer matrix obtains a symplectic structure as it satisfies the relation

$$\mathcal{M}_N^\dagger \Omega_0 \mathcal{M}_N = \Omega_0. \quad (7.34)$$

As usual for symplectic matrices, this implies that the eigenvalues of $\mathcal{M}_N$ appear in pairs of λ and $1/\lambda^*$. In particular, if v is a right eigenvector of $\mathcal{M}_N$ with eigenvalue λ ,

$$\mathcal{M}_N v = \lambda v, \quad (7.35)$$

then $(\Omega_0 v)^\dagger$ is a left eigenvector of $\mathcal{M}_N$ with eigenvalue $1/\lambda^*$,

$$(\Omega_0 v)^\dagger \mathcal{M}_N = \frac{1}{\lambda^*} (\Omega_0 v)^\dagger. \quad (7.36)$$

This separates the $2M$ eigenvalues of the transfer matrix into M eigenvalues λ_j ($j = 1, \dots, M$) with $|\lambda_j| > 1$, and another M eigenvalues $1/\lambda_j^*$ with modulus smaller than unity. The Lyapunov exponents $1/\xi_j$ are then defined through

$$\frac{1}{\xi_j} = \lim_{N \rightarrow \infty} \frac{1}{N} \ln |\lambda_j| \quad (7.37)$$

and are strictly positive at the Fermi energy ($\omega = 0$) for a gapped bulk [25]. At the topological phase transition, the bulk gap closes and therefore at least one Lyapunov exponent vanishes for some j , corresponding to $|\lambda_j| = 1$ or a diverging ξ_j . This is explicitly shown below.

Condition for a vanishing bulk gap

In this appendix, we show that when the bulk band gap vanishes, the matrix T defined in Eq. (7.48) has at least one eigenvalue with unit modulus and thus one vanishing Lyapunov exponent. This implies that the topological phase transition with its associated gap closing is signalled by at least one vanishing Lyapunov exponent.

Consider a homogeneous, quasi-1D system with periodic boundary conditions. According to the Bloch theorem, we can write the wave function as $\psi(n) = e^{ikn} \mathbf{u}$, where $\mathbf{u}$ is an M -component column vector. Using Eq. (7.19), we find a zero-energy eigenstate when

$$[g^{-1}(\omega = 0) - V e^{-ik} - V^\dagger e^{ik}] \mathbf{u} = 0. \quad (7.38)$$

This implies

$$\det [g^{-1}(\omega = 0) - V e^{-ik} - V^\dagger e^{ik}] = 0. \quad (7.39)$$

Thus, $\lambda = e^{-ik}$ is an eigenvalue of $T(\omega = 0)$ defined in Eq. (7.48), which implies that the corresponding Lyapunov exponent vanishes.

Continue of the theory

So far, we did not assume that the system is uniform along the direction with open boundary conditions. Let us now consider models with translational invariance of period a . (In the SSH model discussed above, this period was equal to 2 due to the dimerization of the hopping.) Considering a chain of length Na , the transfer matrix becomes

$$\mathcal{M}_{aN} = T^N, \quad (7.40)$$

where we defined the matrix

$$T = M_a M_{a-1} \dots M_1 \quad (7.41)$$

associated with one period of length a .

In the translationally invariant case, the evaluation of the boundary Green function can be simplified by diagonalizing the matrix T . Again, T satisfies a symplectic structure and its eigenvalues can be separated into $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_M)$ with $|\lambda_j| > 1$, and a second set $(\Lambda^{-1})^*$. Thus, we can write

$$T \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} \begin{pmatrix} \Lambda & 0 \\ 0 & (\Lambda^{-1})^* \end{pmatrix} \quad (7.42)$$

We can use this to express T and hence $\mathcal{M}_{aN}$, and insert this expression into the Eq. (7.30) for the boundary Green function. Upon taking the thermodynamic limit $N \rightarrow \infty$ and using that the system is gapped (see below for details), this yields the desired explicit expression

$$G = \lim_{N \rightarrow \infty} G_{aN} = U_{21} U_{11}^{-1} (V_0^\dagger)^{-1}, \quad (7.43)$$

for the fixed-point boundary Green function (see also Ref. [24]). While we have suppressed the argument ω in this expression, it is valid for any ω which falls inside the bulk gap.

Derivation of Eq. (7.43)

From Eq. (7.40) and (7.42), we have

$$\mathcal{M}_{aN} \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} \begin{pmatrix} \Lambda^N & 0 \\ 0 & (\Lambda^{-N})^* \end{pmatrix} \quad (7.44)$$

which gives

$$\mathcal{M}_{aN,11} = U_{11}\Lambda^N(U^{-1})_{11} + U_{12}(\Lambda^*)^{-N}(U^{-1})_{21} \quad (7.45)$$

$$\mathcal{M}_{aN,21} = U_{21}\Lambda^N(U^{-1})_{11} + U_{22}(\Lambda^*)^{-N}(U^{-1})_{21}. \quad (7.46)$$

Hence

$$\begin{aligned} \mathcal{M}_{aN,21}\mathcal{M}_{aN,11}^{-1} &= [U_{21}\Lambda^N(U^{-1})_{11} + U_{22}(\Lambda^*)^{-N}(U^{-1})_{21}] [U_{11}\Lambda^N(U^{-1})_{11} + U_{12}(\Lambda^*)^{-N}(U^{-1})_{21}]^{-1} \\ &= [U_{21}\Lambda^N(U^{-1})_{11} + O(\Lambda^{-N})] [(U^{-1})_{11}^{-1}\Lambda^{-N}U_{11}^{-1} + O(\Lambda^{-2N})] \\ &= U_{21}U_{11}^{-1} + O((\Lambda^{-2N})). \end{aligned} \quad (7.47)$$

Taking the thermodynamic limit $N \rightarrow \infty$, we have $\Lambda^{-2N} \rightarrow 0$ and thus arrive at Eq. (7.43).

Continue of the theory

For a homogeneous system in which the on-site potential h_n and the hopping amplitude V_n are independent of the site index n along the chain, i.e., a system with $a = 1$, the evaluation of the boundary Green functions can be simplified further. In this case, the columns of the matrix U in Eq. (7.43) consist of the right eigenvectors of

$$T = \begin{pmatrix} (V^\dagger)^{-1}g^{-1} & -(V^\dagger)^{-1}V \\ \mathbb{I} & 0 \end{pmatrix}. \quad (7.48)$$

The simple structure of T implies that the right eigenvector with eigenvalue λ_i can be written as $(\lambda_i z_i^T, z_i^T)^T$, where the vector z_i obeys

$$\left(\frac{V}{\lambda_i} + V^\dagger \lambda_i - g^{-1} \right) z_i = 0. \quad (7.49)$$

Let us also denote the eigenvectors corresponding to eigenvalues $|\lambda_i| > 1$ as x_i and the corresponding eigenvectors with eigenvalues $1/\lambda_i^*$ as y_i . We can now identify $U_{21} = (x_1, \dots, x_M)$ and $U_{11} = (\lambda_1 x_1, \dots, \lambda_M x_M)$. Then, Eq. (7.43) implies

$$G = (x_1, \dots, x_M) \Lambda^{-1} (x_1, \dots, x_M)^{-1} (V^\dagger)^{-1}. \quad (7.50)$$

for the fixed-point boundary Green function. The essential steps in evaluating G thus reduce to solving the nonlinear eigenvalue problem Eq. (7.49) and inserting the result into Eq. (7.50).

For later use, it is useful to collect expressions for all right and left eigenvectors of T for $a = 1$. We write the right eigenvectors of T as $v_i = (x_i^T, x_i^T/\lambda_i)^T$ for $|\lambda_i| > 1$ and as $v_{M+i} = (y_i^T, y_i^T \lambda_i^*)^T$ for the corresponding eigenvalues $1/\lambda_i^*$ with modulus less than unity. The symplectic nature of T implies then that the corresponding left eigenvectors take the form $w_i^T = (-y_i^\dagger V^\dagger \lambda_i, y_i^\dagger V)$ with eigenvalue λ_i and $w_{M+i}^T = (-x_i^\dagger V^\dagger/\lambda_i^*, x_i^\dagger V)$ with eigenvalue $1/\lambda_i^*$.

7.5 Application to 1D and 2D models

In this section, we apply the formalism developed in the previous section to two homogeneous tight-binding models with $a = 1$. We consider the Kitaev chain as an example of a 1D model and a Chern insulator as an example of a 2D model.

7.5.1 Perturbation theory

We are mainly interested in the fixed-point boundary Green function for small ω as this limit captures the topological properties and the effective low-energy excitations. This limit can be most directly treated by means of perturbation theory. Denoting $T_0 = T(\omega = 0)$, we have

$$T = T_0 + \omega A, \quad A = \begin{pmatrix} (V^\dagger)^{-1} & 0 \\ 0 & 0 \end{pmatrix}. \quad (7.51)$$

For small ω , we can treat ωA in first-order perturbation theory (properly generalized to non-hermitian matrices, for which left and right eigenvectors are no longer identical).

If we denote by w_i^T and v_i the left and right eigenvectors of M_0 with eigenvalues λ_i (not restricted in magnitude), we obtain the first-order corrections

$$\delta\lambda_i = \omega \frac{w_i^T A v_i}{w_i^T v_i}, \quad \delta v_i = \omega \sum_{j \neq i} \frac{v_j}{\lambda_i - \lambda_j} \frac{w_j^T A v_i}{w_j^T v_j}. \quad (7.52)$$

These expressions are evident generalizations of the corresponding results in quantum mechanics.

We can now use the explicit structure of the v_i discussed at the end of the previous section as well as the simple form of A . This yields

$$\delta\lambda_i = -\frac{\omega \lambda_i y_i^\dagger x_i}{y_i^\dagger (V/\lambda_i - V^\dagger \lambda_i) x_i} \quad (7.53)$$

for the shift in the eigenvalues with $|\lambda_i| > 1$ and

$$\delta v_i = \omega \begin{pmatrix} \delta r_i \\ \delta s_i \end{pmatrix}, \quad (7.54)$$

with

$$\delta r_i = -\sum_{j \neq i} \frac{y_j^\dagger x_i}{(\lambda_i \lambda_j^* - 1) y_j^\dagger (V/\lambda_j - V^\dagger \lambda_j) x_j} x_j - \sum_j \frac{x_j^\dagger x_i}{(\lambda_i \lambda_j^* - 1) x_j^\dagger (V \lambda_j^* - V^\dagger / \lambda_j^*) y_j} y_j, \quad (7.55)$$

for the corresponding first-order corrections to the right eigenvectors. In the last expression, the sum is only over those indices j with $|\lambda_j| > 1$.

At small ω , the fixed-point boundary Green function in Eq. (7.50) then becomes

$$G = (x_1, \dots, x_M) \Lambda^{-1} \times (x_1 + \omega \delta r_1, \dots, x_M + \omega \delta r_M)^{-1} (V^\dagger)^{-1}. \quad (7.56)$$

Notice that we retain the correction due to a nonzero ω only in the matrix inverse because this is where a singular ω dependence can originate from. Notice also that δs_i does not enter this expression so that we refrain from giving an explicit expression.

7.5.2 Kitaev chain

Model and boundary Green function

As a first example, consider the Kitaev chain which is a lattice model for a spinless p -wave superconductor. The Hamiltonian for a chain of N sites takes the form

$$H = -\mu \sum_{j=1}^N c_j^\dagger c_j + \sum_{j=1}^{N-1} \left[t c_{j+1}^\dagger c_j + \Delta c_{j+1}^\dagger c_j^\dagger + \text{h.c.} \right], \quad (7.57)$$

where c_j ($c_j^\dagger$) is the fermion annihilation (creation) operator at site j . We assume below that both the hopping t and the pairing Δ are real.

This model has trivial and topological phases, with the latter hosting Majorana zero modes at both ends of the chain. The trivial phase occurs for $|\mu| > 2|t|$ when the chemical potential μ falls outside the normal-state band. The topological phase occurs for $|\mu| < 2|t|$. The model supports a chiral symmetry $\mathcal{U}_S = \tau_x$, so that it falls into class BDI. Indeed, in addition to the topological phase transition lines at $|\mu| = 2|t|$, the model also becomes gapless in the absence of pairing correlations, $\Delta = 0$, and the topological phases with $\Delta > 0$ and $\Delta < 0$ have topological indices ± 1 .

To apply the approach developed in Sec. 7.4, we introduce a Nambu spinor $\psi_j^\dagger = (c_j^\dagger, c_j)$. Then, we can identify

$$g^{-1}(\omega) = \omega + \mu\tau_z \quad (7.58)$$

$$V = t\tau_z + i\Delta\tau_y, \quad (7.59)$$

where $\tau_{x,y,z}$ are Pauli matrices in Nambu space.

Applying the perturbative approach developed in Sec. 7.5.1, we first determine the eigenvalues and eigenvectors of T for $\omega = 0$. The eigenvalues follow from the condition

$$\det [V/\lambda + V^\dagger\lambda - g^{-1}(\omega = 0)] = 0, \quad (7.60)$$

which yields

$$(t \mp \Delta)\lambda^2 - \mu\lambda + (t \pm \Delta) = 0. \quad (7.61)$$

Denote the pair of roots for the upper sign as $\lambda_{1,2}$, the roots for the lower sign as $\lambda_{3,4}$, and fix the labeling by imposing $\lambda_1\lambda_4^* = 1 = \lambda_2\lambda_3^*$. The corresponding eigenvectors fulfilling Eq. (7.49) are

$$z_1 = z_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad z_3 = z_4 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (7.62)$$

To determine the two eigenvectors corresponding to eigenvalues with $|\lambda| > 1$, we consider the quantity

$$(\lambda_1^2 - 1)(\lambda_2^2 - 1) = (\lambda_1\lambda_2 + 1)^2 - (\lambda_1 + \lambda_2)^2 = \frac{4t^2 - \mu^2}{(t - \Delta)^2} \quad (7.63)$$

Thus, $|\lambda_1|$ and $|\lambda_2|$ are both larger than or both smaller than unity when $|2t| > |\mu|$.

Let us first consider the situation that $|\lambda_1|$ and $|\lambda_2|$ are both larger than unity. This occurs for $|2t| > |\mu|$ and $\Delta t < 0$. Then, the matrix $(x_1, x_2) = 1 + \tau_x$ is noninvertible and the fixed-point boundary Green function G becomes singular in the limit $\omega \rightarrow 0$. This is in agreement with expectations as the model exhibits a topological phase for these parameters.

Including the first-order corrections in ω , we find

$$G(\omega) = \frac{(1 + \tau_x)}{2\omega(t - \Delta)a_{12}} \quad (7.64)$$

where

$$a_{12} = \frac{1}{(|\lambda_1|^2 - 1)(\lambda_2 - \lambda_4)(\lambda_1^*(t - \Delta) - \lambda_4(t + \Delta))} + \frac{1}{(|\lambda_2|^2 - 1)(\lambda_1 - \lambda_3)(\lambda_2^*(t - \Delta) - \lambda_3(t + \Delta))}. \quad (7.65)$$

Thus, the fixed-point boundary Green function takes a form which implies the existence of a zero-energy end state. The rank of the numerator $1 + \tau_x$ confirms that there is one zero-energy

end state at each end, and the matrix structure imposes that the end-state wavefunction has particle-hole symmetry as it should for a Majorana bound state. The boundary Green function of the Kitaev chain was also derived in Ref. [26] by a different method.

As for the SSH model, the quasiparticle weight of the fixed-point boundary Green function (7.64) encodes information on the localization length of the Majorana end state. In particular, one expects that the Majorana localization length diverges at the topological phase transition. For the parameters of Eq. (7.64), we have either $|\lambda_1| \rightarrow 1$ or $|\lambda_2| \rightarrow 1$, so that a_{12} is dominated by one of the two terms and does indeed diverge.

Similarly, we find $|\lambda_{3,4}| > 1$ for $|2t| > |\mu|$ and $\Delta t > 0$. Then, the matrix $(x_1, x_2) = 1 - \tau_x$ (exploiting that the eigenvectors are only defined up to an overall prefactor) and we find

$$G(\omega) = \frac{(1 - \tau_x)}{2\omega(t + \Delta)a_{34}} \quad (7.66)$$

where a_{34} follows from the expression for a_{12} by the replacements $(1, 2) \leftrightarrow (3, 4)$. As expected, this expression is again consistent with the particle-hole symmetry of Majorana end states.

Now, consider the nontopological phase in the complementary parameter range $|2t| < |\mu|$. Assuming for definiteness that $|\lambda_{1,3}| > 1$, we have the relation

$$\frac{\lambda_1}{t + \Delta} - \frac{\lambda_3}{t - \Delta} = 0. \quad (7.67)$$

and find that all λ_j are real. We then observe that the matrix (x_1, x_2) is invertible and we can evaluate the fixed-point boundary Green function for $\omega = 0$. This yields

$$G(\omega = 0) = \frac{2\lambda_1}{t + \Delta}\tau_z, \quad (7.68)$$

which is nonsingular indicating that there are no zero-energy end states.

Topological index

The Kitaev chain with the Hamiltonian given in Eq. (7.57) belongs to symmetry class BDI, as it has a chiral symmetry operator $\mathcal{U}_S = \tau_x$. For 1D systems, class BDI is characterized by a topological $\mathbb{Z}$ -invariant which is given by

$$\mathcal{Q} = \nu(\lim_{\omega \rightarrow 0} \mathcal{U}_S \frac{\mathbb{I} - iVGV^\dagger}{\mathbb{I} + iVGV^\dagger}) - \frac{M}{2}. \quad (7.69)$$

Here, M is even and $\nu(F)$ counts the number of negative eigenvalues of a matrix F . This expression follows from Ref. [9] in conjunction with the relation (7.17) between reflection matrix and boundary Green function (see App. 7.8 for a summary).

The existence of zero-energy boundary modes implies that for $\omega \rightarrow 0$, the boundary Green function has the form

$$G \sim \frac{K}{\omega} \quad (7.70)$$

with a singular Hermitian matrix K whose rank corresponds to the number of zero modes. This general form for symmetry class BDI is consistent with the expressions (7.64) and (7.66) which we found above for the Kitaev chain. The chiral symmetry implies that $\mathcal{U}_S$ anticommutes with the Hamiltonian, so that $[U_S, VKV^\dagger] = 0$ (see App. 7.8 for more details). Thus, we can choose a basis in which the Hermitian matrices $\mathcal{U}_S$ and $VKV^\dagger$ are simultaneously diagonalized and order the eigenvalues such that

$$\mathcal{U}_S = \text{diag}(\mathbb{I}_{M/2}, -\mathbb{I}_{M/2}) \quad (7.71)$$

with $\mathbb{I}_n$ the identity matrix of dimension n , and

$$VKV^\dagger = \text{diag} (D_{+,p}, 0_{M/2-p}, D_{-,q}, 0_{M/2-q}) \quad (7.72)$$

with $D_{\pm,p}$ a diagonal matrix containing the nonzero eigenvalues of $VKV^\dagger$. Here, p (q) denotes the number of zero-energy modes with positive (negative) chirality. We also use the notation 0_n for an $n \times n$ zero matrix.

Using the asymptotic form of the boundary Green function in Eq. (7.70), we thus find

$$F = \lim_{\omega \rightarrow 0} \mathcal{U}_S \frac{\mathbb{I} - iVGV^\dagger}{\mathbb{I} + iVGV^\dagger} = \text{diag} (-\mathbb{I}_p, \mathbb{I}_{M/2-p}, \mathbb{I}_q, -\mathbb{I}_{M/2-q}), \quad (7.73)$$

and therefore

$$\mathcal{Q} = \nu(F) - M/2 = p - q. \quad (7.74)$$

Thus, the topological index of symmetry class BDI can be computed as the difference between the numbers of zero-energy modes with positive and negative chiralities.

In the absence of zero modes, $\lim_{\omega \rightarrow 0} G(\omega) = G(0)$ is generically an invertible, full-rank matrix. Then, chiral symmetry implies $\{\mathcal{U}_S, VG(0)V^\dagger\} = 0$. (Notice that unlike for the singular boundary Green function in the topological phase, the leading contribution to G for $\omega \rightarrow 0$ is independent of, and thus even in, ω .) Hence, we can choose a basis in which

$$\mathcal{U}_S = \sigma_x \otimes \mathbb{I}_{M/2} \quad (7.75)$$

and

$$VG(0)V^\dagger = \sigma_z \otimes D_{M/2}, \quad (7.76)$$

where the diagonal $M/2 \times M/2$ matrix $D_{M/2}$ collects eigenvalues of $VG(0)V^\dagger$. We obtain the matrix

$$F = \lim_{\omega \rightarrow 0} \mathcal{U}_S \frac{\mathbb{I} - iVGV^\dagger}{\mathbb{I} + iVGV^\dagger} = (\sigma_x \otimes \mathbb{I}_{M/2}) \frac{\mathbb{I}_M - i(\sigma_z \otimes D_{M/2})}{\mathbb{I}_M + i(\sigma_z \otimes D_{M/2})}. \quad (7.77)$$

As $\{\sigma_z \otimes \mathbb{I}_{M/2}, F\} = 0$, the spectrum of F must be symmetric with respect to zero. Hence, we conclude $\mathcal{Q} = 0$ as expected.

For the Kitaev chain in the topological phase, the asymptotic form of the boundary Green function is $G \sim (1 \pm \tau_x)/\omega$. This leads to the topological invariant $\mathcal{Q} = \pm 1$, corresponding to the existence of one Majorana zero mode with chirality ± 1 . When this model is in its trivial phase, we have $G \sim \tau_z$ and the topological invariant becomes $\mathcal{Q} = 0$.

In realistic 1D topological superconductors, the chiral symmetry will typically be absent. This lowers the symmetry class from BDI to D, which has a $\mathbb{Z}_2$ topological invariant. In this case, there can be at most one zero-energy boundary mode when the system is in the topological phase. Then, the asymptotic Green function $G \sim K/\omega$ involves a matrix K of rank one. The topological invariant for this case is (see App. 7.8 for details)

$$\mathcal{Q} = \det \left(\lim_{\omega \rightarrow 0} \frac{\mathbb{I} - iVGV^\dagger}{\mathbb{I} + iVGV^\dagger} \right), \quad (7.78)$$

which takes the value $+1$ in the trivial phase and -1 in the topological phase.

In the presence of one boundary zero mode, diagonalizing the matrix $VKV^\dagger$ results in only one non-zero eigenvalue. In this diagonal basis, one finds

$$\lim_{\omega \rightarrow 0} \frac{\mathbb{I} - iVGV^\dagger}{\mathbb{I} + iVGV^\dagger} = \text{diag}(-1, \mathbb{I}_{M-1}). \quad (7.79)$$

which leads to $\mathcal{Q} = -1$.

In the trivial phase, $\lim_{\omega \rightarrow 0} G(\omega) = G(0)$ is generically a full rank matrix. Because of particle-hole symmetry (see App. 7.8 for further details), the spectrum of $VG(0)V^\dagger$ is symmetric about zero, which leads to $\mathcal{Q} = 1$.

As an aside, this implies that the topological invariant for class D signals the presence or absence of Majoranas in the Kitaev chain despite its being in class BDI.

7.5.3 Chern insulator

Model and boundary Green function

In this section, we consider a Chern insulator as an example of a 2D topological phase. A simple model Hamiltonian for a Chern insulator on a square lattice is given by [3]

$$H = \sin k_x \sigma_x + \sin k_y \sigma_y + B(2 - M - \cos k_x - \cos k_y) \sigma_z. \quad (7.80)$$

The corresponding spectrum is gapped except for $M = 0, 2, 4$, and the model exhibits topological phases for $0 < M < 2$ – with edge states around $k = 0$ – and for $2 < M < 4$ – with edge states around $k = \pi$.

To compute the boundary Green function, we rewrite the Hamiltonian in real space along the x -direction (and retain the momentum representation in the y -direction with periodic boundary conditions). This yields

$$H(k_y) = \sum_{n=0}^N \left[\frac{1}{2i} c_{n+1}^\dagger \sigma_x c_n - \frac{B}{2} c_{n+1}^\dagger \sigma_z c_n + \text{h.c.} + c_n^\dagger (\sin k_y \sigma_y + B\Lambda(k_y) \sigma_z) c_n \right] \quad (7.81)$$

with $\Lambda(k_y) = 2 - M - \cos k_y$. We can then readily read off

$$g^{-1}(k_y, \omega) = \omega - \sin k_y \sigma_y - B\Lambda(k_y) \sigma_z \quad (7.82)$$

$$V = \frac{1}{2i} \sigma_x - \frac{B}{2} \sigma_z. \quad (7.83)$$

It is useful to compute the fixed-point boundary Green function in a perturbation theory in $(\omega - \sin k_y \sigma_y)$, starting with

$$T = T_0 + A(\omega - \sin k_y \sigma_y), \quad A = \begin{pmatrix} (V^\dagger)^{-1} & 0 \\ 0 & 0 \end{pmatrix}, \quad (7.84)$$

where $T_0 = T(\sin k_y = 0, \omega = 0)$. For $\omega - \sin k_y \sigma_y = 0$, the eigenvalues satisfy

$$\lambda^2(1 \pm B) \mp 2B\Lambda\lambda - (1 \mp B) = 0. \quad (7.85)$$

We denote the roots for the upper sign as $\lambda_{1,2}$ and for the lower sign as $\lambda_{3,4}$, and fix the labeling by imposing $\lambda_1 \lambda_4^* = 1 = \lambda_2 \lambda_3^*$. The corresponding eigenvectors are

$$z_1 = z_2 = \begin{pmatrix} 1 \\ i \end{pmatrix}, \quad z_3 = z_4 = \begin{pmatrix} 1 \\ -i \end{pmatrix}. \quad (7.86)$$

Similar to the Kitaev chain, we expect a topological phase when $|\lambda_1|$ and $|\lambda_2|$ are either both larger or both smaller than unity. The relevant parameter ranges can again be deduced from the quantity

$$(\lambda_1^2 - 1)(\lambda_2^2 - 1) = \frac{4B^2}{(1 + B)^2} (1 - \Lambda^2). \quad (7.87)$$

Thus, the topological phase requires $\Lambda^2 < 1$, which implies $|\lambda_{1,2}| > 1$ for $B < 0$ and $|\lambda_{3,4}| > 1$ for $B > 0$. Using that $\cos k_y = \pm 1$ when $\sin k_y = 0$, we recover that the model is topological for $0 < M < 4$.

We can now evaluate the fixed-point boundary Green function in perturbation theory by following essentially the same steps as for the Kitaev chain. When $|\lambda_{1,2}| > 1$, a straight-forward calculation yields

$$G(k_y, \omega) = \frac{(1 + \sigma_y)}{2(B + 1)(\omega + \sin k_y)a_{12}} \quad (7.88)$$

with

$$a_{12} = \frac{1}{[-(B+1)\lambda_1^* + (B-1)\lambda_4](|\lambda_1|^2 - 1)(\lambda_2 - \lambda_4)} + \frac{1}{[-(B+1)\lambda_2^* + (B-1)\lambda_3](|\lambda_2|^2 - 1)(\lambda_1 - \lambda_3)}. \quad (7.89)$$

Similarly, when $|\lambda_{3,4}| > 1$, we have

$$G(k_y, \omega) = \frac{(1 - \sigma_y)}{2(B - 1)(\omega - \sin k_y)a_{34}} \quad (7.90)$$

where

$$a_{34} = \frac{1}{(|\lambda_3|^2 - 1)(\lambda_4 - \lambda_1)[(1 - B)\lambda_4^* + (B + 1)\lambda_1]} + \frac{1}{(|\lambda_4|^2 - 1)(\lambda_3 - \lambda_2)[(1 - B)\lambda_3^* + (B + 1)\lambda_2]}. \quad (7.91)$$

Thus, we find that when $0 < M < 2$, the system has a gapless chiral boundary mode near $k_y = 0$ with dispersion $\omega = \text{sgn}(B) \sin k_y$. When $2 < M < 4$, the gapless chiral boundary mode occurs around $k_y = \pi$ with dispersion $\omega = \text{sgn}(B) \sin k_y$.

In the nontopological phase for $M < 0$ or $M > 4$, we have $\Lambda^2 > 1$ and can thus choose $|\lambda_1|, |\lambda_3| > 1$. As for the Kitaev chain, the λ_j are all real and obey the relation

$$\frac{\lambda_1}{1 - B} + \frac{\lambda_3}{1 + B} = 0. \quad (7.92)$$

The boundary Green function is nonsingular and can hence be evaluated approximately for $\omega - \sin k_y = 0$. This yields

$$G(k_y \rightarrow 0, \pi, \omega \rightarrow 0) = -\frac{2\lambda_1}{1 - B}\sigma_z. \quad (7.93)$$

Topological index

The Chern insulator with Hamiltonian (7.80) belongs to class A. The topological invariant for a two-dimensional system in this symmetry class takes the form [9] (see also App. 7.8)

$$\mathcal{Q} = \frac{1}{2\pi i} \int_0^{2\pi} dk \frac{d}{dk} \text{tr} \log r(k), \quad (7.94)$$

where the zero-energy reflection matrix is expressed in terms of the boundary Green function as

$$r(k_y) = \lim_{\omega \rightarrow 0} \frac{\mathbb{I} - iVG(k_y, \omega)V^\dagger}{\mathbb{I} + iVG(k_y, \omega)V^\dagger}. \quad (7.95)$$

In line with Laughlin's argument [27] and the familiar expression for quantum pumping of charge [28], this can be viewed as the pumped charge in units of the electron charge when threading

the effectively cylinder-shaped sample (periodic boundary conditions in the y -direction) by one flux quantum.

When $0 < M < 2$, the boundary Green function diverges at $k_y = 0$ and has rank one,

$$G(k_y, \omega \rightarrow 0) \sim \frac{1 + \sigma_y}{k_y}. \quad (7.96)$$

Away from $k_y = 0$ the boundary Green function is invertible and finite. This yields a winding number $\mathcal{Q} = 1$.

When $2 < M < 4$, the boundary Green function diverges at $k_y = \pi$ and has rank one,

$$G(k_y, \omega \rightarrow 0) \sim -\frac{1 + \sigma_y}{k_y - \pi}. \quad (7.97)$$

Away from $k_y = \pi$ the boundary Green function is well behaved. We thus obtain the opposite winding number $\mathcal{Q} = -1$.

When $M < 0$ or $M > 4$, the boundary Green function is well behaved at all k_y and the winding number is zero.

7.6 Higher-order topological insulators

7.6.1 Model

Recent work introduced higher-order topological insulators [19]. In 2D, these phases have gapped edges but gapless corner states. In this section, we show that our approach readily accommodates such systems and provides a rather transparent picture.

The model investigated in Ref. [19] takes the form

$$H = \sum_{n,j} \left(t_{x,1} c_{2n,j}^\dagger c_{2n-1,j} + t_{x,2} c_{2n+1,j}^\dagger c_{2n,j} \right) - \sum_{i,m} (-1)^i \left(t_{y,1} c_{i,2m}^\dagger c_{i,2m-1} + t_{y,2} c_{i,2m+1}^\dagger c_{i,2m} \right). \quad (7.98)$$

This square-lattice Hamiltonian has dimerized hopping amplitudes in the x and y -directions and an added π -flux per plaquette, as illustrated in Fig. 60. The hopping amplitudes $t_{\alpha,j}$ ($\alpha = x, y$ and $j = 1, 2$) are assumed to be real. As a consequence of the π flux, the model has an insulating bulk which makes it amenable to our boundary Green function approach.

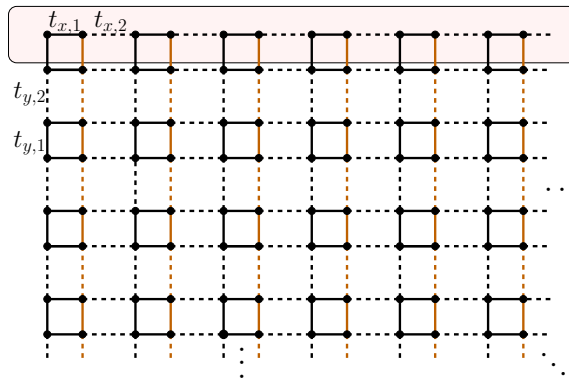


Figure 60: Tight-binding model for a second order topological insulator on a square lattice. Solid and dashed lines in the x and y directions denote alternating nearest neighbor hoppings $t_{x,1}$, $t_{x,2}$ and $t_{y,1}$, $t_{y,2}$. The hopping amplitudes for the bonds colored in brown include a relative minus sign. The effective Hamiltonian of the highlighted boundary parallel to the x -direction is topologically equivalent to an SSH model.

While the edges of this model are gapped, there are zero-energy corner states [19]. In the presence of C_4 symmetry, these corner states can be characterized by a quadrupole moment which is a $\mathbb{Z}_2$ topological index. In the absence of C_4 symmetry, there is a $\mathbb{Z}_2 \times \mathbb{Z}_2$ topological index which can be associated with the edge polarizations.

The main observation underlying the current section is that even if the fixed-point boundary Green function is gapped, it defines a low-energy boundary Hamiltonian in a natural manner. This boundary Hamiltonian can be either trivial or topological. Gapless corner states (or corresponding generalizations to higher dimensions) appear if this boundary Hamiltonian is topological.

7.6.2 Effective boundary Hamiltonian

We now derive the fixed-point boundary Green function of this model. The dimerization of the hopping amplitudes makes this a model with period $a = 2$, with $T = M_2 M_1$ being the product of two matrices. As described in Sec. 7.4, we can deduce the fixed-point boundary Green function from the eigenvectors and eigenvalues of T .

Consider a system with $2N \times 2N$ sites and focus on the Green function for a boundary parallel to the x -direction (in which we thus assume periodic boundary conditions). Introducing a two-component spinor $\psi_{n,j}^T = (c_{2n-1,j}^T, c_{2n,j}^T)$ and its Fourier transform

$$\psi_{n,j} = \frac{1}{\sqrt{N}} \sum_k e^{ink} \psi_j(k), \quad (7.99)$$

the Hamiltonian can be rewritten as

$$H = \sum_k \left\{ \sum_{j=1}^{2N} \psi_j^\dagger(k) h_j(k) \psi_j(k) + \sum_{j=1}^{2N-1} [\psi_{j+1}^\dagger(k) V_j \psi_j(k) + \text{h.c.}] \right\}. \quad (7.100)$$

Here, we define

$$h_n(k) = t_{x,1} \tau_x + t_{x,2} (e^{-ik} \tau_+ + e^{ik} \tau_-), \quad (7.101)$$

and

$$V_j = \begin{cases} t_{y,1} \tau_z & \text{odd } j \\ t_{y,2} \tau_z & \text{even } j \end{cases}. \quad (7.102)$$

The matrices $\tau_{x,y,z}$ are Pauli matrices in the sublattice space of even and odd sites along the x -direction, with $\tau_\pm = (\tau_x \pm i\tau_y)/2$.

With these ingredients, we can construct the transfer matrix $T = M_2 M_1$ using Eq. (7.20). As the edge is gapped, we can set $\omega = 0$ from the outset. This yields

$$T = \begin{pmatrix} -\frac{t_{x,1}^2 + t_{x,2}^2 + t_{y,1}^2 + 2t_{x,1}t_{x,2} \cos k}{t_{y,1}t_{y,2}} & 0 & 0 & \frac{t_{x,1}}{t_{y,1}} + \frac{t_{x,2}}{t_{y,1}} e^{-ik} \\ 0 & -\frac{t_{x,1}^2 + t_{x,2}^2 + t_{y,1}^2 + 2t_{x,1}t_{x,2} \cos k}{t_{y,1}t_{y,2}} & -\frac{t_{x,1}}{t_{y,1}} - \frac{t_{x,2}}{t_{y,1}} e^{ik} & 0 \\ 0 & -\frac{t_{x,1}}{t_{y,1}} - \frac{t_{x,2}}{t_{y,1}} e^{-ik} & -\frac{t_{y,2}}{t_{y,1}} & 0 \\ \frac{t_{x,1}}{t_{y,1}} + \frac{t_{x,2}}{t_{y,1}} e^{ik} & 0 & 0 & -\frac{t_{y,2}}{t_{y,1}} \end{pmatrix}. \quad (7.103)$$

The matrix T decomposes into two independent 2×2 blocks, which we can treat separately. Both blocks have the same eigenvalues λ (with $|\lambda| > 1$) and $1/\lambda^*$. We find that λ fulfills the equation

$$\lambda^2 + \lambda \frac{t_{x,1}^2 + t_{x,2}^2 + t_{y,1}^2 + t_{y,2}^2 + 2t_{x,1}t_{x,2} \cos k}{t_{y,1}t_{y,2}} + 1 = 0. \quad (7.104)$$

Evaluating the corresponding eigenvectors, we can then compute the fixed-point boundary Green function based on Eq. (7.43). This yields

$$G(0, k) = \frac{1}{t_{y,2}(t_{y,2} + \lambda t_{y,1})} \begin{pmatrix} 0 & t_{x,1} + t_{x,2}e^{-ik} \\ t_{x,1} + t_{x,2}e^{ik} & 0 \end{pmatrix}, \quad (7.105)$$

which is invertible and purely real, reflecting the fact that this describes the gapped edge of a gapped bulk.

We can now define an effective boundary Hamiltonian

$$H_{\text{eff}} = -[G(0, k)]^{-1} = f_k \begin{pmatrix} 0 & t_{x,1} + t_{x,2}e^{-ik} \\ t_{x,1} + t_{x,2}e^{ik} & 0 \end{pmatrix} \quad (7.106)$$

from the fixed-point boundary Green function. Here, the prefactor

$$f(k) = \frac{t_{y,2}(t_{y,2} + \lambda t_{y,1})}{t_{x,1}^2 + t_{x,2}^2 + 2t_{x,1}t_{x,2} \cos k} \quad (7.107)$$

has a fixed sign as a function of k . It is our central observation that up to the prefactor f_k , this is just the Hamiltonian of the SSH model (7.6). Thus, the boundary Hamiltonian is topologically equivalent to the bulk Hamiltonian of the SSH model, with t_1 and t_2 replaced by $t_{x,1}$ and $t_{x,2}$.

A similar analysis can be performed for the edge parallel to the y -direction. In view of the fact that the π -flux can be equally included via the hopping amplitudes in the x -direction, the calculation is entirely equivalent and just results in exchanging the subscripts x and y . Thus, the edge along the y -direction can also be characterized by a $\mathbb{Z}_2$ invariant. Altogether, model is thus characterized by a $\mathbb{Z}_2 \times \mathbb{Z}_2$ invariant, which encodes the polarizations of the edges in the two directions [19].

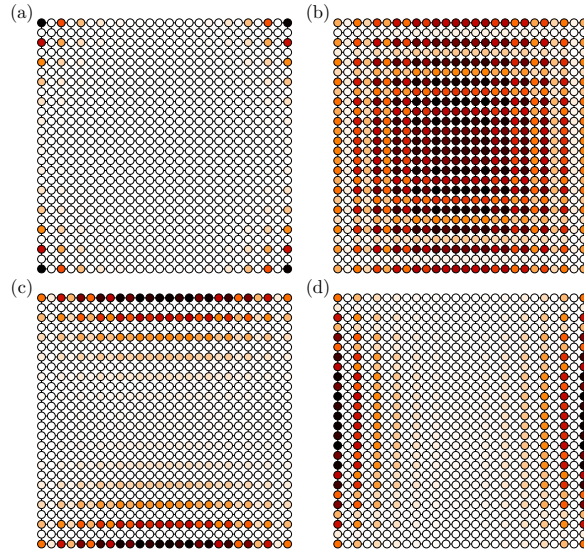


Figure 61: Wave function (modulus squared) of the single-particle excitation closest to zero energy for the second-order topological insulator in Eq. (7.98), evaluated on a finite-size lattice. Darker color indicates larger magnitude of the wave functions. The parameters are chosen as: (a) $t_{x,1} = 0.7$, $t_{y,1} = 0.6$; (b) $t_{x,1} = 1.2$, $t_{y,1} = 1.5$; (c) $t_{x,1} = 1.2$, $t_{y,1} = 0.7$; and (d) $t_{x,1} = 0.7$, $t_{y,1} = 1.5$. We choose $t_{x,2} = t_{y,2} = 1$ in all cases.

The various cases can be illustrated by plotting the wave function (modulus squared) of the excitation which is closest to zero energy, see Fig. 61. When the boundary Hamiltonians for both the x - and the y -edges are in the topological phase, the model exhibits corner states

[panel (a)]. The excitation corresponds to a pure bulk state when both edge Hamiltonians are in the nontopological phase [panel (b)]. When only one of the two boundary Hamiltonians is in the topological phase, the excitations are localized at the boundary in the topological direction, but delocalized in the other one [panels (c) and (d)].

7.7 A1: Boundary Green function and reflection matrix

In this appendix, we present an alternative derivation of the relation (7.17) between reflection matrix and boundary Green function.

We start by noting that if we multiply $\psi_n^\dagger$ from the left in Eq. (7.19), we find that $\text{Im}\psi_n^\dagger V_n^\dagger \psi_{n+1}$ is independent of n . Thus, we can define the current operator

$$\hat{\mathcal{I}}_n = i\Omega_n = i \begin{pmatrix} 0 & -V_n \\ V_n^\dagger & 0 \end{pmatrix}, \quad (7.108)$$

so that the current $\mathcal{I} = \Psi(n)^\dagger \hat{\mathcal{I}}_n \Psi(n)$ becomes independent of n . Here, we use the notation $\Psi(n) = [\psi(n+1)^T, \psi(n)^T]^T$.

It is more convenient to have a current operator which is independent of cell index n . To this end, we introduce the transformation

$$\Psi(n) \rightarrow O_n \Psi(n), \quad \hat{\mathcal{I}}_n \rightarrow (O_n^\dagger)^{-1} \hat{\mathcal{I}}_n O_n^{-1} = \sigma_y \otimes \mathbb{I} \quad (7.109)$$

with $O_n = \text{diag}(1, V_n)$. Here, the Pauli matrices $\sigma_{x,y,z}$ act in the two-component space of $\Psi(n)$. In the transformed basis, the matrix M_n defined in Eq. (7.20) becomes

$$M_n \rightarrow O_n M_n O_{n-1}^{-1} = \begin{pmatrix} (V_n^\dagger)^{-1} g_n^{-1} & -(V_n^\dagger)^{-1} \\ V_n & 0 \end{pmatrix}, \quad (7.110)$$

so that the relation $\Psi(n+1) = M_n \Psi(n)$ remains unchanged.

We can further diagonalize the current operator by introducing a unitary transformation

$$P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}, \quad (7.111)$$

such that

$$\Psi(n) \rightarrow P \Psi(n), \quad M_n \rightarrow P M_n P^\dagger. \quad (7.112)$$

Then, the current operator becomes

$$\hat{\mathcal{I}} = \sigma_z \otimes \mathbb{I}. \quad (7.113)$$

In this new basis, the upper and lower components (of length M each) of $\Psi(n)$ describe the right- and left-moving states at unit cell n .

If we attach normal metal leads from left and right to the first and the last unit cell of the quasi-1D system of N cells, we can define the scattering matrix

$$S_N = \begin{pmatrix} r'_N & t_N \\ t'_N & r_N \end{pmatrix}, \quad (7.114)$$

where the 2×2 block structure reflects the two leads, r_N, r'_N are reflection matrices and t_N, t'_N are transmission matrices. The scattering matrix relates the outgoing states on left and right to the incoming states. After the basis transformations, we have

$$\Psi(n) = \begin{pmatrix} \psi(n+1) - iV_n \psi(n) \\ \psi(n+1) + iV_n \psi(n) \end{pmatrix}. \quad (7.115)$$

Here the upper and lower components are outgoing (incoming) and incoming (outgoing) states on the right (left).

In the new basis, the transfer matrix can be written as [29]

$$\mathcal{M}_N = \begin{pmatrix} t'_N - r_N t_N^{-1} r'_N & r_N t_N^{-1} \\ -t_N^{-1} r'_N & t_N^{-1} \end{pmatrix}. \quad (7.116)$$

If we add one unit cell to the system, we have $\mathcal{M}_{N+1} = M_{N+1} \mathcal{M}_N$, where M_N in the new basis becomes

$$M_N = \frac{1}{2} \begin{pmatrix} (V_N^\dagger)^{-1} g_N^{-1} - iV_N - i(V_N^\dagger)^{-1} & (V_N^\dagger)^{-1} g_N^{-1} - iV_N + i(V_N^\dagger)^{-1} \\ (V_N^\dagger)^{-1} g_N^{-1} + iV_N - i(V_N^\dagger)^{-1} & (V_N^\dagger)^{-1} g_N^{-1} + iV_N + i(V_N^\dagger)^{-1} \end{pmatrix} \quad (7.117)$$

by straightforward algebra. We then find the recurrence relations

$$r_{N+1} t_{N+1}^{-1} = \frac{1}{2} [(V_{N+1}^\dagger)^{-1} g_{N+1}^{-1} - iV_{N+1}] (r_N + \mathbb{I}) t_N^{-1} - \frac{i}{2} (V_{N+1}^\dagger)^{-1} (r_N - \mathbb{I}) t_N^{-1} \quad (7.118)$$

$$t_{N+1}^{-1} = \frac{1}{2} [(V_{N+1}^\dagger)^{-1} g_{N+1}^{-1} + iV_{N+1}] (r_N + \mathbb{I}) t_N^{-1} - \frac{i}{2} (V_{N+1}^\dagger)^{-1} (r_N - \mathbb{I}) t_N^{-1}. \quad (7.119)$$

for the reflection and transmission matrices. By adding and subtracting these two equations, we find

$$(r_{N+1} + \mathbb{I}) t_{N+1}^{-1} = (V_{N+1}^\dagger)^{-1} g_{N+1}^{-1} (r_N + \mathbb{I}) t_N^{-1} - i(V_{N+1}^\dagger)^{-1} (r_N - \mathbb{I}) t_N^{-1} \quad (7.120)$$

$$(r_{N+1} - \mathbb{I}) t_{N+1}^{-1} = -iV_{N+1} (r_N + \mathbb{I}) t_N^{-1}. \quad (7.121)$$

From the last equation, we have

$$t_N^{-1} = i(r_N + \mathbb{I})^{-1} V_{N+1}^{-1} (r_{N+1} - \mathbb{I}) t_{N+1}^{-1}. \quad (7.122)$$

Inserting this expression back into the previous one, we find

$$(r_{N+1} + \mathbb{I}) t_{N+1}^{-1} = i(V_{N+1}^\dagger)^{-1} [g_{N+1}^{-1} (r_N + \mathbb{I}) - i(r_N - \mathbb{I})] (r_N + \mathbb{I})^{-1} V_{N+1}^{-1} (r_{N+1} - \mathbb{I}) t_{N+1}^{-1}. \quad (7.123)$$

Multiplying t_{N+1} from the right, we obtain

$$(r_{N+1} + \mathbb{I})(r_{N+1} - \mathbb{I})^{-1} = i(V_{N+1}^\dagger)^{-1} [g_{N+1}^{-1} - i(r_N - \mathbb{I})(r_N + \mathbb{I})^{-1}] V_{N+1}^{-1}. \quad (7.124)$$

Defining

$$\mathcal{R}_N = (r_N - \mathbb{I})(r_N + \mathbb{I})^{-1}, \quad (7.125)$$

we have

$$i\mathcal{R}_{N+1} = V_{N+1} [g_{N+1}^{-1} - i\mathcal{R}_N]^{-1} V_{N+1}^\dagger. \quad (7.126)$$

Comparing with Eq. (7.4), we find that the boundary Green function can be written as

$$G_N = V_N^{-1} i\mathcal{R}_N (V_N^\dagger)^{-1}. \quad (7.127)$$

Solving for r_N , we recover the relation (7.17) between the reflection matrix r_N and the boundary Green function G_N .

7.8 A2: Topological invariants

Single particle Hamiltonians can be classified based on nonspatial and nonunitary symmetries, i.e., the time-reversal, particle-hole, and chiral symmetries. This defines the ten Altland-Zirnbauer symmetry classes [30]. Ref. [9] discussed the topological classification of the ten symmetry classes in terms of the reflection matrix at the Fermi energy. In this section, we adapt this discussion to the language of the boundary Green functions and provide explicit expressions for the topological indices.

7.8.1 Bott periodicity

Symmetries

Time-reversal symmetry requires the existence of an antiunitary operator $\hat{\mathcal{T}} = \mathcal{U}_T \hat{\mathcal{K}}$ with a unitary matrix $\mathcal{U}_T$ and complex conjugation operator $\hat{\mathcal{K}}$, such that

$$\hat{\mathcal{T}}O(\mathbf{k}_\perp) = O(-\mathbf{k}_\perp)\hat{\mathcal{T}}, \quad O = h_n, V_n \quad \forall n. \quad (7.128)$$

Correspondingly, the boundary Green function obeys the relation

$$\hat{\mathcal{T}}G_n(\mathbf{k}_\perp, \omega) = G_n(-\mathbf{k}_\perp, \omega^*)\hat{\mathcal{T}}, \quad (7.129)$$

which can be proven iteratively using the Dyson equation [see Eq. (7.4)]. If we apply the time-reversal operator twice, we obtain

$$[\mathcal{U}_T^* \mathcal{U}_T, h_n] = [\mathcal{U}_T^* \mathcal{U}_T, V_n] = [\mathcal{U}_T^* \mathcal{U}_T, V_n^\dagger] = 0, \quad \forall n. \quad (7.130)$$

By Schur's lemma, we conclude that $\mathcal{U}_T^* \mathcal{U}_T = \exp(i\alpha)\mathbb{I}$ is a multiple of the identity matrix. Moreover, because of the unitarity of $\mathcal{U}_T$, we find $\exp(2i\alpha) = \pm 1$. Thus, there are two types of time-reversal symmetries,

$$\mathcal{U}_T^* \mathcal{U}_T = \pm \mathbb{I} \quad (7.131)$$

or $\mathcal{T}^2 = \pm 1$.

Particle-hole symmetry requires the existence of an antiunitary operator $\hat{\mathcal{C}} = \mathcal{U}_C \hat{\mathcal{K}}$ such that

$$\hat{\mathcal{C}}O(\mathbf{k}_\perp) = -O(-\mathbf{k}_\perp)\hat{\mathcal{C}}, \quad O = h_n, V_n \quad \forall n. \quad (7.132)$$

Here, $\mathcal{U}_C$ is a unitary matrix. The constraint of particle-hole symmetry on the boundary Green function is

$$\hat{\mathcal{C}}G_n(\mathbf{k}_\perp, \omega) = -G_n(-\mathbf{k}_\perp, -\omega^*)\hat{\mathcal{C}}. \quad (7.133)$$

Similar to the case of time-reversal symmetry, we find that there are two types of particle-hole symmetries given by

$$\mathcal{U}_C^* \mathcal{U}_C = \pm \mathbb{I}. \quad (7.134)$$

or $\hat{\mathcal{C}}^2 = \pm 1$.

Chiral symmetry requires the existence of a unitary operator $\hat{\mathcal{S}} = \mathcal{U}_S$ represented by a unitary matrix $\mathcal{U}_S$, such that

$$\{\hat{\mathcal{S}}, h_n\} = \{\hat{\mathcal{S}}, V_n\} = \{\hat{\mathcal{S}}, V_n^\dagger\} = 0, \quad \forall n. \quad (7.135)$$

Because of chiral symmetry, the boundary Green function fulfills

$$\mathcal{U}_S G_n(\mathbf{k}_\perp, \omega) = -G_n(\mathbf{k}_\perp, -\omega) \mathcal{U}_S. \quad (7.136)$$

Applying chiral symmetry twice, we find that $\mathcal{U}_S^2 = \exp(i\alpha)\mathbb{I}$. Redefining $\mathcal{U}_S \rightarrow \mathcal{U}_S \exp(-i\alpha/2)$, we have

$$\mathcal{U}_S^2 = \mathbb{I}, \quad (7.137)$$

which implies $\mathcal{U}_S = \mathcal{U}_S^\dagger$. When both time-reversal and particle-hole symmetry exist, the system also has chiral symmetry with $\mathcal{U}_S = \mathcal{U}_T \mathcal{U}_C^*$. If either time-reversal or particle-hole symmetry is absent, chiral symmetry is also absent. However, chiral symmetry can exist when both time-reversal and particle-hole symmetry are absent.

The absence or presence of time-reversal ($\hat{\mathcal{T}}$), particle-hole ($\hat{\mathcal{C}}$), and chiral symmetries ($\hat{\mathcal{S}}$) of the various kinds according to Eqs. (7.131), (7.134), and (7.137) leads to the ten Altland-Zirnbauer symmetry classes, as summarized in the first two columns of Table 1.

Dimensional reduction

In all symmetry classes with a chiral symmetry, the reflection matrix (the subscript N is omitted for simplicity) satisfies the constraint

$$\mathcal{U}_S r(\mathbf{k}_\perp, \omega) = r^\dagger(\mathbf{k}_\perp, -\omega) \mathcal{U}_S, \quad (7.138)$$

which follows from Eqs. (7.17) and (7.136). This implies that in the limit $\omega \rightarrow 0$, the matrix $\mathcal{U}_S r$ is hermitian where we use the shorthand $r = \lim_{\omega \rightarrow 0} r(\omega)$. This enables one to define an effective Hamiltonian in $d - 1$ dimensions as [9]

$$H_{d-1} = \mathcal{U}_S r = \lim_{\omega \rightarrow 0} \mathcal{U}_S \frac{\mathbb{I} - iVGV^\dagger}{\mathbb{I} + iVGV^\dagger}. \quad (7.139)$$

This Hamiltonian is again gapped as it has eigenvalues ± 1 and has no chiral symmetry.

In the absence of chiral symmetry, the eigenvalues of the reflection matrix r at the Fermi energy are in general complex. One can then construct an effective Hamiltonian in $d - 1$ dimensions by doubling the degrees of freedom, [9]

$$H_{d-1} = \begin{pmatrix} 0 & r \\ r^\dagger & 0 \end{pmatrix} = \lim_{\omega \rightarrow 0} \begin{pmatrix} 0 & \frac{\mathbb{I} - iVGV^\dagger}{\mathbb{I} + iVGV^\dagger} \\ \frac{\mathbb{I} + iVGV^\dagger}{\mathbb{I} - iVGV^\dagger} & 0 \end{pmatrix}. \quad (7.140)$$

Because of the identity

$$\{H_{d-1}, \sigma_z \otimes \mathbb{I}\} = 0, \quad (7.141)$$

the effective Hamiltonian acquires a chiral symmetry with $\mathcal{U}_S = \sigma_z \otimes \mathbb{I}$.

Classes A and AIII.—Hamiltonians in classes A and AIII are thus transformed into one another under this process of dimensional reduction. This is referred to as complex Bott periodicity (mod 2).

Now we take time-reversal and particle-hole symmetries into account, in order to reproduce the real Bott periodicity for the remaining eight classes of the ten-fold table.

Classes AI and AII.—We first focus on classes AI and AII with only time-reversal symmetry. Let us denote $r(\mathbf{k}_\perp) = \lim_{\omega \rightarrow 0} r(\mathbf{k}_\perp, \omega)$. By Eqs. (7.128) and (7.129), time-reversal symmetry restricts the reflection matrix at the Fermi energy to obey

$$r^\dagger(\mathbf{k}_\perp) \mathcal{U}_T = \mathcal{U}_T r^*(-\mathbf{k}_\perp). \quad (7.142)$$

Taking the hermitian conjugate and multiplying by $\mathcal{U}_T$ from both sides, we obtain

$$r(\mathbf{k}_\perp) \mathcal{U}_T = \mathcal{U}_T r^T(-\mathbf{k}_\perp). \quad (7.143)$$

According to Eq. (7.140), the effective Hamiltonian H_{d-1} has chiral symmetry. In addition, we can confirm using Eqs. (7.142) and (7.143) that it also acquires time-reversal symmetry with $\hat{\mathcal{T}} = (\sigma_x \otimes \mathcal{U}_T) \hat{\mathcal{K}}$ since

$$(\sigma_x \otimes \mathcal{U}_T) H_{d-1} (-\mathbf{k}_\perp)^* = H_{d-1}(\mathbf{k}_\perp) (\sigma_x \otimes \mathcal{U}_T), \quad (7.144)$$

and particle-hole symmetry with $\hat{\mathcal{C}} = (i\sigma_y \otimes \mathcal{U}_T) \hat{\mathcal{K}}$ since

$$(i\sigma_y \otimes \mathcal{U}_T) H_{d-1} (-\mathbf{k}_\perp)^* = -H_{d-1}(\mathbf{k}_\perp) (i\sigma_y \otimes \mathcal{U}_T). \quad (7.145)$$

Notice that time-reversal and particle-hole symmetries combine into the chiral symmetry as expected. The type of time-reversal and particle-hole symmetry of the effective Hamiltonian H_{d-1} can be determined from

$$(\sigma_x \otimes \mathcal{U}_T)^* (\sigma_x \otimes \mathcal{U}_T) = \mathbb{I}_2 \otimes (\mathcal{U}_T^* \mathcal{U}_T) \quad (7.146)$$

$$(i\sigma_y \otimes \mathcal{U}_T)^* (i\sigma_y \otimes \mathcal{U}_T) = -\mathbb{I}_2 \otimes (\mathcal{U}_T^* \mathcal{U}_T). \quad (7.147)$$

Hence, dimensional reduction transforms the classes AI and AIII into CI and DIII, respectively.

Classes C and D.—Next, consider classes C and D with only particle-hole symmetry. Similar to the previous case, particle-hole symmetry implies that the reflection matrix at the Fermi energy obeys

$$r(\mathbf{k}_\perp)\mathcal{U}_C = \mathcal{U}_C r^*(-\mathbf{k}_\perp) \quad (7.148)$$

$$r^\dagger(\mathbf{k}_\perp)\mathcal{U}_C = \mathcal{U}_C r^T(-\mathbf{k}_\perp). \quad (7.149)$$

The effective Hamiltonian H_{d-1} according to Eq. (7.140) has chiral symmetry. In addition, we can establish using Eqs. (7.148) and (7.149) that H_{d-1} has time-reversal symmetry with $\hat{\mathcal{T}} = (\mathbb{I}_2 \otimes \mathcal{U}_T)\hat{\mathcal{K}}$ since

$$(\mathbb{I}_2 \otimes \mathcal{U}_C)H_{d-1}(-\mathbf{k}_\perp)^* = H_{d-1}(\mathbf{k}_\perp)(\mathbb{I}_2 \otimes \mathcal{U}_C) \quad (7.150)$$

and particle-hole symmetry with $\hat{\mathcal{C}} = (\sigma_z \otimes \mathcal{U}_T)\hat{\mathcal{K}}$ since

$$(\sigma_z \otimes \mathcal{U}_C)H_{d-1}(-\mathbf{k}_\perp)^* = -H_{d-1}(\mathbf{k}_\perp)(\sigma_z \otimes \mathcal{U}_C). \quad (7.151)$$

Time-reversal and particle-hole symmetries combine as

$$(\mathbb{I}_2 \otimes \mathcal{U}_C)(\sigma_z \otimes \mathcal{U}_C)^* = \pm \sigma_z \otimes \mathbb{I}, \quad (7.152)$$

and hence produce the chiral symmetry as expected.

The types of the new time-reversal and particle-hole symmetry are determined from

$$(\mathbb{I}_2 \otimes \mathcal{U}_C)(\mathbb{I}_2 \otimes \mathcal{U}_C)^* = \mathbb{I}_2 \otimes (\mathcal{U}_C \mathcal{U}_C^*) \quad (7.153)$$

$$(\sigma_z \otimes \mathcal{U}_C)(\mathbb{I}_2 \otimes \mathcal{U}_C)^* = \mathbb{I}_2 \otimes (\mathcal{U}_C \mathcal{U}_C^*). \quad (7.154)$$

Hence, classes C and D transform into CII and BDI respectively under the procedure of dimensional reduction.

Classes CI, CII, DIII, and BDI.—Finally, we focus on the symmetry classes where time-reversal, particle-hole, and chiral symmetries are all present. In this situation, we have $\mathcal{U}_S = \mathcal{U}_T \mathcal{U}_C^*$. Then, $\mathcal{U}_S^2 = 1$ implies $\mathcal{U}_T^* \mathcal{U}_C \mathcal{U}_S^* = 1$. This can be used to show that

$$\mathcal{U}_S \mathcal{U}_C = \mathcal{U}_C \mathcal{U}_S^* (\mathcal{U}_T^* \mathcal{U}_C) (\mathcal{U}_T^* \mathcal{U}_T). \quad (7.155)$$

Notice that $\mathcal{U}_C^* \mathcal{U}_C = \pm 1$ and $\mathcal{U}_T^* \mathcal{U}_T = \pm 1$ are just numbers.

The effective Hamiltonian H_{d-1} , in this case defined by Eq. (7.139), has no chiral symmetry and can thus have either time-reversal or particle-hole symmetry, but not both. Using Eq. (7.155), we find that H_{d-1} has the property

$$H_{d-1}(\mathbf{k}_\perp)\mathcal{U}_C = (\mathcal{U}_C^* \mathcal{U}_C)(\mathcal{U}_T^* \mathcal{U}_T)\mathcal{U}_C H_{d-1}(-\mathbf{k}_\perp)^*. \quad (7.156)$$

Thus, $\mathcal{U}_C \hat{\mathcal{K}}$ defines a time-reversal symmetry or particle-hole symmetry for H_{d-1} , depending on whether the quantity $(\mathcal{U}_C^* \mathcal{U}_C)(\mathcal{U}_T^* \mathcal{U}_T)$ takes on the value +1 or -1, respectively. Whether this symmetry squares to +1 or -1 is determined by $\mathcal{U}_C^* \mathcal{U}_C$. This implies that under the procedure of dimensional reduction, classes CI, CII, DIII, and BDI transform into C, AII, D, and AI, respectively.

Using this procedure of dimensional reduction, a d -dimensional Hamiltonian in one symmetry class is related to a $(d-1)$ -dimensional Hamiltonian H_{d-1} in another symmetry class. Yet, both Hamiltonians have the same topological invariants [9]. This reproduces the Bott periodicity of the topological classification of symmetry classes [4, 5, 6]. In particular, classes A and AIII are transformed into one another, producing the complex Bott periodicity (mod 2). The remaining eight classes with antiunitary symmetries are shifted by one in Table 1, giving rise to the real Bott periodicity (mod 8). This is summarized in Table 1.

Table 1: Periodic table of topological insulators and superconductors in $d = 0, \dots, 7$ dimensions. The first column denotes the ten symmetry classes of fermionic Hamiltonians, characterized by the absence (0) or presence ($\pm$ or 1) of time-reversal ($\hat{\mathcal{T}}$), particle-hole ($\hat{\mathcal{C}}$), and chiral symmetries ($\hat{\mathcal{S}}$). Time reversal and particle hole symmetry exist in two types denoted by $\pm$. Classes which support only trivial phases are denoted by "-", while classes with nontrivial topological classifications are indicated by the type of topological invariant ($\mathbb{Z}, 2\mathbb{Z}, \mathbb{Z}_2$).

Class	$\hat{\mathcal{T}}$	$\hat{\mathcal{C}}$	$\hat{\mathcal{S}}$	0	1	2	3	4	5	6	7
A	0	0	0	$\mathbb{Z}$	-	$\mathbb{Z}$	-	$\mathbb{Z}$	-	$\mathbb{Z}$	-
AIII	0	0	1	-	$\mathbb{Z}$	-	$\mathbb{Z}$	-	$\mathbb{Z}$	-	$\mathbb{Z}$
AI	+	0	0	$\mathbb{Z}$	-	-	-	$2\mathbb{Z}$	-	$\mathbb{Z}_2$	$\mathbb{Z}_2$
BDI	+	+	1	$\mathbb{Z}_2$	$\mathbb{Z}$	-	-	-	$2\mathbb{Z}$	-	$\mathbb{Z}_2$
D	0	+	0	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}$	-	-	-	$2\mathbb{Z}$	-
DIII	-	+	1	-	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}$	-	-	-	$2\mathbb{Z}$
AII	-	0	0	$2\mathbb{Z}$	-	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}$	-	-	-
CII	-	-	1	-	$2\mathbb{Z}$	-	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}$	-	-
C	0	-	0	-	-	$2\mathbb{Z}$	-	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}$	-
CI	+	-	1	-	-	-	$2\mathbb{Z}$	-	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}$

7.8.2 Topological invariants

From 1D Hamiltonians, dimensional reduction produces 0D Hamiltonians. For classes without symmetry between positive and negative eigenenergies (classes A, AI, and AII), the topological invariant in 0D is given by the imbalance between positive and negative eigenenergies (measured relative to the Fermi energy). This corresponds to a $\mathbb{Z}$ index, except in class AII, where this number is always even because of Kramers degeneracy ($2\mathbb{Z}$ index). Of the remaining classes with symmetry between positive and negative-energy spectra, D and BDI are special as there is no level repulsion between the levels of a pair with positive and negative energies [30, 31]. These classes have a $\mathbb{Z}_2$ topological invariant, which is associated with fermion parity and can be expressed as the sign of a Pfaffian

$$\mathcal{Q} = \text{sgn}[\text{Pf}(iH_0)], \quad (7.157)$$

where the Hamiltonian H_0 is the zero-dimensional Hamiltonian, which is antisymmetric in an appropriate basis. The remaining five symmetry classes are always topologically trivial.

1D systems

The dimensional reduction implies that in one dimension, there are topological phases in symmetry classes AIII, BDI, CII, DIII, and D. The topological invariants in 0D immediately provide explicit expressions for boundary topological invariants in 1D.

Dimensional reduction connects the classes AIII, BDI, and CII in 1D to classes with $\mathbb{Z}$ or $2\mathbb{Z}$ indices in 0D. Using Eq. (7.139), the 1D classes are therefore characterized by the boundary topological invariant

$$\mathcal{Q} = \nu\left(\lim_{\omega \rightarrow 0} \mathcal{U}_S \frac{\mathbb{I} - iVGV^\dagger}{\mathbb{I} + iVGV^\dagger}\right) - \frac{M}{2}, \quad (7.158)$$

where $\nu(A)$ denotes the number of negative eigenvalues of the hermitian matrix A . The constant $M/2$ ensures that the trivial phase has topological invariant $\mathcal{Q} = 0$. This invariant indeed counts the imbalance between positive and negative eigenenergies of H_{d-1} .

By the bulk-boundary correspondence, these topological invariants can also be computed from the bulk 1D Hamiltonian $H(k)$. Due to the presence of chiral symmetry, the bulk Hamil-

tonian can be written as

$$H(k) = \begin{pmatrix} 0 & h(k) \\ h^\dagger(k) & 0 \end{pmatrix}, \quad (7.159)$$

and the corresponding bulk topological invariant is given by [32, 33]

$$\mathcal{Q} = \frac{1}{2\pi i} \int_0^{2\pi} dk \frac{d}{dk} \log \det h(k), \quad (7.160)$$

which is a winding number.

In class DIII, H_{d-1} has particle-hole symmetry realized via a unitary matrix $\mathcal{U}_C$ with $\mathcal{U}_C^* \mathcal{U}_C = 1$. This implies that $\mathcal{U}_C = \mathcal{U}_C^T$, so that it is possible to write $\mathcal{U}_C = \mathcal{V}_C \mathcal{V}_C^T$ with a unitary matrix $\mathcal{V}_C$. By the unitary transformation

$$H_{d-1} \rightarrow \mathcal{V}_C^\dagger H_{d-1} \mathcal{V}_C,$$

H_{d-1} becomes antisymmetric and purely imaginary. According to Eq. (7.157), the boundary topological invariant for class DIII can thus be expressed as

$$\mathcal{Q} = \text{Pf}(iH_{d-1}) = \text{Pf}(\lim_{\omega \rightarrow 0} i\mathcal{V}_C^\dagger \mathcal{U}_S \frac{\mathbb{I} - iVGV^\dagger}{\mathbb{I} + iVGV^\dagger} \mathcal{V}_C). \quad (7.161)$$

Again, this invariant can also be computed from the bulk Hamiltonian $H(k)$ as [34]

$$\mathcal{Q} = \frac{\text{Pf}(\mathcal{U}_T h(\pi))}{\text{Pf}(\mathcal{U}_T h(0))} \frac{\sqrt{\det h(0)}}{\sqrt{\det h(\pi)}}, \quad (7.162)$$

where $h(k)$ is defined as in Eq. (7.159).

Class D also allows us to transform H_{d-1} as

$$H_{d-1} \rightarrow \begin{pmatrix} \mathcal{V}_C & 0 \\ 0 & i\mathcal{V}_C \end{pmatrix}^\dagger H_{d-1} \begin{pmatrix} \mathcal{V}_C & 0 \\ 0 & i\mathcal{V}_C \end{pmatrix} = \begin{pmatrix} 0 & i\mathcal{V}_C^\dagger r \mathcal{V}_C \\ -i\mathcal{V}_C^\dagger r^\dagger \mathcal{V}_C & 0 \end{pmatrix}, \quad (7.163)$$

where $\mathcal{V}_C^\dagger r \mathcal{V}_C$ is purely real. Thus, the boundary topological invariant takes the form

$$\mathcal{Q} = \text{Pf}(iH_{d-1}) = \det \left(\lim_{\omega \rightarrow 0} \frac{\mathbb{I} - iVGV^\dagger}{\mathbb{I} + iVGV^\dagger} \right) \quad (7.164)$$

according to Eq. (7.157). The corresponding bulk topological invariant is [18]

$$\mathcal{Q} = \text{sgn} \left(\frac{\text{Pf} H(0)}{\text{Pf} H(\pi)} \right). \quad (7.165)$$

2D systems

By the procedure of dimensional reduction, the boundary invariants of 2D topological phases correspond to bulk invariants of 1D Hamiltonians. Using the 1D bulk invariant for classes AIII, BDI, and CII in Eq. (7.160), we obtain the boundary topological invariants for 2D systems in classes A, C, and D expressed in terms of the boundary Green function,

$$\mathcal{Q} = \frac{1}{2\pi i} \int_0^{2\pi} dk \frac{d}{dk} \text{tr} \log \left(\lim_{\omega \rightarrow 0} \frac{\mathbb{I} - iVGV^\dagger}{\mathbb{I} + iVGV^\dagger} \right). \quad (7.166)$$

These boundary topological invariants can again alternatively be computed from a bulk topological index expressed in terms of the bulk Hamiltonian. This bulk topological index takes the form of a Chern number [10],

$$\mathcal{Q} = \frac{1}{2\pi} \int dk_x dk_y F_{k_x, k_y} \quad (7.167)$$

where the Berry curvature is defined as

$$F_{k_x, k_y} = \sum_{\alpha \in \text{occ.}} i \partial_{k_x} \langle u_\alpha(\mathbf{k}) | \partial_{k_y} | u_\alpha(\mathbf{k}) \rangle - (k_x \leftrightarrow k_y). \quad (7.168)$$

Here, $|u_\alpha(\mathbf{k})\rangle$ denotes the Bloch wave function of the α band. Note that in computing the Berry curvature, one sums over occupied bands only.

For 2D system in class DIII, the boundary topological invariant follows from the 1D bulk invariant for class D in Eq. (7.164). This yields

$$\mathcal{Q} = \text{sgn} \left(\lim_{\omega \rightarrow 0} \frac{\text{Pf}(\mathcal{V}_C^\dagger \mathcal{U}_S \frac{\mathbb{I} - iVG(0)V^\dagger}{\mathbb{I} + iVG(0)V^\dagger} \mathcal{V}_C)}{\text{Pf}(\mathcal{V}_C^\dagger \mathcal{U}_S \frac{\mathbb{I} - iVG(\pi)V^\dagger}{\mathbb{I} + iVG(\pi)V^\dagger} \mathcal{V}_C)} \right). \quad (7.169)$$

The corresponding bulk topological invariant of the bulk Hamiltonian $H(k_1, k_2)$ can be written as a product of topological invariants of 1D Hamiltonians [34]

$$\mathcal{Q}(H(k_1, k_2)) = \mathcal{Q}(H(k_1, 0)) \mathcal{Q}(H(k_1, \pi)), \quad (7.170)$$

where $\mathcal{Q}(H(k_1, k_2))$ with $k_2 = 0, \pi$ is given in Eq. (7.162).

Finally, the topological invariant of 2D system in class AII is given by

$$\mathcal{Q} = \lim_{\omega \rightarrow 0} \frac{\text{Pf}(\mathcal{U}_T \frac{\mathbb{I} - iVG(\pi)V^\dagger}{\mathbb{I} + iVG(\pi)V^\dagger})}{\text{Pf}(\mathcal{U}_T \frac{\mathbb{I} - iVG(0)V^\dagger}{\mathbb{I} + iVG(0)V^\dagger})} \frac{\sqrt{\det \frac{\mathbb{I} - iVG(0)V^\dagger}{\mathbb{I} + iVG(0)V^\dagger}}}{\sqrt{\det \frac{\mathbb{I} - iVG(\pi)V^\dagger}{\mathbb{I} + iVG(\pi)V^\dagger}}}, \quad (7.171)$$

according to the 1D bulk topological invariant for class DIII in Eq. (7.162). This topological invariant can be computed from the bulk via [35]

$$\mathcal{Q} = \prod_{\mathbf{K}} \frac{\text{Pf}(w(\mathbf{K}))}{\sqrt{\det(w(\mathbf{K}))}}, \quad (7.172)$$

where the matrix element of w is defined as

$$w_{mn}(\mathbf{k}) = \langle u_m(-\mathbf{k}) | \hat{\mathcal{T}} | u_n(\mathbf{k}) \rangle, \quad (7.173)$$

with occupied Bloch bands $|u_n(\mathbf{k})\rangle$.

3D systems

The 3D topological invariant in class AII, which transforms into DIII under dimensional reduction, takes the form

$$\mathcal{Q} = \mathcal{Q}(0) \mathcal{Q}(\pi) \quad (7.174)$$

with

$$\mathcal{Q}(k) = \lim_{\omega \rightarrow 0} \frac{\text{Pf}(\mathcal{U}_T \frac{\mathbb{I} - iVG(\pi, k)V^\dagger}{\mathbb{I} + iVG(\pi, k)V^\dagger})}{\text{Pf}(\mathcal{U}_T \frac{\mathbb{I} - iVG(0, k)V^\dagger}{\mathbb{I} + iVG(0, k)V^\dagger})} \frac{\sqrt{\det \frac{\mathbb{I} - iVG(0, k)V^\dagger}{\mathbb{I} + iVG(0, k)V^\dagger}}}{\sqrt{\det \frac{\mathbb{I} - iVG(\pi, k)V^\dagger}{\mathbb{I} + iVG(\pi, k)V^\dagger}}}. \quad (7.175)$$

The topological invariants in classes AIII, CI, DIII, and CII can be written in terms of 2D bulk topological invariants. For simplicity, we do not write them down explicitly here.

7.9 Easy illustrative examples

7.9.1 Example 1

7.9.2 Example 2

8 The Non-Equilibrium Green Function (NEGF) Method

Kerem Y. Camsari, Shuvro Chowdhury and Supriyo Datta

The Non-Equilibrium Green Function (NEGF) method was established in the 1960's through the classic work of Schwinger, Kadanoff, Baym, Keldysh and others using many-body perturbation theory (MBPT) and the diagrammatic theory for non-equilibrium processes. Much of the literature is based on the original MBPT-based approach and this makes it inaccessible to those unfamiliar with advanced quantum statistical mechanics. We obtain the NEGF equations directly from a one-electron Schrödinger equation using relatively elementary arguments. These equations have been used to discuss many problems of great interest such as quantized conductance, (integer) quantum Hall effect, Anderson localization, resonant tunneling and spin transport without a systematic treatment of many-body effects. But it goes beyond purely coherent transport allowing us to include phase-breaking interactions (both momentum-relaxing and momentum-conserving as well as spin-conserving and spin-relaxing) within a self-consistent Born approximation. We believe that the scope and utility of the NEGF equations transcend the MBPT-based approach originally used to derive it. NEGF teaches us how to combine quantum dynamics with “contacts” much as Boltzmann taught us how to combine classical dynamics with “contacts”, using the word “contacts” in a broad figurative sense to denote all kinds of entropy-driven processes. We believe that this approach to “contact-ing” the Schrödinger equation should be of broad interest to anyone working on device physics or non-equilibrium statistical mechanics in general.

8.1 Introduction

The non-equilibrium Green's function (NEGF) method was pioneered in the 1960's by the classic work of Martin, Schwinger [1], Kadanoff, Baym [2], Keldysh [3] and others, which have been discussed in many review articles such as Danielewicz [4] and Mahan [5]. After the advent of mesoscopic physics in the 1980's this method was combined with the Landauer approach (See Refs. [6], [7], [8]) and this “NEGF-Landauer method” has been widely used in the field of nanoelectronics for device modeling and technology development (See [9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23] as a small subset of a vast literature). The chapter by Klimeck and Boykin in this volume provides an introduction to state-of-the-art quantum transport simulation tools based on this method. The present chapter is intended to serve a complementary purpose, namely to introduce the conceptual underpinnings of NEGF and to answer some frequently asked questions regarding them.

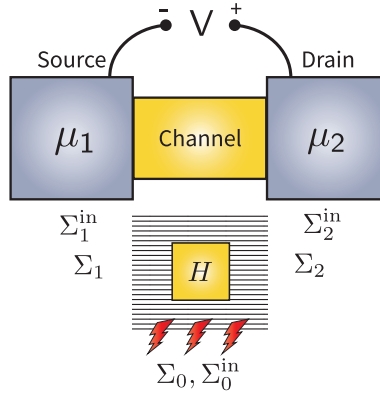


Figure 62: **A resistor:** Physical structure comprising a channel with two contacts, labeled source and drain. In the non-equilibrium Green function (NEGF) method the channel is described by a Hamiltonian $\mathbf{H}$ while its connection to the “contacts” are described by the self-energy functions $\mathbf{\Sigma}$, $\mathbf{\Sigma}^{\text{in}}$: subscripts 1 and 2 denote the physical contacts labeled source and drain, while the subscript 0 represents abstract contacts like the phonon bath.

The iconic device in modern electronics is the field effect transistor (FET), billions of which are a part of every smartphone. An oversimplified sketch of an FET is shown in Fig. 62 consisting of an active region marked the *channel* sandwiched between two highly conductive regions labeled the *source* and the *drain*. An FET also has a third terminal (not shown) which can be used to control the resistance $R = V/I$ of the channel, but we will not get into such “details”. We will focus just on the problem of calculating the steady-state charge current I for a given voltage V taking *quantum* effects into account.

Quantum mechanical descriptions usually start from a Hamiltonian matrix $\mathbf{H}$ whose eigenvalues give us the allowed energy levels. However, if we treat the structure like a closed system described just by an $\mathbf{H}$, an applied voltage will cause opposite charges to pile up in the source and drain, and we will have a capacitor and not a resistor. What makes it a resistor is the external battery that continually takes electrons out of the drain and inserts electrons back into the source to maintain an electrochemical potential difference: $\mu_1 - \mu_2 = qV$.

One way to describe such an open system is to describe the two contacts denoted 1 and 2 through self-energy functions $\mathbf{\Sigma}_{1,2}$, $\mathbf{\Sigma}_{1,2}^{\text{in}}$. The first of these $\mathbf{\Sigma}_{1,2}$ has a Hermitian part that modifies the Hamiltonian and an anti-Hermitian part that represents the rate at which electrons escape from the channel. The second of these $\mathbf{\Sigma}_{1,2}^{\text{in}}$ represents the inflow of the electrons into the channel from the contacts. In addition to these physical contacts there are abstract “contacts” representing the interactions of the electrons with the lattice and with other electrons as they traverse the channel and these are described the self-energy functions $\mathbf{\Sigma}_0$, $\mathbf{\Sigma}_0^{\text{in}}$. To apply the NEGF method to a given problem, there are two steps:

1. Identify the appropriate Hamiltonian $\mathbf{H}$ and self-energy functions $\mathbf{\Sigma}_m$, $\mathbf{\Sigma}_m^{\text{in}}$
2. Use these in the NEGF Eqs. (8.1-8.3) summarized in Sect. 8.2 to calculate quantities of interest like current, electron density or density of states.

The NEGF method is commonly viewed as an esoteric tool accessible only to a small group of specialists. To quote from two popular books [24, 25]:

...In this respect, the most difficult topic is probably the non-equilibrium Green’s function formalism of Chapter 4 ...

...Because of my aim of avoiding too heavy formalism I have chosen not to describe non-equilibrium Green’s function approaches to transport phenomena.

Courses detailing these approaches usually spend half the time on the formalism, finding the poles of the various Green's functions, and figuring out analytic continuations and so on ...

Both books are pedagogically outstanding and we cite them only to stress that NEGF is considered difficult even by the foremost practitioners in the field. And the reason is that the traditional discussion of NEGF is based on many-body perturbation theory (MBPT) which takes many semesters of advanced quantum mechanics to master.

However, as we have often noted [26], the only aspect of NEGF that really requires MBPT is in writing down the self-energy functions $\Sigma_0, \Sigma_0^{\text{in}}$ describing interactions within the channel. Everything else can be accomplished without the use of MBPT, including deriving the NEGF Eqs. (8.1-8.3) and writing down the non-interacting channel Hamiltonian $\mathbf{H}$ along with the self-energy functions for the physical contacts $\Sigma_{1,2}$ and $\Sigma_{1,2}^{\text{in}}$. This is the approach one of us has used in all his books starting from Chapter 8 of Ref. [1] and later in Ref. [28] and Part B of Ref. [29]. The recent book by Ghosh (Ref. [30]) is unique in describing both our approach and the MBPT-based one along with special non-perturbative approaches.

8.1.1 Decoupling NEGF from MBPT

Why is the standard treatment so heavily dependent on MBPT? We believe the reason is historical. Until 1990 it was common to regard the physical contacts as an unimportant detail and view the channel interactions as the essential physics. And so the original papers in the field dating back to the 1960's described the resistance in terms of $\Sigma_0, \Sigma_0^{\text{in}}$ and never even considered the $\Sigma_{1,2}, \Sigma_{1,2}^{\text{in}}$ describing the physical contacts.

What changed around 1990 was the widespread experimental measurement of the resistance of *ballistic conductors* where electrons zip through the channel like bullets with little or no interactions making Σ_0 and Σ_0^{in} negligible. Do such conductors have a resistance? The experimental answer was a resounding *yes*⁶. A current I flowing through a non-zero resistance R must generate a heat $I^2 R$. Heat is generated when energetic electrons give up some of their energy to the lattice making the atoms jiggle, which manifests itself as an increase in temperature.

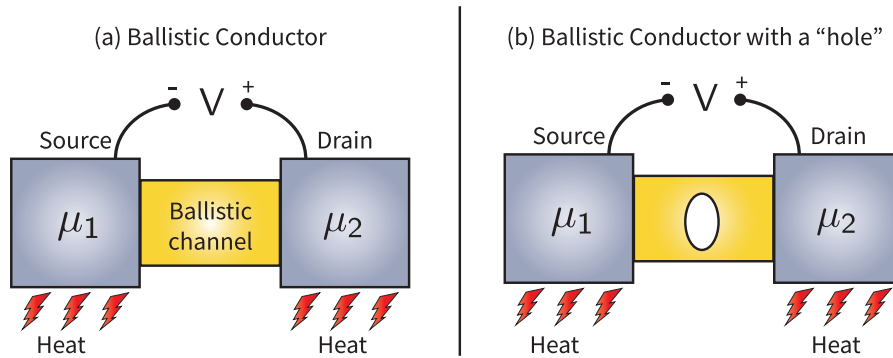


Figure 63: **Left:** In a ballistic conductor the resistance arises from the channel-contact interfaces but the heating occurs deeper in the contacts. **Right:** A ballistic conductor with a hole further accentuates the spatial separation between the microscopic cause of resistance and the resulting generation of heat.

But if there are no interactions in the channel, then how can heat be generated? The answer is that the heat is generated within the contacts (Fig. 63a) which is not surprising since

⁶For example Ref. [31] is one of the first experiments that established the fundamental limits of conductance in such ballistic conductors. Later, scientists have even measured the resistance of a single molecule [32]!

a bullet heats up the material it hits, and not the medium it flies through. A corollary of this observation is that although any resistance R is accompanied by the generation of I^2R heat, the heating need not occur exactly where the resistance is caused.

For example in the ballistic conductor the resistance is caused by the channel-contact interfaces, while the heating could occur away from the interface. This separation of the resistance R and its heat I^2R is even more stark if we consider a ballistic conductor with a hole in the middle (Fig. 63b). We would expect the hole to cause a large increase in resistance by obstructing the flow of electrons. But it does not have the *internal degrees of freedom* needed to get heated which must still occur far away somewhere in the contacts⁷.

However, pre-1990 it was not common to talk of resistance in a model without explicitly including interactions and dissipation in the channel⁸. Indeed much of the work on quantum transport was based on the Kubo formula which equates resistance to dissipation and then relates it to noise through the fluctuation-dissipation theorem. This is a powerful approach to linear response problems but it tends to create the impression that the origin of resistance is energy relaxation rather than momentum relaxation. But it now seems clear that resistance is caused by momentum exchange, and the resulting energy exchange need not occur at the same location. It is possible to create a zero-order model for resistance without including any Σ_0 or Σ_0^{in} and this approach at least has enormous pedagogical value.

The NEGF Eqs. (8.1-8.3) obtained in the next section from elementary quantum mechanics are the *same* as those obtained in the classic paper by Keldysh using MBPT [3]. The difference is that we obtain them considering only the physical contacts described by $\Sigma_{1,2}$ and $\Sigma_{1,2}^{\text{in}}$, while the classic treatment considers only the many-body interactions described by Σ_0 and Σ_0^{in} . Our approach makes it accessible to a much wider audience and for over a decade we have been teaching the NEGF equations to advanced undergraduate and beginning graduate students, both on-campus and online.

We believe the utility of this approach goes beyond the purely pedagogical. It decouples the NEGF Eqs. (8.1-8.3) from MBPT thus clearing the path for alternative approaches to Σ_0 and Σ_0^{in} including non-perturbative methods as well as phenomenological approaches. In this chapter we will present a simple example of the latter.

8.1.2 Outline

In Sect. 8.2 we will briefly summarize the NEGF equations followed by a brief outline of their derivation from the one-electron Schrödinger equation in Sect. 8.3. In Sect. 8.4 we present toy examples to give the reader a feeling for how the method is applied. Corresponding Python codes are also included for interested readers. We end in Sect. 8.5 with answers to a few commonly asked questions. For more detailed examples and discussions we refer the reader to Chapter 8 of Ref. [1], Ref. [28] and Part B of Ref. [29].

8.2 NEGF Equations

There are three NEGF equations as described below.

⁷An experimental observation that supports this viewpoint is current transport through carbon nanotubes. Despite being extremely small in diameter and volume, carbon nanotubes can carry enormous amounts of current and if there were any significant energy dissipation within the channel, they could not get rid of such excessive heat and would burn up.

⁸Though device engineers had started to recognize that momentum and energy relaxation lengths could be very different in late 80's. See for example, Ref. [33].

1. First is the equation for the quantum *density of states* (times 2π)⁹

$$\mathbf{A} = i [\mathbf{G}^R - \mathbf{G}^A]$$

where the advanced Green function $\mathbf{G}^A$ is the Hermitian conjugate of the retarded Green function $\mathbf{G}^R$ given by ($\mathbf{\Sigma} = \mathbf{\Sigma}_0 + \mathbf{\Sigma}_1 + \mathbf{\Sigma}_2$)

$$\mathbf{G}^R = [E\mathbf{I} - \mathbf{H} - \mathbf{\Sigma}]^{-1} \quad (8.1)$$

2. Next is the equation for the quantum *electron density* (times 2π) per unit energy

$$\mathbf{G}^n = \mathbf{G}^R \mathbf{\Sigma}^{\text{in}} \mathbf{G}^A \quad (8.2)$$

3. And finally the current per unit energy $\tilde{I}_m$ at contact m is given by

$$\tilde{I}_m = \frac{q}{h} \text{Trace} [\mathbf{\Sigma}_m^{\text{in}} \mathbf{A} - \mathbf{\Gamma}_m \mathbf{G}^n] \quad (8.3)$$

where $\mathbf{\Gamma}_m = i [\mathbf{\Sigma}_m - \mathbf{\Sigma}_m^\dagger]$.

We use a notation that is slightly different from that used in the conventional literature because it helps provide a physical picture for different quantities. For example we use $\mathbf{G}^n$ instead of the traditional $-i\mathbf{G}^<$ because $\mathbf{G}^n$ represents a matrix version of the *electron density*. The correspondences are summarized in a table to help the reader translate our equations as needed.

Table 2: NEGF symbols used in this chapter and their counterpart in the literature.

Conventional symbol	Physical Interpretation	Symbol used in this book
$-i\mathbf{G}^<$	Matrix electron density	$\mathbf{G}^n$
$+i\mathbf{G}^>$	Matrix hole density	$\mathbf{G}^p$
$-i\mathbf{\Sigma}^<$	In-scattering function	$\mathbf{\Sigma}^{\text{in}}$
$+i\mathbf{\Sigma}^>$	Out-scattering function	$\mathbf{\Sigma}^{\text{out}}$

8.3 NEGF equations from one-electron Schrödinger equation

Let us briefly indicate how the NEGF Eqs. (8.1-8.3) are obtained directly from the one-electron Schrödinger equation $E\psi = \mathbf{H}\psi$. The first step is to incorporate the effect of the *open boundary conditions* imposed by the contacts through two additional terms

$$E\psi = \mathbf{H}\psi + \underbrace{\mathbf{\Sigma}\psi}_{\text{OUTFLOW}} + \underbrace{\mathbf{s}}_{\text{INFLOW}} \quad (8.4)$$

Note that the modification of the Schrödinger equation by the $\mathbf{\Sigma}$ and the $\mathbf{s}$ terms lead to a conceptually different viewpoint of energy. Instead of being the resonant energies of a given system described by $\mathbf{H}$, energy now becomes an independent variable related to the incoming electrons from the contacts that excite the channel. (see Chapter 8 of [28] for a detailed discussion).

⁹Note that the quantum density of states is generally a dense matrix of size $N \times N$ for a lattice size of N ($2N \times 2N$ if electron spin is included). Its diagonal entries correspond to the classical local density of states but off-diagonal elements also contain useful information. For example, when the quantum density of states is multiplied by Pauli spin matrices, the diagonal elements provide local density of states *for a particular spin* as we show later by an example in Section 8.4.6. The matrix nature of $\mathbf{A}$ and $\mathbf{G}^n$ compared to their classical counterparts that are of size $N \times 1$ is due to the quantum generalization of these quantities.

8.3.1 NEGF Eqs. (8.1-8.2)

From Eq. (8.4) we write the wavefunction as

$$\psi = [EI - \mathbf{H} - \mathbf{\Sigma}]^{-1} \mathbf{s} = \mathbf{G}^R \mathbf{s}$$

making use of the definition of $\mathbf{G}^R$ from NEGF Eq. (8.1). Note that $\mathbf{\Sigma}$ is non-Hermitian so that $E\mathbf{I} - \mathbf{H} - \mathbf{\Sigma}$ is not singular for any real value of E and has a well-defined inverse $\mathbf{G}^R$.

Since different sources ‘ $\mathbf{s}$ ’ are incoherent, we cannot superpose the resulting ψ ’s from multiple sources. So we define bilinear quantities that *can* be superposed:

$$\underbrace{\psi\psi^\dagger}_{\mathbf{G}^n/(2\pi)} = \mathbf{G}^R \underbrace{\mathbf{s}\mathbf{s}^\dagger}_{\mathbf{\Sigma}^{\text{in}}/(2\pi)} \mathbf{G}^A$$

giving us NEGF Eq. (8.2).

In our description ψ is the one-electron wavefunction whose square gives the probability of finding an electron. When averaged over all electrons in an ensemble it gives the electron density. More generally $\psi\psi^\dagger$ is a matrix whose diagonal elements give the probabilities, and the off-diagonal elements give the correlations. For a more formal justification using field operators the reader can check the Appendix in Ref. [28].

Note that $\mathbf{s}\mathbf{s}^\dagger$ reflects the availability of electrons at a given energy in the contacts which fill the available states in the channel giving rise to $\psi\psi^\dagger$. Fermi functions are invoked only in the contacts and not in the channel.

8.3.2 NEGF Eq. (8.3)

Starting from the time-dependent version of the modified Schrödinger equation, Eq. (8.4),

$$i\hbar \frac{d}{dt} \psi = [\mathbf{H} + \mathbf{\Sigma}] \psi + \mathbf{s}$$

and its conjugate transpose (noting that $\mathbf{H}$ is a Hermitian matrix)

$$-i\hbar \frac{d}{dt} \psi^\dagger = \psi^\dagger [\mathbf{H} + \mathbf{\Sigma}^\dagger] + \mathbf{s}^\dagger$$

and making use of the relations

$$\psi = \mathbf{G}^R \mathbf{s} \quad \text{and} \quad \psi^\dagger = \mathbf{s}^\dagger \mathbf{G}^A,$$

we can write

$$\begin{aligned} i\hbar \frac{d}{dt} \psi\psi^\dagger &= \left(i\hbar \frac{d}{dt} \psi \right) \psi^\dagger + \psi \left(i\hbar \frac{d}{dt} \psi^\dagger \right) \\ &= ([\mathbf{H} + \mathbf{\Sigma}] \psi + \mathbf{s}) \psi^\dagger - \psi (\psi^\dagger [\mathbf{H} + \mathbf{\Sigma}^\dagger] + \mathbf{s}^\dagger) \\ &= [(\mathbf{H} + \mathbf{\Sigma}) \psi\psi^\dagger - \psi\psi^\dagger (\mathbf{H} + \mathbf{\Sigma}^\dagger)] + [\mathbf{s}\mathbf{s}^\dagger \mathbf{G}^A - \mathbf{G}^R \mathbf{s}\mathbf{s}^\dagger] \end{aligned}$$

Making use of

$$2\pi \psi\psi^\dagger = \mathbf{G}^n \quad \text{and} \quad 2\pi \mathbf{s}\mathbf{s}^\dagger = \mathbf{\Sigma}_{\text{in}}$$

we can write

$$\frac{d}{dt} \psi\psi^\dagger = \frac{[\mathbf{H}\mathbf{G}^n - \mathbf{G}^n\mathbf{H}] + [\mathbf{\Sigma}\mathbf{G}^n - \mathbf{G}^n\mathbf{\Sigma}^\dagger] + [\mathbf{\Sigma}^{\text{in}}\mathbf{G}^A - \mathbf{G}^R\mathbf{\Sigma}^{\text{in}}]}{i2\pi\hbar}. \quad (8.5)$$

We can interpret $\psi\psi^\dagger$ as the number operator so that its derivative represents the current operator. To find the change of any quantity A , we can multiply by the corresponding operator $\mathbf{A}_{\text{op}}$ and take the trace to obtain

$$\frac{dA}{dt} = -\frac{i}{h} \text{Trace} \left([\mathbf{H}\mathbf{G}^n - \mathbf{G}^n\mathbf{H}]\mathbf{A}_{\text{op}} + [\boldsymbol{\Sigma}\mathbf{G}^n - \mathbf{G}^n\boldsymbol{\Sigma}^\dagger]\mathbf{A}_{\text{op}} + [\boldsymbol{\Sigma}^{\text{in}}\mathbf{G}^A - \mathbf{G}^R\boldsymbol{\Sigma}^{\text{in}}]\mathbf{A}_{\text{op}} \right)$$

The first term represents the change in the quantity A inside the channel due to the action of the Hamiltonian $\mathbf{H}$ while the last two terms represent the “ A current” injected from the terminals. We can define the terminal current operator $\mathbf{I}_{\text{op}}$ as

$$\mathbf{I}_{\text{op}} = -\frac{i}{h} \left([\boldsymbol{\Sigma}\mathbf{G}^n - \mathbf{G}^n\boldsymbol{\Sigma}^\dagger] + [\boldsymbol{\Sigma}^{\text{in}}\mathbf{G}^A - \mathbf{G}^R\boldsymbol{\Sigma}^{\text{in}}] \right) \quad (8.6)$$

such that the “ A current” can be obtained from the trace of $\mathbf{A}_{\text{op}}\mathbf{I}_{\text{op}}$.

If we are interested only in the charge current, then we need the trace of $\mathbf{I}_{\text{op}}$ since the charge operator is an identity matrix (times the electronic charge). Noting that $\text{Trace}[\mathbf{AB}] = \text{Trace}[\mathbf{BA}]$ we can write

$$\text{Trace}(\mathbf{I}_{\text{op}}) = \frac{1}{h} \text{Trace}(\boldsymbol{\Sigma}^{\text{in}}\mathbf{A} - \boldsymbol{\Gamma}\mathbf{G}^n).$$

Both the left and the right hand sides of this equation are zero, since we are discussing steady state charge transport with no time variation. But the terms on the left can be separated into two parts, one associated with contact 1 and one with contact 2. They tell us the currents at contacts 1 and 2 respectively and the fact that they add up to zero is simply a statement of *Kirchhoff's law* for steady-state currents in circuits¹⁰.

With this in mind we can write for the current at contact m ($m = 1, 2$)

$$\tilde{I}_m = \frac{q}{h} \text{Trace}(\boldsymbol{\Sigma}_m^{\text{in}}\mathbf{A} - \boldsymbol{\Gamma}_m\mathbf{G}^n)$$

as stated in NEGF Eq. (8.3).

8.4 A simple example

In this Section we will outline the application of the NEGF Eqs. (8.1-8.2) to a simple example centered around a one-dimensional conductor with a barrier U and connected to two contacts with Fermi functions $f_1(E) = 1$ and $f_2(E) = 0$. The reader can consult Part B of Ref. [29] for more details.

¹⁰Note that Eq. (8.6) can also be used to obtain terminal *spin* currents if both sides of the equation is multiplied by Pauli spin matrices corresponding to a desired spin direction. In this case however the standard Kirchhoff's Laws for charge transport may not hold but it is still possible to design “spin-circuits” that obey Kirchhoff's laws [34].

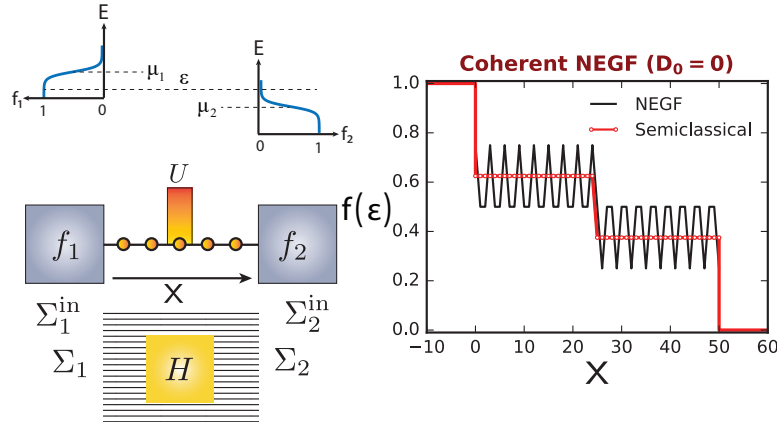


Figure 64: **Left-Top:** A simple one-dimensional conductor with a barrier U and connected to two contacts with Fermi functions $f_1(E) = 1$ and $f_2(E) = 0$. **Left-Bottom:** To analyze it using the NEGF Eqs. (8.1-8.2) we need the $\mathbf{H}$ to describe the channel and the $\mathbf{\Sigma}$'s to describe the physical contacts. Interactions in the channel are ignored in this example ($D_0 = 0$). **Right:** Plot of the occupation factor across the channel.

8.4.1 Hamiltonian $\mathbf{H}$

We describe the 1D channel using a simple tight-binding Hamiltonian $\mathbf{H}$ with ϵ on the diagonal and t on the upper and lower diagonals, as shown pictorially in Fig. 65. This leads to a dispersion relation of the form

$$E(k) = \epsilon + 2t \cos(ka) \quad (8.7a)$$

This cosine dispersion can approximate a parabolic one

$$E = E_c + \frac{\hbar^2 k^2}{2m} \quad (8.7b)$$

if we choose

$$E_c = \epsilon + 2t \quad (8.8a)$$

$$\text{and} \quad -t \equiv t_0 \equiv \frac{\hbar^2}{2ma^2}. \quad (8.8b)$$

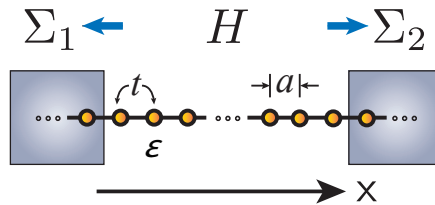


Figure 65: For the one-dimensional examples discussed in this chapter, we use the 1D Hamiltonian shown with on-site elements ϵ and nearest neighbor coupling t .

8.4.2 Self-energy due to contacts

The self-energy function for each contact has only one non-zero element corresponding to the point that is connected to that contact:

$$\mathbf{\Sigma}_1 = \begin{bmatrix} te^{ika} & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad \mathbf{\Sigma}_2 = \begin{bmatrix} 0 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 \\ 0 & \cdots & 0 & 0 & 0 \\ 0 & \cdots & 0 & 0 & te^{ika} \end{bmatrix}$$

8.4.3 Inscattering from contacts

The inscattering functions are given simply by the broadening functions times the corresponding occupation factor:

$$\mathbf{\Sigma}_1^{\text{in}} = \mathbf{\Gamma}_1 f_1 \quad \text{and} \quad \mathbf{\Sigma}_2^{\text{in}} = \mathbf{\Gamma}_2 f_2 \quad (8.9)$$

where the broadening functions are obtained from $\mathbf{\Gamma}_{1,2} = i[\mathbf{\Sigma}_{1,2} - \mathbf{\Sigma}_{1,2}^\dagger]$:

$$\mathbf{\Gamma}_1 = \frac{\hbar\nu}{a} \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad \mathbf{\Gamma}_2 = \frac{\hbar\nu}{a} \begin{bmatrix} 0 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 \\ 0 & \cdots & 0 & 0 & 0 \\ 0 & \cdots & 0 & 0 & 1 \end{bmatrix}$$

noting that the velocity $\nu = dE/(\hbar dk)$ which from the dispersion relation Eq. (8.7a) equals $-2at/\hbar \sin(ka)$.

Given these matrices it is straightforward to use the NEGF Eqs. (8.1-8.2) to evaluate the “electron density” $\mathbf{G}^n/2\pi$ and the “density of states” $\mathbf{A}/2\pi$ from which we can obtain the ℓ -th element of the occupation factor

$$f_\ell = \frac{G_{\ell,\ell}^n}{A_{\ell,\ell}}$$

which is plotted in Fig. 64. Note that the NEGF occupation factor shows oscillations due to quantum interference around the average values expected from a semiclassical picture. The semiclassical curve is based on three localized resistors in series, one each at the two interfaces, and one at the barrier:

$$\frac{R}{h/q^2} = \underbrace{\frac{1}{2}}_{\text{channel-source interface}} + \underbrace{\frac{1-T}{T}}_{\text{barrier}} + \underbrace{\frac{1}{2}}_{\text{channel-drain interface}} \quad (8.10)$$

where T is the transmission probability through the barrier given by

$$T(E) = \frac{(\hbar\nu/a)^2}{U^2 + (\hbar\nu/a)^2} \quad (8.11)$$

This is a standard result that can be obtained from scattering theory on a discrete lattice [28], next we will show how the NEGF equations lead to the same result.

8.4.4 Current

If we are not interested in the spatial variation of quantities like the occupation factor, but are only interested in the current that flows for a given voltage, we can simply let the channel represent one point, and treat the rest as contacts. This makes all matrices collapse to (1×1) reducing them to just numbers:

$$\begin{aligned}
 H &= \varepsilon + U \\
 \Sigma_1(E) &= te^{+ika} = \Sigma_2(E) \quad \text{where } ka = \cos^{-1} \left(\frac{E - \varepsilon}{2t} \right) \\
 \Sigma(E) &= \Sigma_1(E) + \Sigma_2(E) = 2te^{+ika} \\
 \Gamma_1 &= \frac{\hbar\nu}{a} = \Gamma_2 \\
 \Sigma_1^{\text{in}} &= \Gamma_1 f_1, \quad \Sigma_2^{\text{in}} = \Gamma_2 f_2 \quad \text{and} \quad \Sigma^{\text{in}} = \Gamma_1 f_1 + \Gamma_2 f_2
 \end{aligned}$$

Now we do not need a computer to evaluate the NEGF Eqs. (8.1-8.3). We can just do it by hand. First we calculate G^R from Eq. (8.1):

$$\begin{aligned}
 G^R(E) &= \frac{1}{E - (\varepsilon + U) - 2te^{ika}} \\
 &= \frac{1}{-U - i2t \sin(ka)} = \frac{1}{-U + i\hbar\nu/a}
 \end{aligned}$$

so that

$$\begin{aligned}
 G^A(E) &= \frac{1}{-U - i\hbar\nu/a} \\
 A(E) &= \frac{2\hbar\nu/a}{U^2 + (\hbar\nu/a)^2}
 \end{aligned}$$

Next from Eq. (8.2):

$$G^n = \frac{\hbar\nu/a}{U^2 + (\hbar\nu/a)^2} (f_1 + f_2)$$

so that from Eq. (8.3):

$$\begin{aligned}
 \tilde{I}(E) &= \frac{q}{h} \frac{(\hbar\nu/a)^2}{U^2 + (\hbar\nu/a)^2} (2f_1 - \overline{f_1 + f_2}) \\
 &= \frac{q}{h} \underbrace{\frac{(\hbar\nu/a)^2}{U^2 + (\hbar\nu/a)^2}}_{T(E)} (f_1 - f_2)
 \end{aligned}$$

The quantity $T(E)$ represents the transmission probability through the barrier U . Integrating $\tilde{I}(E)$ we get the total current. The quantity $T(E)$ is the same as the semiclassical transmission that was shown in Eq. (8.11).

8.4.5 Dephasing interactions

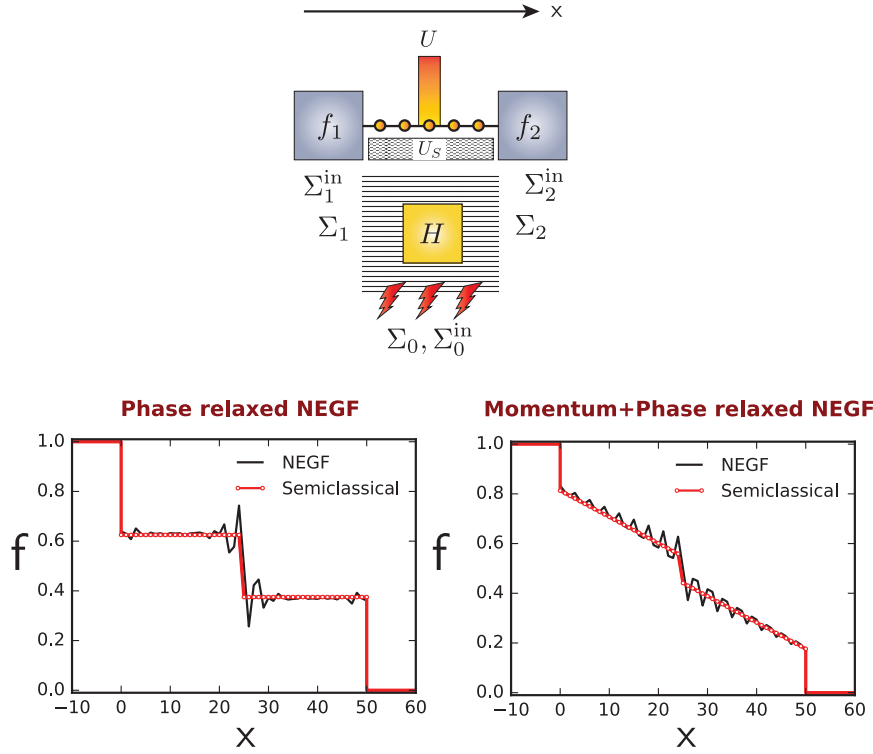


Figure 66: **Top:** Same problem as in Fig. 64 but we now include the effect of a random potential U_S . **Middle:** The effect of the random potential is incorporated through Σ_0 and Σ_0^{in} where the magnitude of scattering, D_0 , is $0.09 t_0^2$. **Bottom Left:** Plot of occupation factor across channel for phase-relaxing scattering. **Bottom right:** Plot of occupation factor across channel for phase and momentum-relaxing scattering.

What we have done so far represents coherent transport with no interactions in the channel. In this limit, the results could also have been obtained from the scattering theory of transport, sometimes referred to as the Landauer-Büttiker formalism. The NEGF method provides a convenient way to get the same results. However, the NEGF method also allows us to include channel interactions through the self-energy functions Σ_0 , Σ_0^{in} .

It is important to include these interactions because the purely coherent model predicts quantum interference effects like the oscillations in Fig. 64 which are usually not seen experimentally except at extremely low temperatures. The reason is that the phase of individual electrons is destroyed by the random scattering potential it sees due to the other electrons. Such electron-electron interactions normally do not lead to any overall loss of momentum from the system of electrons and so have little effect on the mobility or the current. But they do lead to dephasing which suppresses the coherent oscillations. Dephasing may also arise from other sources such as electron-phonon interactions [35]. However, they usually cause momentum relaxation as well, thereby affecting the resistance. We specifically mention electron-electron interaction because it can cause dephasing without having any first-order effect on the resistance. Resistance typically becomes temperature independent around 10 K, once phonons have frozen out. But electron-electron interactions continue to cause dephasing and it usually takes much lower temperatures to see interference effects.

The classic NEGF provides definite prescriptions for including all such effects from first principles starting from a microscopic Hamiltonian using MBPT as described in many available references. Here we will use a phenomenological approach that incorporates dephasing in terms

of a single parameter D_0 which can be adjusted to reflect experimental dephasing times:

$$\mathbf{\Sigma}_0 = D_0 \mathbf{G}^R \quad (8.12a)$$

$$\mathbf{\Sigma}_0^{\text{in}} = D_0 \mathbf{G}^n \quad (8.12b)$$

Note that we now need to solve Eqs. (8.1) and (8.2) self-consistently since the same quantity appears on both the left and right hand sides:

$$\mathbf{G}^R = [\mathbf{EI} - \mathbf{H} - \mathbf{\Sigma}_1 - \mathbf{\Sigma}_2 - D_0 \mathbf{G}^R]^{-1} \quad (8.13a)$$

$$\mathbf{G}^n = \mathbf{G}^R [\mathbf{\Sigma}_1^{\text{in}} + \mathbf{\Sigma}_2^{\text{in}} + D_0 \mathbf{G}^n] \mathbf{G}^A \quad (8.13b)$$

The self-consistent calculation is straightforward to implement numerically and a Python code is provided for readers interested in reproducing the results shown in Figs. 64 and 66. Note that the results for pure dephasing agree well with the simple semiclassical model based on series resistors (Eq. 8.10) as described earlier. Fig. 66 also shows results for a slightly different model of interactions which relaxes both phase and momentum leading to a linear drop in the occupation factor across the channel like an ohmic resistor. In this slightly different model, the diagonal elements of $\mathbf{\Sigma}_0$ and $\mathbf{\Sigma}_0^{\text{in}}$ are the same as before, but all off-diagonal elements are set to zero [36, 37].

Intuitively, we can justify the pure phase relaxation model by thinking of it as distributed “Büttiker probes” that take electrons out of the channel and reinject them back into it to break their phase coherence [38]. But because the electron density ($\mathbf{G}^n$) is preserved *exactly* by constant multiplication, no property other than phase is relaxed. In contrast, just keeping the diagonal elements of the $\mathbf{G}^n$ in the real space representation breaks momentum in addition to phase since the reinjected electrons have no definite momentum and can flow in any direction. This intuition allows an immediate extension to spin preserving but momentum relaxing $\mathbf{\Sigma}_0$ as follows: Supposing we express $\mathbf{H}$, $\mathbf{G}^n$ and lead $\mathbf{\Sigma}$ ’s in $2N \times 2N$ matrices for a lattice of N points, following the same principle we could write down a spin-conserving but momentum and phase relaxing $\mathbf{\Sigma}_0$ as:

$$\mathbf{\Sigma}_0 = \begin{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} 0 & \cdots & \cdots & 0 \\ 0 & \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \end{bmatrix},$$

where the 2×2 blocks in each lattice point preserve spin information but phase and momentum are relaxed. Indeed, using this principle we can write down $\mathbf{\Sigma}_0$ ’s that can preserve information within *any* unit cell with B basis functions by having $B \times B$ blocks of all 1’s that preserve local correlations in $\mathbf{G}^n$. In the next subsection, we show an example of spin conserving dephasing in the context of spin transport.

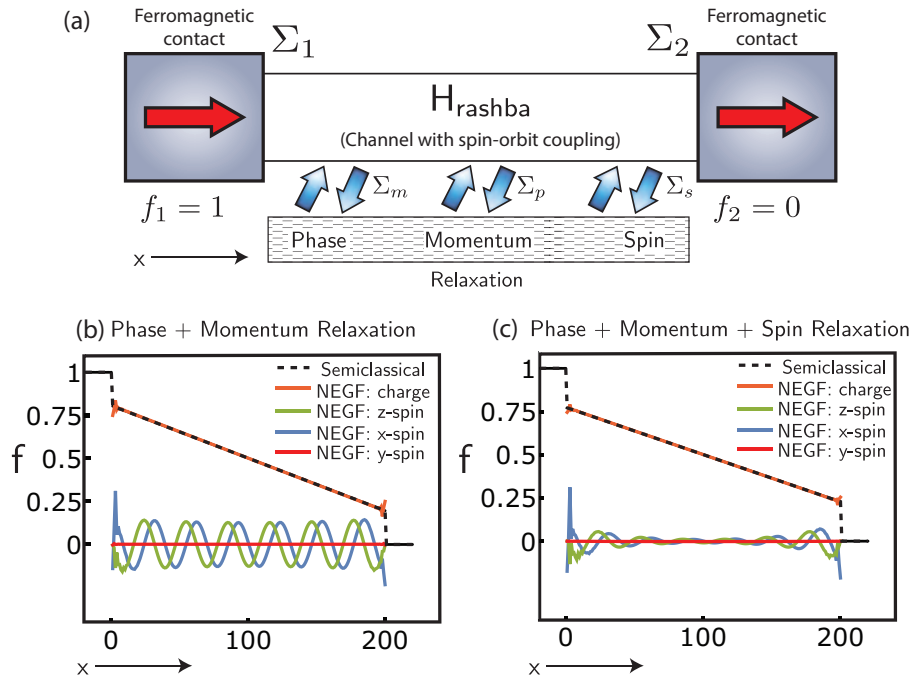


Figure 67: (a) Rashba channel with ferromagnetic leads. Phase, momentum and spin relaxation can be introduced to the system independently. (b) Spin precession is possible when $+x$ spins are injected to the channel even in the presence of phase and momentum relaxation. (c) Adding spin relaxation attenuates spin precession. (Python code available at the end of the article).

8.4.6 Diffusive spin transport using NEGF

Spin electronics or spintronics that combine spin transport with magnetics has attracted much attention in recent years [39]. Of particular interest are the so-called high spin orbit materials such as Rashba channels or interfaces [40, 41, 42, 43] as well as topological insulators [44] or heavy metals with high spin orbit coupling [45, 46]. NEGF allows a natural extension to incorporate spin transport in the tight-binding framework which makes the Hamiltonian for a lattice of N points a $2N \times 2N$ matrix. For example, consider a channel with a Rashba spin-orbit interaction that is described by the Hamiltonian,

$$\mathbf{H} = \mathbf{H}_0 + \eta (\sigma_x k_y - \sigma_y k_x) \quad (8.14)$$

where $\mathbf{H}_0$ represents the usual effective mass Hamiltonian, η is the Rashba coefficient, σ 's are Pauli spin matrices. We can think of the Rashba interaction as an effective magnetic field that depends on the momentum of an electron if we associate $|\eta \vec{k}|$ as the magnitude of the effective magnetic field. Indeed, injecting x directed spins by using ferromagnetic contacts with momentum ($\pm k_x$) can cause periodic oscillations of their spin, since these electrons would feel an effective magnetic field in the y direction. This effect was initially proposed in Ref. [47] and experimentally confirmed in Ref. [48], followed later by full demonstrations including the contacts [49, 50, 51].

Here, our purpose is to simply show how the NEGF formalism can be useful in modeling this effect in the presence of *momentum*, *phase* and *spin* relaxations in a toy model. Fig. 67a shows an NEGF simulation of 1D wire with Rashba spin orbit coupling, attached to two ferromagnetic leads that point in the x -direction. The left lead is an x -directed ferromagnet that injects x -spins in the channel. In the presence of momentum and phase relaxation, the chemical potential for charge goes steadily down as in Fig. 66. We can calculate the corresponding *spin* potentials

by:

$$f_{\ell}^{x,y,z} = \frac{\text{Trace} [\mathbf{G}_{\ell,\ell}^n \boldsymbol{\sigma}_{x,y,z}]}{\text{Trace} [\mathbf{A}_{\ell,\ell}]}$$

where ‘Trace’ indicates tracing the 2×2 blocks of $\mathbf{G}^n$ and $\mathbf{A}$ at lattice point ℓ and $\boldsymbol{\sigma}$ ’s are 2×2 Pauli spin matrices. Interestingly, if there is little to no spin relaxation, we see that the x -spins rotate in the channel even in the presence of high momentum and phase relaxation due to the effective magnetic field that is exerted on the x -spins by the Rashba channel. In practice, however, momentum relaxation in high spin-orbit materials is known to give rise to strong spin relaxation through the Elliott-Yafet [52] mechanism. Our purpose is simply to show that it is possible to include spin relaxation *independently* of momentum relaxation. This allows us to flexibly choose phase, momentum and spin relaxation rates to match experimentally measured quantities when modeling real experiments. As discussed in Ref. [37], a possible spin flip relaxation model, $\boldsymbol{\Sigma}_0^{sf}$ can be written as:

$$\boldsymbol{\Sigma}_{0,i,j}^{sf} = D_0(\sigma_x \mathbf{G}_{i,j}^n \sigma_x + \sigma_y \mathbf{G}_{i,j}^n \sigma_y + \sigma_z \mathbf{G}_{i,j}^n \sigma_z) \quad (8.15)$$

where the $\boldsymbol{\Sigma}_0^{sf}$ acts on all 2×2 sub-blocks of the $\mathbf{G}^n$ matrix that are labeled $\mathbf{G}_{i,j}^n$. By multiplying the terms out, we can easily see the effect of this dephasing is to reinject an electron *with an opposite spin* back to the channel that relaxes spin. In Fig. 67c we show how a pure spin relaxation model can be added in addition to momentum and phase breaking to reduce the spin signals in the channel.

8.4.7 Application to superconducting devices

An important application of the NEGF method is to devices where one of the contacts or the channel is superconducting. This is of great current interest due to the excitement surrounding the search for Majorana bound states (MBS) or equivalently Majorana zero modes (MZM) in Rashba nanowire systems in proximity with superconductors [53, 68, 15, 8, 57]. It would take us too far afield to discuss the significant conceptual extensions needed for these problems and so we will refer interested readers to a couple of old papers [58, 59] and some recent ones [60, 61, 62].

8.5 Frequently asked questions

8.5.1 Shouldn’t we also consider the Poisson equation for modeling real devices?

Yes. Earlier we mentioned the role of electron-electron interactions in causing dephasing which can be included through the anti-Hermitian part of the self-energy function $\boldsymbol{\Sigma}_0$. A far more important effect of electron-electron interactions is manifested through the Hermitian part of $\boldsymbol{\Sigma}_0$ which effectively modifies the Hamiltonian $\mathbf{H}$. The most important part of $\boldsymbol{\Sigma}_0$ is just the Coulomb potential U arising from the electron density $\mathbf{G}^n/2\pi$, and can be obtained by solving the Poisson equation and is often called the Hartree term. In addition there are more subtle contributions to $\boldsymbol{\Sigma}_0$ generally referred to as exchange and correlation effects which subtract from the Hartree term, see for example, Chapter 2 of Ref. [28].

Since $\boldsymbol{\Sigma}_0$ depends on $\mathbf{G}^n$, this requires a self-consistent calculation: We start by assuming $\boldsymbol{\Sigma}_0 = 0$, use Eq. 8.1,8.2 to calculate $\mathbf{G}^n$, use $\mathbf{G}^n$ to calculate $\boldsymbol{\Sigma}_0 = 0$, recalculate $\mathbf{G}^n$ from Eq. 8.1,8.2 and so on till the results converge. This requirement for an iterative solution makes it much more difficult to calculate the energy levels even for Helium which has two electrons compared to Hydrogen which has only one electron and hence no electron-electron interactions.

Two more points to note. Firstly, in device simulations it is common to use the Hamiltonian $\mathbf{H}$ from semi-empirical methods that already include the Σ_0 for a uniform channel material in equilibrium. As such it is necessary to include *only* the *change* in Σ_0 when the channel is driven out-of-equilibrium by the applied bias. Secondly, this correction is not important for low applied voltages, but plays a very significant role at high voltages. Indeed one can say that the saturation current for transistors is largely controlled by the Hartree term, making it primarily an electrostatically controlled device.

8.5.2 Can you arbitrarily designate the ends as contacts?

Designating a region as a contact implies that there is a significant increase in the number of conducting channels due to an increase in width and/or in the density of states. When analyzing a given structure, it is important to scrutinize the results carefully, to make sure that the assumed contacts are indeed functioning as nearly ideal reservoirs.

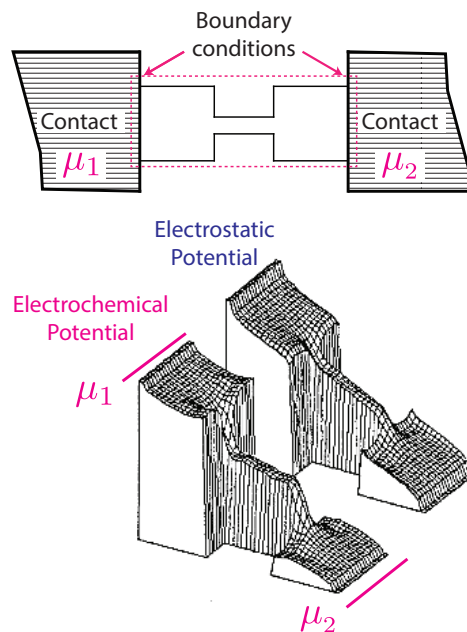


Figure 68: A wide-narrow-wide ballistic constriction where the contacts are imposed inside the widest region. A numerical simulation with phase scattering shows the 2D profile of electrochemical and electrostatic potentials. Reprinted Fig. 27b of Ref. [7] with permission from Michael J. McLennan, Yong Lee, and Supriyo Datta, Physical Review B, 43, 13846, 1991 Copyright (1991) by the American Physical Society.

For example, Fig. 68 is an excerpt from an old paper showing a wide-narrow-wide structure (see Fig. 27a of Ref. [7]). We impose the contact boundary conditions inside the widest region as shown. Note the sharp drops in the occupation factor (or electrochemical potential) at the narrow-wide interfaces due to the interface resistance. The electrostatic potential is obtained by convolving the electrochemical potential with a screening function as explained in the paper (Fig. 27b of Ref. [7]). Note also the small drops where the wide regions meet the contacts indicating a small interface resistance.

One could save computation time by imposing the contact in the wide region rather than in the widest region. This may or may not be acceptable depending on how conductive the wide region is relative to the narrow region. A related phenomenon called the “source-starvation” effect where contact-like wide regions are driven out of equilibrium has been discussed in the context of ballistic nanotransistors [63].

8.5.3 Does NEGF give the correct coherent and semiclassical limits?

Yes, with small dephasing, it matches the scattering theory for coherent transport, also called the Landauer-Büttiker formalism. With sufficient dephasing, it reproduces results from the semiclassical Boltzmann method, thus providing a bridge from quantum to semiclassical transport. The example presented earlier shows how the NEGF result for coherent transport includes oscillations around the Boltzmann result, which damp out when dephasing is included. Many such examples are included in Ref. [29] essentially as homework problems that are easily reproduced.

8.5.4 Questions about Kubo formalism and NEGF

How is this related to the Kubo formalism?

The Kubo formalism is a linear response theory based on small perturbations close to equilibrium. Many of the problems in modern nanoelectronics require a framework that can handle transport far from equilibrium. The results from NEGF agree with the Kubo formalism in the linear response regime close to equilibrium.

Isn't it a problem to have separate electrochemical potentials in a single system?

On the contrary, we would argue that current flow requires separate electrochemical potentials, just as the flow of heat requires separate temperatures. There is nothing fundamentally wrong with invoking two large reservoirs held at two different electrochemical potentials and/or temperatures connected by a nanoscale conductor.

Doesn't the Kubo formula use a single electrochemical potential?

Yes, the Kubo formalism calculates an equilibrium quantity, namely the noise, and relates it to the linear transport coefficients through the fluctuation-dissipation theorem. But this only works for transport close to equilibrium and is not possible in general. NEGF on the other hand addresses the non-equilibrium problem directly. Like the Boltzmann equation, NEGF too permits the use of separate electrochemical potentials as boundary conditions.

Can't you have equilibrium currents with a single electrochemical potential?

Yes, but these are distinct from the transport currents we are discussing. For example in the quantum Hall regime, there are circulating edge currents which involve all electrons and not just the ones with energies close to μ . But these are normally not measured with external probes.

8.5.5 Questions about many-body processes and NEGF

Isn't the flow of electricity essentially a many-body process?

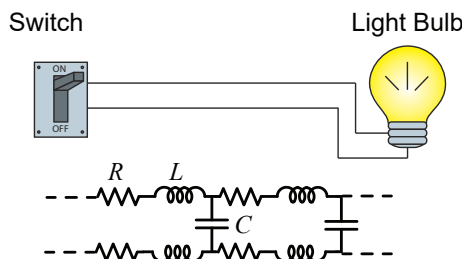


Figure 69: When we turn the switch on a signal propagates to the light bulb which can be described by a distributed RLC model. The capacitance C and the inductance L are determined primarily by electrostatics and magnetostatics respectively.

As undergraduates we learn that current flow, at least in metals, is a many-body interacting process: When we turn the switch on, a signal propagates to the light bulb at the speed of light. One electron pushes the next electron, which pushes the next electron and so on. If we had to wait for an electron to transmit from the switch to the bulb it would take far too long. And so it might seem futile to describe current flow without invoking many-body interactions. However, we believe that while signal propagation cannot be analyzed without interactions, the low frequency resistance can be.

This signal propagation can be described quantitatively by a distributed RLC model [64]¹¹. The L 's and C 's in this model have a transport component, but usually represent primarily the interactions and the velocity is given by $1/\sqrt{LC}$. However, these parameters are not involved in determining the steady-state or dc current since at low frequencies the L 's appear as short circuits, and the C 's are open circuits. The low frequency resistance R can be understood at least approximately without invoking interactions.

This observation appears to apply to quantum transport as well. For example, in 1985 Richard Webb and collaborators from IBM reported [66] the Aharonov-Bohm (AB) effect in metallic rings with a period described quite well by the one-electron theory without interactions, as if we were talking about AB paths in vacuum! The one-electron picture works so well possibly because there is a Landau-style (We mean Landau, not Landauer!) dressed quasi-particle that “actually” moves from left to right. We have some mental pictures of what such a quasi-electron looks like, but we are not aware of any serious discussions along these lines.

Are many-body effects irrelevant in transport experiments?

That is not at all what we mean. Our point is that many-body effects are fairly obvious at high frequencies starting with the well-known fact that signals travel at the speed of light and not at the speed of electrons. By contrast their role at low frequencies are more subtle, like Coulomb blockade and the Kondo effect. More recently experiments on graphene have revealed intriguing evidence of collective motion [67].

In a volume published on the occasion of the 50th anniversary of *Anderson localization*, Phil Anderson commented “... What might be of modern interest is the channel concept which is so important in localization theory. The transport properties at low frequencies can be reduced to a sum over one dimensional channels.”

Note the qualification *low frequencies*. Transmission of an electron from the left to the right contact is essential only for steady-state charge transport. High frequency currents can

¹¹For a spin generalization of similar ideas, see [65].

flow capacitively without end-to-end transfer of electrons. Also, experiments have revealed that spins can be transported across insulators by magnons without the actual transmission of electrons¹².

Would NEGF be suitable for describing such effects?

NEGF provides a clear prescription for including arbitrary interactions to any desired order in perturbation theory. But some effects may require us to go beyond MBPT: perhaps a non-perturbative approach, or a different perturbation parameter. Single-electron charging effects provide a good example [70].

8.5.6 How can we teach NEGF without advanced quantum statistical mechanics?

The NEGF eqs. (8.1-8.2) are essentially the same as Eqs. (75)-(77) of L. V. Keldysh, Sov. Phys. JETP, vol. 20, p. 1018 (1965) [3], which is one of the seminal papers on the NEGF method that obtained these equations using MBPT and established the “diagram technique for non-equilibrium processes” for calculating the self-energy functions Σ_0 and Σ_0^{in} . Much of the literature is based on the original MBPT-based approach and this makes it inaccessible to those unfamiliar with advanced quantum statistical mechanics.

We obtain the NEGF equations directly from a one-electron Schrödinger equation using relatively elementary arguments. These equations have been used to discuss many problems of great interest like quantized conductance, (integer) quantum Hall effect, Anderson localization, resonant tunneling and spin transport without a systematic treatment of many-body effects. But it goes beyond purely coherent transport allowing us to include phase-breaking interactions (both momentum-relaxing and momentum-conserving) within a self-consistent Born approximation.

The NEGF equations provide a unified framework for such problems, spanning a wide range of materials and phenomena all the way from molecular to ballistic to diffusive transport and has been widely adopted by the nanoelectronics community for device analysis and design.

8.5.7 But is this the real NEGF?

- The answer is NO, if we associate NEGF with the MBPT commonly used to obtain Σ_0 and Σ_0^{in} appearing in Eqs. (8.1-8.3)
- The answer is YES, if we associate NEGF with Eqs. (8.1-8.3) irrespective of how the Σ ’s are obtained.

Which answer we choose is clearly a matter of perspective, but the second viewpoint seems more in keeping with semiclassical transport theory, where the (steady-state) Boltzmann approach is identified with the equation:

$$\boldsymbol{\nu} \cdot \nabla f + \mathbf{F} \cdot \nabla_p f = S_{\text{op}} f$$

and NOT with the evaluation of the scattering operator S_{op} which is analogous to the Σ ’s in NEGF. While the concept of the collision integral involved in S_{op} was well-known, the quantitative details of its evaluation has evolved significantly since the days of Boltzmann. Similarly, totally new approaches that go beyond many-body perturbation theory (MBPT) for evaluating the Σ ’s have been and will be developed as we apply NEGF to newer problems.

¹²See for example Ref. [68] and [69] for possible practical applications of such pure spin conductors.

8.5.8 Contact-ing Schrödinger

In summary, we feel that the scope and utility of the NEGF equations Eq. (8.1-8.3) transcend the MBPT-based approach originally used to derive it. NEGF teaches us how to combine quantum dynamics with “contacts” much as Boltzmann taught us how to combine classical dynamics with “contacts”, using the word “contacts” in a broad figurative sense to denote all kinds of entropy-driven processes. There are many problems of fundamental and applied significance that can be simulated straightforwardly within the NEGF method without a first principles treatment of many-body effects which in any case would be perturbative in nature. This includes novel phenomena like the (integer) quantum Hall effect as well as new devices like magnetic tunnel junctions. As such, we feel that this elegant framework for “contact-ing” the Schrödinger equation should be of broad interest to anyone working on device physics or non-equilibrium statistical mechanics in general.

9 Tunneling current I: Formalism

A simple model Hamiltonian is proposed for a metal-insulator-metal tunneling junction, which permits the direct calculation of the tunneling current without introducing any effective Hamiltonian. The model rests on the use of localized functions; this procedure avoids the difficult matching problem at the boundary between the barrier and the electrodes. Moreover the use of Keldysh’s perturbation theory for nonequilibrium systems allows an explicit calculation of the current to all orders in the applied bias. It is found that the transfer coefficient appearing in the expression for the current is energy dependent. This model can be systematically extended to include many body effects.

(тут абстрактные формулы, без других знаний вообще не разобраться, как ими пользоваться и что конкретно они значат?)

9.0.1 1. Introduction

In the past few years, a great deal of experimental work has been performed on various types of tunneling junctions with more and more refined techniques. This now allows the observation of fine structures in the $I - V$ curves, such as phonon effects in mim (metal-insulator-metal) junctions (Klein and Léger 1968, Rowell et al. 1969), or vibration spectra of radicals located in the insulating barrier (Lambe and Jaklevic 1968). On the other hand, the theoretical approach to the problem of tunneling still remains founded essentially on the theory of Bardeen (1961), that is, on the transfer Hamiltonian model. It is well known that this model gives, in most cases, a very good description of the observed effects, but that it has not been possible yet to deduce it firmly from first principles. Moreover, as has been recently pointed out by several authors (Appelbaum and Brinkman 1969), this model is not sufficient to treat many body effects properly.

In most treatments of tunneling, one tries, starting from the initial Hamiltonian of the total MIM system, to transform it into an effective ‘tunneling Hamiltonian’, H_{eff} , of the form

$$H_{\text{eff}} = H_L + H_R + \mathbf{T} \quad (1)$$

where H_L and H_R respectively describe, as exactly as possible, the left and right isolated electrodes, while $\mathbf{T}$ is a transfer term which describes the probability for an electron to tunnel through the barrier. One of the problems related to this formulation is that $\mathbf{T}$ obviously corresponds, not to an ordinary potential, but to the $\mathbf{T}$ -matrix for the tunneling process, that is, it is in general energy dependent. The energy dependence is neglected in the transfer

Hamiltonian theory, which makes this theory unsuitable for describing many body effects; many body processes, for instance phonon emission and absorption in the barrier, must appear essentially as retardation effects on the transfer term of the effective Hamiltonian.

In this paper, starting from a very simple model Hamiltonian for a mim junction, we calculate directly the tunneling current, without having to make use of an intermediate effective Hamiltonian of the type given by equation (1). Our method takes advantage of the following two points:

(i) We describe the system, not in terms of quasi-free electron wave functions, but of functions localized on atomic sites (analogous to Wannier functions). The decomposition of the system into three parts (left electrode, barrier, right electrode) results immediately from this picture, without giving rise to the usual mathematical difficulties.

(ii) It is usually considered that a description of the system as a whole does not permit calculation of the current, since tunneling is a nonequilibrium process. This is one of the main reasons for using an effective Hamiltonian of type (1); H_L and H_R correspond to two disconnected systems, each of which is separately at equilibrium, the transfer term T induces transitions between these two systems, and one calculates the transition probability by means of first order time dependent perturbation theory. Actually, the current can be calculated directly, even at finite voltage, with the help of the theory of perturbations in nonequilibrium systems (Kjeldysh 1965, Craig 1968), which generalizes the diagram techniques well known for the equilibrium case. Moreover, within this framework it is also possible to take into account many body effects systematically, without any of the difficulties encountered in the usual tunneling formalism.

In § 2, we describe and discuss our model Hamiltonian and calculate the corresponding electron Green function at zero voltage. We then compare this result with what would be obtained from an effective Hamiltonian model.

In § 3, we derive the tunneling current for such a system, and we examine briefly the possible generalizations of our model.

9.1 2. Model Hamiltonian and Green function at zero voltage

9.1.1 2.1. The model Hamiltonian

We first consider, for simplicity, a one dimensional system, made of semi-infinite metallic electrodes M and M' , and of an insulating barrier I ; for the time being, we neglect all many body interactions. We assume that any one electron wave function may be described in terms of an orthonormal basis $\varphi_l^{(n)}$, the function $\varphi_l^{(n)}$ being localized in the vicinity of site l (n is a band index). In terms of the corresponding creation and annihilation operators C^+ and C , the Hamiltonian has the form

$$H = \sum_{\substack{l,l' \\ n,n'}} T_{ll'}^{nn'} C_l^{(n)+} C_{l'}^{(n')} \quad (2)$$

If the system were homogeneous, the $\varphi_l^{(n)}$ would be chosen as the ordinary Wannier functions. We assume that a similar set may be constructed for a non-homogeneous system, which retains the same properties, that is, each function $\varphi_l^{(n)}$ is essentially localized within the l th atomic cell, and there is one $\varphi_l^{(n)}$ corresponding to each level of the l th atom †.

One can think of these pseudo-Wannier functions as being built by the following process: one would first build localized functions by means for instance of a LCAO method, then orthogonalize these localized functions. The main effect of the orthogonalization condition is to mix the LCAO function with the corresponding functions centred on the nearest neighbours of l (or

more generally on a few neighbours, the amount of admixture of $\phi_{l+j}^{(n)}$ to $\phi_l^{(n)}$ decreasing exponentially with $|j|$). The contribution of the ϕ of high energy (for instance those which belong to the continuum and therefore are not localized) does not affect seriously the localized character of the orthogonalized function $\varphi_l^{(n)}$ corresponding to $\phi_l^{(n)}$; on the one hand they lie much higher in energy than the ϕ_i we are interested in, so the corresponding amount of admixture is very small; on the other hand the sum of the contributions of all these quasi-free functions gives rise to destructive interferences so that the total contribution is essentially localized within one (or at most a few) atomic distance. It is clear that the φ_i centred on sites which are not too close to the metal-insulator boundaries (typically, farther than a few atomic distances) are practically identical to the Wannier functions of the bulk medium. Close to the boundaries, the modification of these functions is related to the surface effects due to the metal-insulator contacts.

In the calculations, we neglect all interband processes, so we retain in equation (2) only $\dagger$ This amounts to saying that, in a homogeneous system there is, on a given site i , one Wannier function corresponding to each band.

the φ_l constructed from one given LCAO level: this corresponds, for the metals M and M', to a one band approximation. The insulator is also described in the equivalent of a one band model; we retain only the φ_l corresponding to the band which is closest in energy to the energy level of the tunneling electrons $\dagger$.

In a homogeneous material, the matrix elements of the Hamiltonian T_{lm} between two Wannier functions decrease rapidly with $|l - m|$. This remains true in the M - I - M' junction, so that we can safely neglect the T_{lm} which would couple a site of M with a site of M'. (This assumes that the barrier is at least 3 or 4 atomic layers thick, which is usually the case in the experiments.)

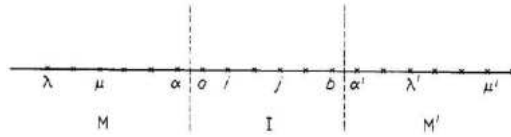


Figure 1. Schematic representation of a M-I-M' junction. The dashed lines represent the boundaries between the sites of the insulator and the two metals.

Moreover, we will assume for simplicity that the electrode M and the insulator are coupled by only one matrix element $T_{\alpha a}$, where α is the last site of M and a the first site in the barrier (see figure 1). In the same way, M' and I are coupled only through $T_{b\alpha'}$. All the calculations can be carried out with an arbitrary range for the M - I and I - M' couplings, they lead to a much more heavy algebra but to the same qualitative results as the simple nearest neighbour assumption.

The Hamiltonian can therefore be written as

$$H = \sum_{i\mu} T_{i\mu} C_i^+ C_\mu + \sum_{\lambda\mu'} T_{\lambda\mu'} C_{\lambda'}^+ C_{\mu'} + \sum_{ij} T_{ij} C_i^+ C_j + (T_{\alpha\mu} C_\alpha^+ C_\mu + \text{hc}) + (T_{\alpha'b} C_{\alpha'}^+ C_b + \text{hc}) \quad (3)$$

The unprimed greek indices describe the atomic sites of M (M'), the italic indices label the sites in the barrier.

The first term of equation (3) may be viewed as a 'left electrode Hamiltonian' (near the barrier, $T_{\lambda\mu}$ depends on both λ and μ , partly because of ordinary surface effects, but also because the Wannier functions are distorted by the presence of the insulator). The other terms of equation (3) respectively correspond to the right electrode, to the barrier and to the transfers between them. The separation of these various terms is here very natural; in order to establish the connection of the present approach with the usual description based on continuous wave functions throughout the system, one would need to go to the limit where the intersite distance

goes to zero. At the moment, we have not been able to overcome mathematical difficulties due to the matching of wave functions at the boundaries.

Limits of applicability (????)

Examples, where this approach works (????)

Examples, where this approach doesn't work (????)

9.1.2 2.2. The Green functions at zero bias

We now calculate the equilibrium Green function of the system described by the Hamiltonian of equation (3). We introduce the following notations

$$T_{xa} = T \quad T_{a'b} = T' \quad (4)$$

Let $g_{\lambda\mu}(\epsilon)$ be the Green function of the left electrode alone (corresponding to the Hamiltonian $\Sigma T_{\lambda\mu} C_{\lambda}^+ C_{\mu}$). The propagator $g_{\lambda\mu}$ is given by all the diagrams that go from site λ to site μ without ever entering the barrier. Similarly, we denote by $g_{i'\mu'}(\epsilon)$ and $g_{ij}(\epsilon)$ the propagators for the right electrode and for the barrier alone. Since the barrier is insulating, g_{ij} decreases exponentially with $|i - j|$.

The full propagator $G_{pq}(\epsilon)$ (where the site indices p and q run over the whole system) is the solution of

$$\epsilon G_{pq}(\epsilon) - \sum_{p_1} T_{pp_1} G_{p_1q}(\epsilon) = \delta_{pq} \quad (5)$$

In order to calculate the full propagator $G_{\lambda\mu}(\epsilon)$, we must take into account possible excursions of the electron across the (M – I) interface. If we single out the first and last paths followed in the M region, we may write

$$G_{\lambda\mu}(\epsilon) = g_{\lambda\mu}(\epsilon) + g_{\lambda\alpha}(\epsilon) T G_{aa}(\epsilon) T g_{x\mu}(\epsilon)$$

Similarly, we obtain for the (M' – M') and (M – M') propagators

$$\begin{aligned} G_{\lambda'\mu'}(\epsilon) &= g_{\lambda'\mu'}(\epsilon) + g_{\lambda'\alpha'}(\epsilon) T' G_{bb}(\epsilon) T' g_{\alpha'\mu'}(\epsilon) \\ G_{\lambda\mu'}(\epsilon) &= g_{\lambda\alpha}(\epsilon) T G_{ab}(\epsilon) T' g_{\alpha'\mu'}(\epsilon) \end{aligned} \quad (6)$$

We must now calculate the full propagator $G_{ij}(\epsilon)$ inside the barrier. For that purpose, we note that an excursion of the electron into **M** or **M'** may be viewed as a self-energy process superimposed on the barrier propagation. The corresponding self-energy operator has only two nonzero matrix elements

$$\begin{aligned} \Sigma_{aa} &= \Sigma = T^2 g_{\alpha x} \\ \Sigma_{bb} &= \Sigma' = T'^2 g_{\alpha'\alpha'} \end{aligned} \quad (7)$$

¹²† One can easily generalize the calculations to the case of a two band model for the insulating barrier. This will be discussed in § 3 .

‡ This assumption is not crucial. The exponentially small direct coupling between M and M' gives rise simply to a correspondingly small correction to the transmission coefficient of the barrier.

and the Dyson equation for G_{ij} reads

$$G_{ij}(\epsilon) = g_{ij}(\epsilon) + g_{ia}(\epsilon)\Sigma(\epsilon)G_{aj}(\epsilon) + g_{ib}(\epsilon)\Sigma'(\epsilon)G_{bj}(\epsilon) \quad (8)$$

The solution of equation (8) is straightforward,

(!!! here are computations!!!!)

and yields for instance

$$\begin{aligned} G_{ab} &= \frac{g_{ab}}{D} \\ G_{aa} &= \frac{g_{aa}(1 - T'^2 g_{bb} g_{\alpha'\alpha'}) + T'^2 g_{ba} g_{ab} g_{\alpha'x'}}{D} \\ G_{bb} &= \frac{g_{bb}(1 - T^2 g_{aa} g_{\alpha\alpha}) + T^2 g_{ba} g_{ab} g_{\alpha x}}{D} \end{aligned} \quad (9)$$

where the denominator D is given by

$$D = (1 - T^2 g_{aa} g_{\alpha\alpha}) (1 - T'^2 g_{bb} g_{\alpha'\alpha'}) - T^2 T'^2 g_{ab} g_{ba} g_{\alpha\alpha} g_{\alpha'\alpha'} \quad (10)$$

the exact Green function G is thus completely determined.

We note that the M – M' propagator may be written as

$$G_{\lambda\lambda'} = \sum_{\mu\mu'} g_{i\mu\mu'} \mathcal{T}_{\mu\mu'} g_{\mu'\lambda'} \quad (11)$$

where $\mathcal{T}$ is an effective T-matrix given by

$$\mathcal{T}_{\mu\mu'} = \frac{TT'G_{ab}}{D} \delta_{\mu\alpha'} \delta_{\mu'\alpha'} \quad (12)$$

$\mathcal{T}_{\mu\mu'}$ is equivalent to the transfer matrix element introduced by Bardeen.

Note that, as we have already noticed, and as was to be expected, since the barrier is insulating, g_{ab} , and therefore $\mathcal{T}$, decrease exponentially when $|a - b|$, that is, the barrier thickness, increases.

Besides the fact that we have an explicit expression for the T-matrix, our result differs from the usual one in several ways:

(i) $\mathcal{T}$ depends on the energy ϵ , a point which has been recognized to be essential in the proper treatment of inelastic tunneling. Physically, this simply means that the electron spends a finite time inside the barrier.

(ii) $\mathcal{T}$ depends on the electrode propagators $g_{\alpha x}$ and $g_{x'x'}$ essentially because an electron inside the barrier can make excursions into the electrodes.

(iii) A simple minded calculation of the full G in terms of $\mathcal{T}$ usually gives erroneous results. For instance, the usual transfer Hamiltonian theory would yield for the M-M propagator

$$G_{\lambda\mu} = g_{\lambda\mu} + g_{\lambda\alpha} \mathcal{T} g_{\alpha'\alpha'} \mathcal{T} g_{\alpha\mu} = g_{\lambda\mu} + g_{\lambda\alpha} \frac{T^2 T'^2 g_{ab} g_{ba} g_{\alpha'\alpha'}}{D^2} g_{\alpha\mu} \quad (13)$$

which is obviously different from the exact expression (4). The error stems from the fact that the standard method ignores processes in which the propagator turns backward to M inside the barrier, before it has reached the I-M' interface.

9.1.3 Other basic questions about the model

Why couldn't we do easier?

(?????????)

9.2 3. The tunneling current

9.2.1 3.1. The current operator

Let P be a point between site i and $i + 1$; the current J at point P is the difference between the flow of electrons from left to right and from right to left. We thus expect an operator J of the form

$$J_P = \sum_{\substack{l \geq i+1 \\ m \leq i}} A_{lm} C_l^+ C_m - \sum_{\substack{l \leq i \\ m \geq i+1}} A_{lm} C_l^+ C_m \quad (14)$$

In order to calculate A_{lm} , we write the continuity equation, which in our discrete representation reads

$$J_P - J_{P'} + \frac{\partial \rho_i}{\partial t} = 0 \quad (15)$$

P' is a point between sites $(i - 1)$ and i , while $\rho_i = e C_i^+ C_i$ is the electron charge at site i (or, more exactly, around the site i , over a region depending on the localization of the pseudo-Wannier functions). Using the Hamiltonian of equation (2), we find

$$\frac{\partial \rho_i}{\partial t} = \frac{1}{i\hbar} [\rho_i, H] = \frac{e}{i\hbar} \sum_m (T_{im} C_i^+ C_m - T_{mi} C_m^+ C_i) \quad (16)$$

A comparison of equations (14) and (16) yields

$$A_{lm} = \frac{e}{i\hbar} T_{lm} \quad (17)$$

In a stationary state, the current may be calculated at any point in the system. Since in our simple model the barrier is coupled to M (M') through a single matrix element T_{aa} ($T_{\alpha'b}$), it is particularly convenient to compute the current between α and a (or between b and α'). From now on, we shall choose P between sites α and a . In that case, equations (14) and (17) lead to

$$J = \langle J_P \rangle = \frac{eT}{i\hbar} (\langle C_\alpha^+ C_a \rangle - \langle C_a^+ C_\alpha \rangle) \quad (18)$$

The calculation of J is reduced to that of the Green function G_{aa} for equal times.

9.2.2 3.2. Calculation of the current

We now assume that the barrier is biased at a voltage V ; the chemical potentials μ and μ' in both electrodes differ by a quantity

$$\mu' - \mu = eV \quad (19)$$

If the electrodes are of a different nature, there may exist a contact potential, so that the difference in the electrostatic potentials ($\varphi' - \varphi$) is not necessarily equal to V . We shall assume that such contact potentials have been eliminated by appropriate shifts of the band structure; the applied bias may then be described by a Hamiltonian

$$\sum_i V_i C_i^+ C_i \quad (20)$$

where V_i is the 'effective' potential at site i , going from 0 on one side to V on the other.

Ordinary perturbation theory may be used to obtain the linear response of the junction to the bias V ; the term (20) is then treated as a small perturbation, and one only needs to calculate the appropriate correlation function for a junction in thermodynamic equilibrium (that is, such that $\mu = \mu'$). Unfortunately, such an approach does not work if we want to calculate the structure of the $(I - V)$ characteristics; the system is then far from equilibrium ($\mu \neq \mu'$), and the usual methods of perturbation theory do not apply. We shall use instead a generalized perturbation method due to Kjeldysh (1965) which remains valid for off equilibrium situations (5). This technique is briefly described in the Appendix.

We proceed along the same lines as in subsection 2.2, that is, we treat the I-M and I-M' couplings as perturbations; to zeroth order, the three regions I, M and M' are completely disconnected, and no current can flow between them. The applied bias is included in the zeroth order Hamiltonian; the voltage V_i acts to modify the one electron energies, and we assume that the latter have been redefined accordingly; the only effect of the bias is then to maintain a difference in chemical potentials between the two sides. We note that to zeroth order, the existence of such a difference ($\mu' - \mu$) does not preclude the existence of thermodynamic equilibrium (no current can flow across the barrier). We can thus construct our perturbation theory on a well defined initial state. To be more specific, the potentials are constant in M and M' (although different); there is no electric field, and the metals are in equilibrium. Inside the barrier, there is an electric field, but usually the Fermi level lies somewhere in a forbidden band; no current can flow, and again the insulator responds reversibly to the existing electric field. (Such a conclusion would not apply in the FowlerNordheim regime, when the tunneling electron can enter the conduction band of the insulator; our theory as such does not apply to this case.)

The initial state being thus unambiguously defined, we establish adiabatically the I-M and I - M' couplings, and we calculate the current through the barrier. The Kjeldysh method allows us to perform the calculation to all orders in the coupling, thus describing a basically nonequilibrium (although stationary) situation.

One might raise a major objection to the above procedure; it amounts to establishing first the de bias, and only later the coupling between the barrier and the electrode. Physically, it is the reverse that is true; the transfer matrix elements are always there, and the dc bias is established afterwards; it is not obvious that the corresponding limits can be interchanged. We believe that such a difficulty does not appear in the experimental situation in which the two electrodes are connected to a generator which controls the bias and serves as a source (or sink) of carriers when current flows. The relation between bias $V(t)$, current $J(t)$ and conductance $G(t)$ is instantaneous

$$J(t) = V(t)G(t)$$

and it does not matter whether the bias or the conductance is established first. The situation would be different if the junction were isolated: the flow of current would then change the capacitance of the junction, and the bias at time t would depend on the past history of V and $G^\dagger$.

We now turn to the detailed application of Kjeldysh's method to our problem. We introduce three Green functions

$$\begin{aligned}
G_{pq}^a(t, t') &= -i\theta(t - t') \langle [C_p(t), C_q^+(t')]_+ \rangle \\
G_{pq}^r(t, t') &= i\theta(t' - t) \langle [C_p(t), C_q^+(t')]_+ \rangle \\
F_{pq}(t, t') &= -i \langle [C_p(t), C_q^+(t')]_- \rangle
\end{aligned} \tag{21}$$

The site indices p and q run over the whole system, $[A, B]_+ = AB + BA$; $[A, B]_- = AB - BA$ and $\theta(t)$ is the ordinary step function.

The current J given by equation (18) can be expressed as

$$J = \frac{eT}{\hbar} \{G_{\alpha a}^c(t, t_+) - G_{u\alpha}^c(t, t_+)\} \tag{22}$$

where G^c is the usual causal Green function

$$G_{pq}^c(t, t') = -i \langle T \{C_p(t) C_q^+(t')\} \rangle \equiv \frac{1}{2} \{G_{pq}^a(t, t') + g_{pq}^r(t, t') + f_{pq}(t, t')\}$$

We also define the corresponding Green functions $g_{pq}^a(t, t')$, $g_{pq}^r(t, t')$ and $f_{pq}(t, t')$ for the zero order system (no coupling between barrier and electrodes). We note that g^r, g^a and f vanish when p and q belong to two different subparts of the junctions (for example $g_{\lambda i}^{(a)} = 0$).

The time Fourier transforms of G^a, G^r and F obey the following Dyson's equations [‡] (see Appendix):

$$\begin{aligned}
G_{pq}^a(\omega) &= g_{pq_1}^a(\omega) \{ \delta_{q_1 q} + \Sigma_{q_1 q_2} G_{q_2 q}^a(\omega) \} \\
G_{pq}^r(\omega) &= g_{pq_1}^r(\omega) \{ \delta_{q_1 q} + \Sigma_{q_1 q_2}^r G_{q_2 q}^r(\omega) \} \\
F_{pq}(\omega) &= f_{pq}(\omega) + g_{pq_1}^r(\omega) \Sigma_{q_1 q_2} F_{q_2 q}(\omega) + f_{pq_1}(\omega) \Sigma_{q_1 q_2} G_{q_2 q}^a(\omega)
\end{aligned} \tag{24}$$

where the self-energy Σ_{pq} is defined as

$$\Sigma_{pq} = T(\delta_{px'} \delta_{aq} + \delta_{pa} \delta_{\alpha q}) + T'(\delta_{pb} \delta_{\alpha' q} + \delta_{px'} \delta_{bq}) \tag{25}$$

Let us first consider the advanced and retarded functions G^r and G^a ; the corresponding Dyson's equations are identical to those introduced in §2 in our description of the equilibrium system-except that an index r (or a) has been affixed to all propagators. The previous algebra remains valid; for instance, when calculating $G_{\alpha a}^a$, we may start from site α , and single out the first 'crossing' vertex at which the propagator penetrates into the barrier, in $G_{\alpha a}^a$, we pick out the last crossing vertex before the propagator reaches site α . We thus obtain

$$\begin{aligned}
G_{\alpha a}^a &= g_{\alpha\alpha}^a T G_{aa}^a \\
G_{a\alpha}^a &= G_{aa}^a T g_{\alpha\alpha}^a
\end{aligned} \tag{26}$$

^{12‡} In such a case, the correct limit is obtained by letting first the volume of the electrodes go to infinity (which makes the capacitance of the junction infinite), the adiabatic switching time going to infinity only later.

[‡] In the general nonequilibrium situation, all the Green functions depend separately on the two time variables t and t' . However, since we are interested here in a stationary situation, they depend only on $(t - t')$.

Similar equations hold for G^r . We see that G^r and G^a do not contribute to the expression (22) for the current, which reduces to

$$J = \frac{eT}{2\hbar} \{F_{xa}(t, t_+) - F_{ax}(t, t_+)\} \quad (27)$$

We are then left with the calculation of F_{xa} and F_{ax} .

The chain of Dyson's equations for F_{xa} is, according to equation (24) †

$$\begin{aligned} F_{\alpha a} &= f_{\alpha a} T G_{aa}^a + g_{\alpha a}^r T F_{aa} \\ F_{aa} &= f_{aa} + f_{aa} T G_{aa}^a + f_{ab} T' G_{\alpha' a}^a + g_{aa}^r T F_{\alpha a} + g_{ab}^r T' F_{\alpha' a} \\ F_{\alpha' a} &= f_{\alpha' a} T' G_{ba}^a + g_{\alpha' a}^a T' F_{ba} \\ F_{ba} &= f_{ba} + f_{bb} T' G_{\alpha' a}^a + f_{ba} T G_{aa}^a + g_{bb}^r T' F_{\alpha' a} + g_{ba}^r T F_{\alpha a}. \end{aligned} \quad (28)$$

(The various matrix elements of G^r and G^a are obtained from equation (9).) We now assume that, in the range of energy of interest ‡, the density of states in the isolated barrier is strictly zero. This means that we restrict ourselves to electrons tunneling in the true gap of the insulator (for instance, we assume that there is no impurity bound state in this domain). Since the density of states is related to the imaginary part of the Green functions, in this energy range the g functions of the barrier are real, so that

$$\begin{aligned} g_{ij}^a &= g_{ij}^r = g_{ij}^c = g_{ij} \\ f_{ij} &= 0 \end{aligned} \quad (29)$$

The resolution of equation (28) is then straightforward,
(!!! here are computations!!!!)
and yields

$$\begin{aligned} F_{xa} &= T T'^2 f_{x' \alpha'} G_{ba}^a G_{ab}^r g_{\alpha \alpha}^r + \frac{T f_{\alpha \alpha}}{D^a D^r} (1 - T'^2 g_{\alpha' \alpha'}^a g_{bb}) (1 - T'^2 g_{bb} g_{\alpha' \alpha'}^r) g_{aa} + \\ &\quad + T T'^2 G_{ab}^a G_{ba}^r f_{\alpha \alpha} g_{\alpha' \alpha'}^a (1 - T'^2 g_{x' \alpha'}^r g_{bb}) \end{aligned} \quad (30)$$

In the same way, one obtains F_{ax} , and finally

$$\begin{aligned} F_{\alpha a}(\omega) - F_{a \alpha}(\omega) &= T T'^2 G_{ba}^a(\omega) G_{ab}^r(\omega) \left\{ (g_{\alpha x}^r(\omega) - g_{\alpha x}^a(\omega)) f_{\alpha' \alpha'}(\omega) \right. \\ &\quad \left. + f_{\alpha \alpha}(\omega) (g_{\alpha' \alpha'}^a(\omega) - g_{\alpha' x'}(\omega)) \right\} \end{aligned} \quad (31)$$

where $G_{ba}^{a(r)}(\omega) = g_{ba}(\omega)/D^{a(r)}$ is the Green function which describes the propagation of an electron from one side of the barrier to the other in the real system; the renormalizing factor $1/D^{a(r)}$ expresses the probability of an electron path, going from a to b , penetrating into the electrodes.

At this stage, it is convenient to introduce two extra Green functions

$$\begin{aligned} G_{pq}^+(t, t') &= i \langle C_q^+(t') C_p(t) \rangle \\ G_{pq}^-(t, t') &= -i \langle C_p(t) C_q^+(t') \rangle \end{aligned} \quad (32)$$

with similar definitions for the zeroth order limits $g^\pm$. We verify easily that

$$\begin{aligned} g^+ + g^- &= f \\ g^+ - g^- &= g^r - g^a \end{aligned} \quad (33)$$

The expression (31) may then be recast in the form

$$F_{\alpha a}(\omega) - F_{a x}(\omega) = 2T T'^2 G_{ba}^a(\omega) G_{ab}^r(\omega) \{g_{\alpha\alpha}^+(\omega) g_{\alpha'\alpha'}^-(\omega) - g_{\alpha a}^-(\omega) g_{\alpha'x'}^+(\omega)\} \quad (34)$$

Together with equation (27), equation (34) determines the current J .

By performing a spectral analysis of $g_{xx}^\pm$ and $g_{\alpha x}^\pm$, we obtain, in the limit of zero temperature

$$\begin{aligned} g_{zx}^+(\omega) &= 2i\theta(\mu - \omega) \text{Im } g_{\alpha x}(\omega) = 2i\pi\theta(\mu - \omega)\rho_\alpha(\omega) \\ g_{\alpha x}^-(\omega) &= -2i\pi\theta(\omega - \mu)\rho_\alpha(\omega) \end{aligned} \quad (35)$$

where $\rho_\alpha(\omega)$ is the density of states on site α in the isolated electrode. A similar result holds for $g_{x'\alpha'}^\pm$, μ being replaced by μ' .

Inserting equations (34) and (35) into the expression (27) for the current, one obtains [†]

$$J = -\frac{(2\pi)^2 e T^2 T'^2}{\hbar} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} G_{ba}^a(\omega) G_{ab}^r(\omega) \rho_\alpha(\omega) \rho_{\alpha'}(\omega) \{\theta(\omega - \mu)\theta(\mu' - \omega) - \theta(\mu - \omega)\theta(\omega - \mu')\} \quad (36)$$

In the case of figure 2 where $\mu' > \mu$, equation (36) reduces to

$$J = -\frac{(2\pi)^2 e T^2 T'^2}{\hbar} \int_{\mu}^{\mu+eV} \frac{d\omega}{2\pi} G_{ba}^a(\omega) G_{ab}^r(\omega) \rho_\alpha(\omega) \rho_{\alpha'}(\omega) \quad (37)$$

9.2.3 3.3. Discussion

Equation (37) shows that the current is, as usual, built up from the contribution of the states whose energies range from μ to $\mu' = \mu + eV$, that is, which are occupied in electrode M' and empty in electrode M. The main difference between our result and the expression obtained from the conventional tunneling Hamiltonian lies in the form of the 'transfer coefficient', which in our model is equal to

$$|\mathcal{T}|^2 = T^2 T'^2 G_{ba}^a(\omega) G_{ab}^r(\omega) \quad (38)$$

We note that $|\mathcal{T}|^2$ is frequency dependent. In the simple case that we have considered, this dependence is very weak as long as eV remains much smaller than the barrier energy. On the other hand, this fact becomes essential if we are to include self-energy effects, such as electron-phonon interactions. First of all, $|\mathcal{T}|^2$ depends mostly on the energy, and not on the momentum, a point which has been stressed by various authors. Moreover, $|\mathcal{T}|^2$ depends on

^{12†} According to Kjeldysh, the terms linear in f in equation (28) should vanish. This conclusion does not hold here. Mathematically, such a statement amounts to claiming that the product $f(1 + \Sigma^a G^a) = f g^{a-1} G^a$ vanishes because $f(g^a)^{-1} = 0$. In fact in our case the factor G^a is singular, and the product has a finite value, as can be seen by direct inspection. We suspect that the difference with Kjeldysh stems from the fact that our model does not involve any relaxation mechanism in the electrodes.

[‡] We will check on the final result that, as can be expected, the energies of interest range between the Fermi levels of the two electrodes.

the electrode propagators g_{xx} and $g_{\alpha'\alpha'}$, through the renormalizing denominators in equation (9); this is a new feature of the present calculations, which may affect profoundly the shape of the self-energy structures in the $I - V$ characteristics.

It should be stressed that some of the simplifying features of the above model can be easily relaxed, namely:

(a) one can extend the calculation to three dimensions, in the simple case where the component of the wave vector parallel to the M-I and I-M' boundary planes $\mathbf{k}_{\parallel}$ is conserved. This means that we neglect the Moiré patterns which would arise from a difference between

the periodicities in M, I and M' in the yz plane (parallel to the boundaries). Moreover, we assume that each electrode is coupled to the barrier through a Hamiltonian of the form

$$H_{IM} = \sum_{aa} T_{aa} C_{\alpha}^{\dagger} C_a + \text{hc} \quad (39)$$

where α lies in the first atomic plane of the metal, and a in the first atomic plane of the insulator (the range of T_{aa} in the yz plane is unimportant). With this assumption, the x

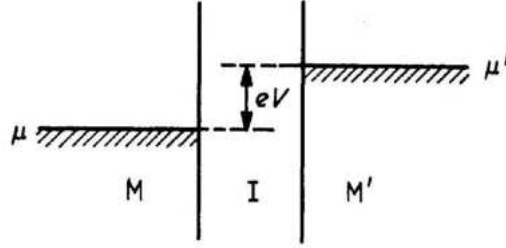


Figure 2. Schematic representation of the Fermi levels of the two electrodes in the presence of an applied voltage V .

component of the current is given by the same formal expression as in the one dimensional case

$$J = \frac{ie}{\hbar} \sum_{\alpha a} (T_{\alpha a} C_{\alpha}^{\dagger} C_a - T_{a\alpha} C_a C_{\alpha}) \quad (40)$$

In order to take advantage of the translational invariance along the yz plane, we Fourier transform all the quantities with respect to y and z ; we define for instance

$$T(\mathbf{k}_{\parallel}) = \sum_{\alpha a} T_{\alpha a} \exp \{i\mathbf{k}_{\parallel} \cdot (\mathbf{R}_{\alpha} - \mathbf{R}_a)\} \quad (41)$$

$$G_{pq}(\mathbf{k}_{\parallel}, \omega) = \sum_{\mathbf{p}_1, \mathbf{q}_1} G_{\mathbf{R}_p, \mathbf{R}_q}(\omega) \exp \{i\mathbf{k}_{\parallel} \cdot (\mathbf{p}_{\parallel} - \mathbf{q}_{\parallel})\}$$

(where p and $\mathbf{p}_{\parallel}$ denote respectively the x and (y, z) components of $\mathbf{R}_p$). For each value of $\mathbf{k}_{\parallel}$ we are left with a one dimensional problem, and hence the current is given by

$$J = \int \frac{d^2 \mathbf{k}_{\parallel}}{(2\pi)^2} \frac{(2\pi)eT^2(\mathbf{k}_{\parallel}) T'^2(\mathbf{k}_{\parallel})}{\hbar} \int_{\mu}^{\mu+eV} \frac{d\omega}{2\pi} G_{ba}^a(\mathbf{k}_{\parallel}, \omega) G_{ab}^r(\mathbf{k}_{\parallel}, \omega) \rho_{\alpha}(\mathbf{k}_{\parallel}, \omega) \rho_{\alpha'}(\mathbf{k}_{\parallel}, \omega) \quad (42)$$

where $\rho_{\alpha}(\mathbf{k}_{\parallel}, \omega)$ is the density of states in the isolated electrode on the last atomic plane and for a given value of the parallel momentum

^{12†} Note that G_{ba} must include the effect of the finite electric field inside the barrier, which presumably accounts for the so called 'background current'. Such an effect is easily calculated if the bias is much smaller than the barrier height.

$$\rho_\alpha(\mathbf{k}_\parallel, \omega) = -\frac{1}{\pi} \text{sgn}(\omega - \mu) \text{Im } g_{\alpha\alpha}^c(\mathbf{k}_\parallel, \omega) \quad (43)$$

(b) We have assumed that a one band model is sufficient to describe the insulating barrier. Our calculations can be generalized without difficulty to the case of a two band model. This could be of interest if the energy range of interest lies close to the middle of the gap separating the valence and conduction band of the insulator. In that case, the transfer coefficient appearing in J could be somewhat lowered, due to interference effects between transfer processes which take place via the two different bands.

(c) The most interesting extension of the present model consists in including many body effects, such as electron-phonon interactions in the electrodes, or interaction with localized excitations or resonant states inside the barrier. This can be done systematically with the help of Keldysh's generalized diagram technique. It is then possible to describe in detail the influence of these various effects on the transfer coefficient $|\mathcal{T}|^2$, as well as on the density of states of the electrodes.

These various extensions are presently under study, and will be published separately. Here we shall only stress the main features of our model. It is based on the following two ingredients:

- (i) a description in terms of localized 'Wannier' functions, which avoids the difficult matching problem at the boundary between the barrier and the electrodes,
- (ii) the use of Keldysh's perturbation theory for nonequilibrium processes, which allows an explicit calculation of the propagators and of the current, to all orders in the applied bias and in the electrode-barrier couplings.

The assumption of nearest neighbour coupling is not essential. Quite generally, the problem reduces to the calculation of the propagator G_{ij} inside the barrier: if the latter contains p layers, we are faced with a $p \times p$ secular equation (which becomes $2p \times 2p$ if we include two energy bands in the insulator). Possible excursions of the electron inside the electrode only appear as self-energy corrections to this barrier secular equation. The expression for the current resembles the usual result, but with a transfer coefficient $|\mathcal{T}|^2$ which we calculate explicitly, and which depends on the energy, and also on the properties of the electrodes.

9.2.4 Appendix

The usual diagram technique involves the calculation of chronological products of the form[†]

$$\langle \phi_0 | T \{ A(t) B(t') \dots \} | \phi_0 \rangle$$

(at zero temperature). $A(t), B(t')$ are operators in the Heisenberg representation. In the interaction representation, this expression becomes

$$\langle \phi_0 | S(-\infty, +\infty) T \{ \tilde{A}(t) \tilde{B}(t') \dots S(+\infty, -\infty) \} | \phi_0 \rangle$$

where the matrix $\mathbf{S}$ is defined by

$$S(t', t) = T \left[\exp \left\{ -i \int_t^{t'} H_{\text{int}}(t_1) dt_1 \right\} \right]$$

The interaction Hamiltonian H_{int} is supposed to be adiabatically turned on from $t = -\infty$, and turned off at $t = +\infty$. The operators in the interaction representation are $\tilde{A}(t), \tilde{B}(t'), H_{\text{int}}(t_1)$.

In the absence of irreversible effects

$$S(+\infty, -\infty)\phi_0 = e^{ix}\phi_0$$

so that

$$\langle \phi_0 | T \{ A(t) B(t') \dots \} | \phi_0 \rangle = \frac{\langle \phi_0 | T \{ \tilde{A}(t) \tilde{B}(t') \dots S(+\infty, -\infty) \} | \phi_0 \rangle}{\langle \phi_0 | S(+\infty, -\infty) | \phi_0 \rangle}$$

This statement is not true if the evolution of the system is irreversible, since the system either absorbs or emits energy during this evolution.

But it is possible to generalize the diagram technique. Indeed

$$\langle \phi_0 | S(-\infty, +\infty) T \{ \tilde{A}(t) \tilde{B}(t') \dots S(+\infty, -\infty) \} | \phi_0 \rangle = \langle \phi_0 | T_c \{ \tilde{A}(t) \tilde{B}(t') \dots S_c \} | \phi_0 \rangle \quad (\text{A.1})$$

where T_c is an operator ordering the times from right to left, not as usual from $-\infty$ to $+\infty$, but along a contour C made of a 'positive' branch going from $-\infty$ to $+\infty$ and with a 'negative' branch going from $+\infty$ to $-\infty$. Thus, on the positive branch, the times are ordered from $-\infty$ to $+\infty$, while they are ordered from $+\infty$ to $-\infty$ on the negative branch. Any time of the negative branch is considered as posterior to any time of the positive branch. The times $t, t' \dots$ of the chronological product are on the positive branch.

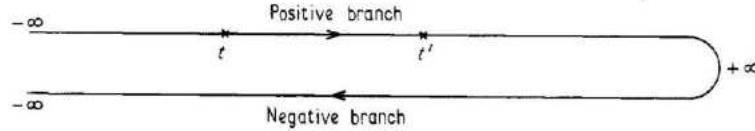


Figure 3. Representation of contour C .

In the preceding expression

$$S_c = S(-\infty, +\infty)S(+\infty, -\infty)$$

where the times of $S(-\infty, +\infty)$ are on the positive branch and the times of $S(+\infty, -\infty)$ on the negative branch. It is clear that the presence of this term will introduce times on the negative branch of the contour in the Dyson equation.

It is now possible to calculate the expression (A.1) with the usual diagram technique; but since the times belonging to the negative branch (noted with subscript $-$) are posterior to the times belonging to the positive branch (noted with subscript $+$), we must use four Green functions

$$\begin{aligned} G_{ij}^+(t, t') &= -i \langle \phi_0 | T_c \{ c_i(t_+) c_j^+(t'_-) \} | \phi_0 \rangle \\ &= i \langle \phi_0 | c_j^+(t') c_i(t) | \phi_0 \rangle \end{aligned}$$

$$\begin{aligned} G_{ij}^-(t, t') &= -i \langle \phi_0 | T_c \{ c_i(t_-) c_j^+(t'_+) \} | \phi_0 \rangle \\ &= -i \langle \phi_0 | c_i(t) c_j^+(t') | \phi_0 \rangle \end{aligned}$$

$$\begin{aligned} G_{ij}^c(t, t') &= -i \langle \phi_0 | T_c \{ c_i(t_+) c_j^+(t'_+) \} | \phi_0 \rangle \\ &= -i \langle \phi_0 | T \{ c_i(t) c_j^+(t') \} | \phi_0 \rangle \end{aligned}$$

$$\begin{aligned} \tilde{G}_{ij}^c(t, t') &= -i \langle \phi_0 | T_c \{ c_i(t_-) c_j^+(t'_+) \} | \phi_0 \rangle \\ &= -i \langle \phi_0 | \tilde{T} \{ c_i(t) c_j^+(t') \} | \phi_0 \rangle \end{aligned}$$

^{12†} Abrikosov, Gorkov, Dzialoshinski, 1965, Quantum Field Theory in Statistical Physics, (Oxford: Pergamon) p. 49.

where T orders the times from $-\infty$ to $+\infty$ and $\tilde{T}$ from $+\infty$ to $-\infty$.

In a diagram calculation the integration is performed along the contour C ; this is equivalent to perform the integration from $-\infty$ to $+\infty$, and to sum over subscripts $+$ and $-$. Because the negative branch of the contour C goes from $+\infty$ to $-\infty$, any point of this branch corresponds to a $(-)$ sign, or equivalently any interaction on this branch corresponds to a $(-)$ sign.

Summing over subscripts $+$ and $-$ is equivalent to using 2×2 matrices. Therefore the usual diagram technique may be used, if one defines a Green function matrix

$$\mathbf{G} = \begin{pmatrix} G^c & G^- \\ G^+ & \tilde{G}^c \end{pmatrix}$$

A Dyson's equation may then be written

$$G_{ij}(t, t') = G_{ij}^0(t, t') + \int G_{ik}^0(t, t_1) \Sigma_{kl}(t_1, t_2) G_{ij}(t_2, t') dt_1 dt_2 \quad (\text{A.2})$$

where

$$\Sigma = \begin{pmatrix} \Sigma^c & \Sigma^- \\ \Sigma^+ & \tilde{\Sigma}^c \end{pmatrix}$$

is the self-energy matrix and $\mathbf{G}^0$ is the Green function matrix in the absence of the interaction.

The functions $G^\pm, G^c, \tilde{G}^c$ are connected by

$$G^c + \tilde{G}^c = G^+ + g^-$$

In the same way

$$\Sigma^c + \tilde{\Sigma}^c = -(\Sigma^+ + \Sigma^-)$$

Thus among the four equations represented by equation (A.2), only two are independent. This becomes clearer if we make the following canonical transformation:

$$\mathbf{G} \rightarrow \frac{1 - i\sigma_y}{2} \begin{pmatrix} G^c & G^- \\ G^+ & \tilde{G}^c \end{pmatrix} \frac{1 + i\sigma_y}{2} = \begin{pmatrix} 0 & G^a \\ G^r & F \end{pmatrix} \quad (\text{A.3})$$

where $\dagger$

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

$$G^a = G^c - G^+ = -\tilde{G}^c + G^-$$

$$G^r = G^c - G^- = -\tilde{G}^c + G^+$$

$$F = G^c + \tilde{G}^c = G^+ + G^-$$

and

$$\begin{pmatrix} \Sigma^c & \Sigma^- \\ \Sigma^+ & \tilde{\Sigma}^c \end{pmatrix} \text{ becomes } \begin{pmatrix} \Omega & \Sigma^r \\ \Sigma^a & 0 \end{pmatrix}$$

with

$$\Sigma^a = \Sigma^c + \Sigma^+ = -(\tilde{\Sigma}^c + \Sigma^-)$$

$$\Sigma^r = \Sigma^c + \Sigma^- = -(\tilde{\Sigma}^c + \Sigma^+)$$

$$\Omega = \Sigma^c + \tilde{\Sigma}^c = -(\Sigma^+ + \Sigma^-)$$

Thus we obtain three equations and the Fourier transforms of the equations for G^a and G^r are evidently complex conjugates.

† Keldysh's notation for G^a and G^r , that we use there, is not the usual one. This last is obtained by exchanging a and r .

In our case, the perturbation T is instantaneous and

$$H_{\text{int}}(t) = \{TC_{\alpha}^{+}(t)C_a(t) + T'C_{\alpha'}^{+}(t)C_b(t)\} + \text{hc}$$

where time t is either on the positive or on the negative branch of contour C so that H_{int} cannot connect the two branches.

Consequently Σ has the following form in the transformed representation of equation (A.3):

$$\sigma = \begin{pmatrix} 0 & T \\ T & 0 \end{pmatrix} (\delta_{p\alpha}\delta_{qa} + \delta_{pu}\delta_{q\alpha}) + \begin{pmatrix} 0 & T' \\ T' & 0 \end{pmatrix} (\delta_{p\alpha'}\delta_{qb} + \delta_{pb}\delta_{q\alpha'})$$

9.3 Easy illustrative examples

9.3.1 Example 1

9.3.2 Example 2

10 Tunneling current II: Free electron description by Caroli et. al.

The theory of tunnelling introduced in a preceding paper is reformulated in the case of free electrons subjected to an arbitrary one-electron potential. An exact expression for the current is obtained, using Keldysh's generalization of perturbation theory. This result is particularly suited to an exact field theoretical treatment of manybody effects in the tunnelling junction. This paper complements our original work based on a discrete representation in terms of localized Wannier orbitals, and avoids the corresponding approximations.

(Все-таки обычную 1д квантовую механику таким способом тоже что ли можно решать??? порешаю пару задач!!)

10.1 1. Introduction

In a recent paper (Caroli et al. 1971, to be referred to as I) a new method for calculating the tunnelling current in metal-insulator-metal (MIM) junctions was proposed, based on Keldysh's nonequilibrium perturbation formalism (Keldysh 1965). This method does not rely on any effective Hamiltonian approximation and gives an exact expression for the current. It rests on the description of the MIM system in terms of localized electron wave functions (analogous to the Wannier functions of an infinite homogeneous system). Each of these localized functions (for a given energy band) is associated with a definite atomic site. The use of such a 'discrete' representation permits the separation of the system into three parts (two electrodes and the barrier) to appear in the formalism naturally and without any ambiguity.

Neglecting many-body effects and taking, for the sake of simplicity, a onedimensional model, the Hamiltonian reads

$$H = \sum_{ij} T_{ij} C_i^{+} C_j \quad (1)$$

The coupling between, for example, one electrode (M) and the barrier (I) is provided by those T_{ij} terms for which $i(j)$ belongs to (M) and $j(i)$ to I.

This model stems from a point of view which is rather different from the usual one. Since the free electron model is, in most cases, the simplest representation of a metal, it has been-quite naturally-assumed that the most simple model of a tunnel junction consists in a free-electron gas submitted to a square-barrier potential. In the absence of any atomic potential, it is impossible to define a discrete representation, since there are no atomic sites $\ddagger$. The Hamiltonian must be written (in the real space representation) as

$$H = \int dx \psi^\dagger(x) \left\{ -\frac{\hbar^2}{2m} \nabla^2 + V(x) \right\} \psi(x) \quad (2)$$

where $V(x)$ is the barrier potential. This can be understood as the limiting form of equation (1) when the intersite distance $(x_{i+1} - x_i) = a \rightarrow 0$. In that limit, the pseudo-Wannier function Φ_i corresponding to C_i^+ becomes a delta function.

As long as the intersite distance a remains finite, it is always possible to decide unambiguously whether a given site belongs to M, I or M', and one can express all properties of the system in terms of the properties of its separate parts. On the contrary, in the square barrier problem (figure 1), since we have to use the continuous representation of equation (2)

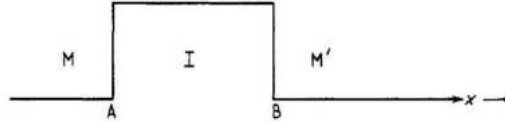


Figure 1

if we want to separate the system into its subparts, we must be able to define each point of the x axis as belonging to a given subpart. This is obviously impossible to do for the two discontinuity points A and B : point A, for instance, belongs to both M and I. We believe that the difficulties one faces when trying to solve the square barrier model, and which are usually avoided by means of an effective Hamiltonian, stem from this boundary problem, and are of a mathematical and not of a physical nature.

These difficulties can be completely avoided in the following way: we replace the Schrödinger equation by an (approximate) difference equation. By this procedure, the initial continuous problem becomes formally equivalent to the discrete problem of I, the role of the intersite distance being now played by the spacing ϵ of the difference equation. Using the results of I, we can solve this ad hoc discrete model exactly. The continuous model is obtained at its $\epsilon = 0$ limit.

In this way, we may unambiguously divide the system into two parts, left and right, and express the full propagator in terms of the individual propagators on each side of the cut. The current across the cut at finite bias may then be evaluated using Keldysh's method, the transfer term between the two halves being treated to all orders of perturbation theory. This approach is somewhat simpler than the one used in I, where the system was divided into three parts, rather than two (namely the electrodes and the barrier). While this could be done easily in the language of this paper, it leads to unnecessary complications: a single cut is enough to isolate the two regions with different chemical potentials, which is all that we need in order to take advantage of Keldysh's method.

In §2, we show that, with the help of this method, one obtains an exact formal expression for the current, which is valid whatever the shape of the potential in the electrodes as well as in the barrier. In §3 this formal result is applied to the square barrier model.

¹² $\ddagger$ This problem also arises when the periodic potential is small: in that case one can still define localized pseudo-Wannier functions, but, due to the smallness of the potential a great number of energy bands have to be retained in the calculation, so that the continuous representation becomes more adequate.

10.2 2. Expression of the current in the continuous representation

10.2.1 2.1. The Green function at zero bias

Let us consider a one dimensional MIM system, in the independent electron approximation, and let us call $I(x)$ the complete one-electron potential: that is V contains the periodic (pseudo) potentials of the two metallic electrodes, as well as the barrier potential. Let $\Phi(x)$ be an electron eigenfunction for this system. The Schrödinger equation may be written

$$\int H(x, x') \Phi(x') dx' = E\Phi(x) \quad (3)$$

where

$$H(x, x') = -\frac{\hbar^2}{2m} \delta''(x - x') + V(x) \delta(x - x') \quad (4)$$

We choose to use the integral representation of the Schrödinger equation in order to make the comparison easier with the discrete representation of I . In the same way, the equation of motion of the electron Green function reads

$$\int dx'' \{ \omega \delta(x - x'') - H(x, x'') \} \mathbf{G}_\omega(x'', x') = \delta(x - x') \quad (5)$$

The essential difference between $H(x, x')$ and the corresponding term T_{ij} of the discrete representation lies in the local (zero space range) character of $H(x, x')$.

If we choose to cut the system described by equation (5) into two subparts $x < a$ and $x > a$, the coupling between these subparts does not appear as a transfer term between the two sides of the $x = a$ plane (as it would be in the discrete case), but comes in only, due to the local character of H , via a $\delta(x - a)\delta(a - x')$ term in $H(x, x')$, ie via boundary conditions. Therefore, in order to make the separation process clear, one is directly led to consider the Hamiltonian (4) as the limit of a Hamiltonian with a finite range ϵ in $x - x'$. We choose

$$H^{(\epsilon)}(x, x') = -\frac{\hbar^2}{2m\epsilon^2} \{ \delta(x - x' - \epsilon) - 2\delta(x - x') + \delta(x - x' + \epsilon) \} + V(x) \delta(x - x') \quad (6)$$

It is clear that

$$H(x, x') = \lim_{\epsilon \rightarrow 0} H^{(\epsilon)}(x, x') \quad (7)$$

The Schrödinger equation becomes

$$-\frac{\hbar^2}{2m\epsilon^2} \{ \Phi^{(\epsilon)}(x + \epsilon) - 2\Phi^{(\epsilon)}(x) + \Phi^{(\epsilon)}(x - \epsilon) \} + V(x) \Phi^{(\epsilon)}(x) = E\Phi^{(\epsilon)}(x) \quad (8)$$

Therefore the introduction of a nonlocal character in H simply amounts to the transformation of the exact Schrödinger equation into a difference equation, as is classical in all methods of numerical integration.

Let us divide the x axis into intervals of width ϵ . From equation (8) it follows that, when $\Phi^{(\epsilon)}(x)$ is known in any two intervals (eg $(x_0 - \epsilon, x_0)$ and $(x_0, x_0 + \epsilon)$), it is determined on the entire x axis. Up to that point the precise shape of $\Phi^{(\epsilon)}$ in the two initial intervals is arbitrary.

^{12†} Note that the discrete model that we introduce here is completely different in nature from the one used in I: while the discrete representation of I is attached to the existence of atomic sites, the one we use here is a purely mathematical construction, which has nothing to do with the Wannier representation.

The only condition on the set $\{\Phi^{(\epsilon)}\}$ is that the $\lim_{\epsilon \rightarrow 0} \{\Phi^{(\epsilon)}(x)\} = \Phi(x)$, where $\Phi(x)$ is the solution of equation (3). This amounts to defining $\Phi(x)$ by means of a limiting process similar to the one used in the theory of integration. It is known that a good $\{\Phi^{(\epsilon)}\}$ set consists of 'stair-like' function (constant in each interval): this corresponds to the Riemann integral (see figure 2).

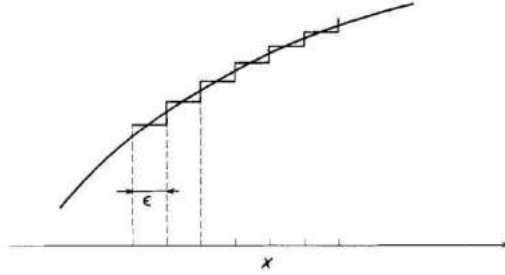


Figure 2

Let us call $\Phi_p^{(\epsilon)}$ the value of $\Phi^{(\epsilon)}$ in the p th interval. Equation (8) now becomes

$$-\frac{\hbar^2}{2m\epsilon^2} \left\{ \Phi_{p+1}^{(\epsilon)} - 2\Phi_p^{(\epsilon)} + \Phi_{p-1}^{(\epsilon)} \right\} + V_p \Phi_p^{(\epsilon)} = E \Phi_p^{(\epsilon)} \quad (9)$$

where V_p is a 'stair-like' approximation of $V(x)$.

In the same way, we define a 'discrete' Green function $G_{pp}^{(\epsilon)}(\omega)$, which is given by

$$\sum_{p''} (\omega \delta_{pp''} - H_{pp''}) G_{p''p'}^{(\epsilon)} = \frac{\delta_{pp'}}{\epsilon} \quad (10)$$

where

$$H_{pp''} = \frac{\hbar^2}{2m\epsilon^2} (2\delta_{pp''} - \delta_{p,p''-1} - \delta_{p,p''+1}) + V_p \delta_{pp''} \quad (11)$$

In order to avoid normalization complications, we also introduce the 'normalized' Green function $G = \epsilon G^{(\epsilon)}$.

We are now left with a problem formally identical to the discrete one of I (with nearest neighbour transfer only), the role of the site indices of the discrete case is played here by the interval indices p . The transfer term has a strength $|\hbar^2/2m\epsilon^2|$.

Let us now cut the system at a point x_0 that we choose as the end of interval i (figure 3).

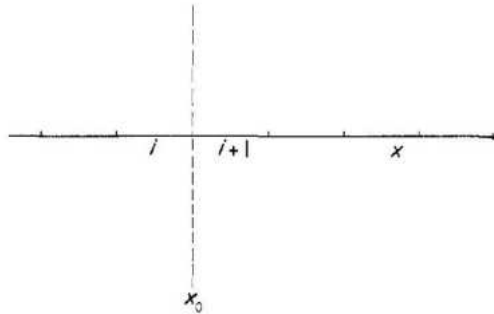


Figure 3

We want to express the Green function G of the entire system in terms of the Green functions g and $\tilde{g}$ of respectively the right and left parts, and of the transfer terms between these two parts. The (normalized) propagator of the right part $g_{pp'}$ (p and $p' \geq i+1$) is the solution of

$$\sum_{p''} (\omega \delta_{pp''} - h_{pp''}) g_{p''p'} = \delta_{pp'} \quad (12)$$

with

$$h_{pp''} = \begin{cases} H_{pp''} & \text{if } p \text{ and } p'' \geq i+1 \\ 0 & \text{if } p \text{ or } p'' < i+1. \end{cases}$$

$g_{pp'}$ is zero if either of the two indices p and p' belongs to the left part. A similar definition holds for the left part propagator $\tilde{g}$ (with $p, p' \leq i$).

The coupling term is provided by $(H_{i,i+1} + H_{i+1,i})$: as already mentioned, this coupling is exactly a nearest neighbour transfer (such a type of coupling was used in I: in the 'discrete problem', this is only an approximation, as the real transfer may have a longer range).

Proceeding as in I, and setting $T = -\hbar^2/2m\epsilon^2$, we obtain

$$G_{lk} = \begin{cases} \frac{\tilde{g}_{li} T g_{i+1,k}}{D} & l \leq i; k \geq i+1; \\ \tilde{g}_{lk} + \frac{\tilde{g}_{li} T g_{i+1,i+1} T \tilde{g}_{ik}}{D} & l, k \leq i \end{cases}$$

and particularly

$$G_{ii} = \frac{\tilde{g}_{ii}}{D}; \quad G_{i,i+1} = \frac{\tilde{g}_{ii} T g_{i+1,i+1}}{D} \quad (13)$$

with $D = 1 - T^2 \tilde{g}_{ii} g_{i+1,i+1}$.

When one goes to the continuous limit $\epsilon \rightarrow 0, T \rightarrow \infty$, $G^{(\epsilon)}$ must be finite. Consequently we have to study in more detail the behaviour of g and $\tilde{g}$ when $\epsilon \rightarrow 0$.

For that purpose, we introduce a more general propagator: $g_l(m, n)$, which is the Green function of the semi-infinite medium $p \geq l$ (ie where a cut has been made between intervals $l-1$ and l , $\tilde{g}_l(m, n)$ is, similarly, the propagator of the semi-infinite medium $p \leq l$ (cut between l and $l+1$).

Let us now cut the system between intervals l and $l+1$ (figure 4).

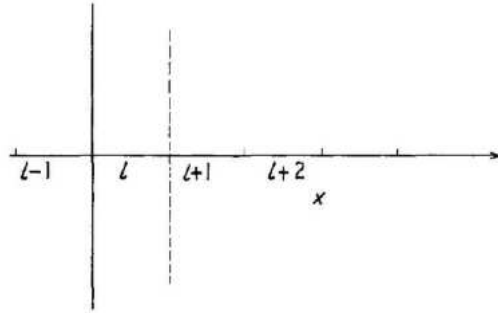


Figure 4

Using the same method that led to equation (13), we express the right-hand side quantity $g_l(l) \equiv g_l$ in terms of the propagators of the two sides of the cut: g_{l+1} for the right side, g_l^0 for the left one, which simply consists of the isolated l th interval.

$$g_l^0 = \frac{1}{\omega + 2T - V_l} \quad (14)$$

We obtain

$$g_l = \frac{g_l^0}{1 - T^2 g_l^0 g_{l+1}} \quad (15)$$

which can be rewritten as

$$\frac{1}{g_l} = \omega + 2T - V_l - T^2 g_{l+1} \quad (16)$$

¹²† Note that if $V(x)$ has a discontinuity at some point x_0 , this point must be taken as the origin of one of the intervals.

We will now prove, using equation (16), that when $\epsilon \rightarrow 0$, the behaviour of g_l is of the following form:

$$g_l = \frac{1}{T} \{1 + \epsilon \gamma_l + \epsilon^2 \delta_l + O(\epsilon^3)\} \quad (17)$$

where γ_l and δ_l have finite limits $\gamma(x), \delta(x)$ when $\epsilon \rightarrow 0$. Indeed, when equation (17) is inserted into equation (16), the $T(\sim \epsilon^{-2})$ terms cancel and we obtain

$$T \{-\epsilon \gamma_l + \epsilon^2 (\gamma_l^2 - \delta_l)\} = \omega - V_l - T(\epsilon \gamma_{l+1} + \epsilon^2 \delta_{l+1}) + O(\epsilon^3). \quad (18)$$

When $\epsilon \rightarrow 0$, $\gamma_{l+1} - \gamma_l \rightarrow \epsilon d\gamma(x)/dx$. Identifying the zeroth-order terms in equation (18), and going to the $\epsilon = 0$ limit, we obtain the equation of definition of the, continuous function $\gamma(x)$

$$-\frac{\hbar^2}{2m} \left(\frac{d\gamma}{dx} + \gamma^2 \right) = \omega - V(x) \quad (19)$$

A more explicit expression for $\gamma(x_0)$ -which is connected with the propagator in the semiinfinite medium $x > x_0$ -will be given later †. In exactly the same way, we obtain for $\tilde{g}_l$

$$\tilde{g}_l = \frac{1}{T} \{1 + \epsilon \tilde{\gamma}_l + O(\epsilon^3)\} \quad (20)$$

and $\lim_{\epsilon \rightarrow 0} \tilde{\gamma}_i = \tilde{\gamma}(x)$, where $\tilde{\gamma}(x)$ is the solution of

$$-\frac{\hbar^2}{2m} \left(\tilde{\gamma}^2 - \frac{d\tilde{\gamma}}{dx} \right) = \omega - V(x) \quad (21)$$

Inserting these results in equation (13), we obtain

$$G(x_{0-}, x_{0+}) = \lim_{\epsilon \rightarrow 0} G_{ii}^{(\epsilon)} = + \frac{2m}{\hbar^2} \frac{1}{\gamma(x_0) + \tilde{\gamma}(x_0)} \quad (22)$$

$G(x_0, x_0)$, a propagator of the entire system, is now expressed in terms of the two functions $\gamma(x_0), \tilde{\gamma}(x_0)$, each of which is directly related to the Green function of one of the subparts.

With the help of equation (20), we can also calculate

$$G(x_{0-}, x_{0+}) = \lim_{\epsilon \rightarrow 0} G_{i,i+1}^{(\epsilon)} = G(x_{0-}, x_{0+}) \quad (23)$$

As it should be, the propagators of the entire system are continuous across the cut. Finally, and in order to be able to express $G(x, x')$ for $x, x' \neq x_0$, we need to study the behavior of $g_i(i)$ and $g_i(jk)$ in the $\epsilon = 0$ limit. This is done in the Appendix.

In particular, it is found that ‡

$$g_i(il) \xrightarrow{\epsilon \rightarrow 0} \frac{1}{T} \exp \int_{x_i}^{x_i} \gamma(t) dt \quad (24)$$

and in the same way

$$\tilde{g}_i(il) \xrightarrow{\epsilon \rightarrow 0} \frac{1}{T} \exp \int_x^{x_i} \tilde{\gamma}(t) dt \quad (25)$$

Making use of equations (24), (25) and (13), we obtain, for the full propagator $G(x_1, x_2)$ connecting two points belonging to different subparts ($x_1 < x_0 < x_2$)

$$G(x_1, x_2) = \frac{2m}{\hbar^2 \{\gamma(x_0) + \tilde{\gamma}(x_0)\}} \exp \left\{ \int_{x_1}^{x_0} \tilde{\gamma}(t) dt + \int_{x_0}^{x_2} \gamma(t) dt \right\} \quad (26)$$

and if $x_1, x'_1 < x_0$

$$G(x_1, x'_1) = \tilde{g}(x_1, x'_1) + \frac{2m}{\hbar^2 \{\gamma(x_0) + \tilde{\gamma}(x_0)\}} \exp \left\{ \int_{x_0}^{x_1} \tilde{\gamma}(t) dt + \int_{x_0}^{x'_1} \tilde{\gamma}(t) dt \right\} \quad (27)$$

Note that equations (19) and (21) are the conventional Riccati equations associated with the Schrödinger equations.

‡ Note that in equation (24) $g_i \equiv g_i(ii)$ is the limit of $g_i(il)$ for $l \rightarrow i$, which will be of help when we want to calculate γ , (ie g_i) explicitly for a given shape of the potential.

Therefore the use of the above limiting process enables us to express all the propagators of the entire system in terms of quantities ($g, \tilde{g}, \gamma$ and $\tilde{\gamma}$) which describe its separate subparts. without any ambiguity. g and $\tilde{g}$ are ordinary Green functions for semi-infinite media. γ and $\tilde{\gamma}$ are solutions of the nonlinear differential equations (19) and (21).

Since nonlinear equations are in general rather difficult to solve, we would like to have some alternative simpler expression of γ and $\tilde{\gamma}$. This can be obtained in the following way: let us call $\Psi_\alpha(x)$ the eigenfunction of the right-hand semi-infinite medium $x > x_0$, with energy E_α . By definition $\Psi_\alpha(x_0) = 0$, and

$$g_\omega(x, x') = \sum_\alpha \frac{\Psi_\alpha^*(x) \Psi_\alpha(x')}{\omega - E_\alpha} \quad (28)$$

On the other hand, γ has been initially defined from $g_i(ii)$ (equation (17)). When we go to the continuous limit $\epsilon \rightarrow 0$,

$$g_i^{(\epsilon)}(ii) \simeq g(x_0 + \epsilon, x_0 + \epsilon) \simeq -\frac{2m\epsilon}{\hbar^2} \{1 + \epsilon\gamma(x_0)\}. \quad (29)$$

There may seem to be a contradiction between equations (28) and (29): since $\Psi_\alpha(x_0) = 0$, we can write $\Psi_\alpha(x_0 + \epsilon) \simeq \epsilon \Psi'_\alpha(0)$, and from equation (23), we then expect $g(x_0 + \epsilon, x_0 + \epsilon)$ to be of order ϵ^2 , and not ϵ . It is shown in the Appendix that, for reasons of convergence of the α sum at high energy, one must first perform the α sum in equation (28) before expanding the Ψ_α around $x_x = x_0$. One then checks that the leading term in $g(x_0 + \epsilon, x_0 + \epsilon)$ is effectively equal to $-2m\epsilon/\hbar^2$, whatever the potential $V(x)$, and that

$$\gamma(x_0) = -\frac{\hbar^2}{2m} \lim_{x \rightarrow x_0} \sum_\alpha \frac{\Psi_\alpha^*(x_0) \Psi'_\alpha(x)}{\omega - E_\alpha} = \lim_{x \rightarrow x_0} \lim_{x' \rightarrow x_0} \left(-\frac{\hbar^2}{2m} \right) \frac{\partial}{\partial x} \frac{\partial}{\partial x'} g(x, x'). \quad (30)$$

A similar definition holds for $\tilde{\gamma} \cdot \gamma$ and $\tilde{\gamma}$ are therefore determined as soon as g and $\tilde{g}$ are known. We thus see that the properties of the entire system are completely expressible in terms of those of its two subparts.

10.2.2 2.2 Formal expression of the current

In the continuous representation, the current density at point x_0 is given (for a single spin direction) by

$$\langle J \rangle = \frac{e\hbar'}{2m} \lim_{x' \rightarrow x} \left(\frac{\partial}{\partial x} - \frac{\partial}{\partial x'} \right) G^+(xt, x't) \Big|_{x=x_0} \quad (31)$$

where we use Keldysh's definition of G^+ :

$$G^+(xt, x't') = i \langle \Psi^+(x't') \Psi(xt) \rangle \quad (32)$$

As in §2.1, we transform the differential definition (31) of $\langle J \rangle$ into a difference relation. The current at a point x_0 lying between 'sites' 0 and 1 is equal to

$$\langle J \rangle = \frac{e\hbar}{2m} \int \frac{\{G_{\omega}^{+(\epsilon)}(x_0 + \epsilon, x_0) - G_{\omega}^{+(\epsilon)}(x_0, x_0 + \epsilon)\}}{\epsilon} d\omega = -\frac{eT}{\hbar} \int \frac{d\omega}{2\pi} (G_{10}^+ - G_{01}^+) \quad (33)$$

where, as above, $T = -\hbar^2/2m\epsilon^2$ and $G_{pp'}^+ = \epsilon G_{pp'}^{+(\epsilon)}$.

As expected, the expression of $\langle J \rangle$ is formally identical to what was obtained in I for the discrete representation. Note that the cutting point x may be chosen anywhere in the medium. However, since we want to compute the response to a finite bias, it is more convenient in practice to take x inside the insulator or at one of its edges.

We are left with the problem of expressing G_{10}^+ and G_{01}^+ in terms of the right-hand and left-hand side propagators in an out of equilibrium system. As in I this is done in the frame of Keldysh's extended perturbation theory (for more detail about this method see Keldysh (1965) and I). The perturbation is provided by the transfer coupling between the two sides of the cut-which have different chemical potential-and we calculate the current to all orders in that coupling. Following Keldysh, we introduce Green functions and self-energy matrices:

$$\hat{G}_{pq} = \begin{pmatrix} G_{pq}^c & G_{pq}^- \\ G_{pq}^+ & \tilde{G}_{pq}^c \end{pmatrix} \quad \hat{\Sigma}_{pq} = \begin{pmatrix} \Sigma_{pq}^c & \Sigma_{pq}^- \\ \Sigma_{pq}^+ & \tilde{\Sigma}_{pq}^c \end{pmatrix} \quad (34)$$

(where p and q are interval indices) and the advanced and retarded functions

$$\begin{aligned} G^a &= G^c - G^+ = -\tilde{G}^c + g^- \\ G^r &= G^c - G^- = -\tilde{G}^c + g^+ \\ \Sigma^a &= \Sigma^c + \Sigma^+ = -(\tilde{\Sigma}^c + \Sigma^-) \\ \Sigma^r &= \Sigma^c + \Sigma^- = -(\tilde{\Sigma}^c + \Sigma^+) \end{aligned} \quad (35)$$

The causal and anticausal functions G^c and $\tilde{G}^c$. as well as G^a and G^r . are related to the dynamical properties of the system, while G^+ and G^- describe the occupation of respectively the hole and particle states.

Since the one-body coupling is restricted to a transfer between intervals 0 and 1, $\hat{\Sigma}_{pq}$ takes the simple form (see Appendix of I (??? я не понял, что там важного???)):

$$\hat{\Sigma}_{pq} = T\hat{1}(\delta_{p0}\delta_{q1} + \delta_{p1}\delta_{q0}). \quad (36)$$

With these notations, the Dyson equation reads

$$\hat{G}_{pq} = \hat{g}_{pq} + \sum_{lm} \hat{g}_{pl} \Sigma_{lm} \hat{G}_{mq} \quad (37)$$

$\hat{g}$ is the zero-order Green matrix, that is, it describes the uncoupled right and left part of the cut.

All the expressions that have been obtained in §2.1 for the equilibrium Green function are transformed into matrix equations. After some simple but tedious algebra, we obtain

$$T[G_{10}^+ - G_{01}^+] = \frac{T}{1 - T^2 g_{11}^a g_{00}^a} [g_{00}^+ g_{11}^- - g_{00}^- g_{11}^+] \frac{T}{1 - T^2 g_{11}^r g_{00}^r} \quad (38)$$

Since the isolated right and left semi-infinite media are separately at equilibrium, $\hat{g}_{00}$ and $\hat{g}_{11}$ are directly obtained from the results of §2.1. For instance (see I)

$$\begin{aligned}
g_{00}^a &= (g_{00}^r)^* = \frac{\tilde{g}_{00}}{\tilde{g}^*} \quad \text{for } \omega > \mu_L \\
&= 2i \operatorname{Im} \tilde{g}_{00} = 0 \quad \text{for } \omega > \mu_L
\end{aligned} \tag{39}$$

where μ_L is the chemical potential of the left electrode. Similar relations (with $\tilde{g}_{00} \rightarrow g_{11}$ and $\mu_L \rightarrow \mu_R$) hold for $\tilde{g}_{11}$. Going to the $\epsilon = 0$ limit, and using equations (17) and (20), we obtain, for both spin directions

$$\langle J \rangle = \frac{8e}{\hbar} \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \frac{\operatorname{Im} \{\gamma(x_0)\} \operatorname{Im} \{\tilde{\gamma}(x_0)\}}{|\gamma(x_0) + \tilde{\gamma}(x_0)|^2} \left[\Theta(\mu_L - \omega) \Theta(\omega - \mu_R) - \Theta(\mu_R - \omega) \Theta(\omega - \mu_L) \right] \tag{40}$$

where Θ is the ordinary step function, and the right and left γ functions are defined by equation (30).

One easily checks, with the help of equations (19) and (21), that

$$\frac{d}{dx} \left\{ \frac{\operatorname{Im} \gamma(x) \operatorname{Im} \tilde{\gamma}(x)}{|\gamma(x) + \tilde{\gamma}(x)|^2} \right\} = 0$$

i.e. that expression (40) for the current is effectively independent of the position of the cutting point x_0 .

If, for instance, we call μ the chemical potential of the right electrode, and if the junction is biased at potential V ($\mu_L = \mu + eV > \mu$), the current reads

$$\langle J \rangle = \frac{8e}{\hbar} \int_{\mu}^{\mu+eV} \frac{d\omega}{2\pi} \frac{\operatorname{Im} \gamma(x_0) \operatorname{Im} \tilde{\gamma}(x_0)}{|\gamma(x_0) + \tilde{\gamma}(x_0)|^2} \tag{41}$$

Equation (40) provides an exact expression of the current, in terms of the properties of two isolated right and left semi-infinite media. As was already emphasized in § 2.1 and since the current is conserved along the x axis, the cutting point x_0 can be chosen anywhere in the system. If we place it for instance at the M-I contact, we see from equation (30) that $\gamma(x_0)$, which is related to the properties of the I – M' subsystem, is exponentially small. The exponential dependence of the current in the insulator thickness, which appears usually via a 'barrier transmission coefficient', is therefore included, in our formalism, in γ , or $\tilde{\gamma}$, or in both, depending on whether x_0 is chosen to lie at one of the barrier-electrode contacts, or inside the barrier. As in I, since γ and $\tilde{\gamma}$ are functions of ω , this 'transmission coefficient' is energy dependent.

Finally, it is important to note that, contrary to what was deduced from the transfer Hamiltonian type of approximations, the properties of the electrodes do not appear in the current via the volume density of states of the isolated **M** and M'. They come into play by means of γ functions: these are related to double derivatives of the propagators of the isolated electrodes at the contact surfaces and cannot therefore be expressed directly in terms of densities of states.

As it stands, the result (40) does not include many-body effects. Indeed, it may seem worthless to build such an involved formalism in order to obtain the well known result of elementary one-electron theories. Actually, the interest of our approach is that it can be generalized to include many-body effects without any conceptual difficulty, in contrast to the usual methods. The expression (40), being explicitly energy dependent, includes retardation effects, and is perfectly fit to describe inelastic phenomena. This will be studied in a separate paper, where we shall show that many-body effects appear through a renormalization of g and γ .

10.3 3. The current through a square barrier

In order to make our results clearer, we apply them to the oversimplified model of the free-electron and square-barrier model. At zero bias, the potential profile is shown on figure 5. We choose to cut the system at the point of symmetry $x = 0$. Clearly, we then have $\gamma_\omega(0) = \tilde{\gamma}_\omega(0)$. For a finite bias, the potential profile becomes that shown by the heavy line of figure 6. The height and shape of the barrier are modified by the bias, an effect which gives rise to the so-called 'background current'. Here we shall neglect this effect and we shall assume that the height of the barrier remains equal to V_0 in the presence of the bias V .

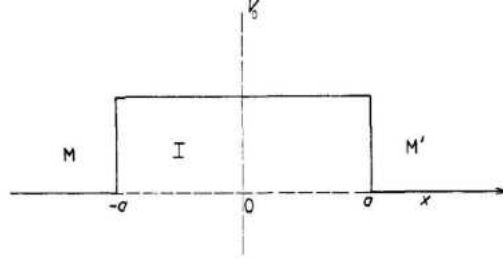


Figure 5

The bias then acts as a mere shift of the energy scale on one side, and we may write

$$\gamma_\omega(0) = \tilde{\gamma}_{\omega+eV}(0). \quad (42)$$

Strictly speaking, this relation would be exact if the potential profile in the barrier was that shown by the broken line of figure 6.

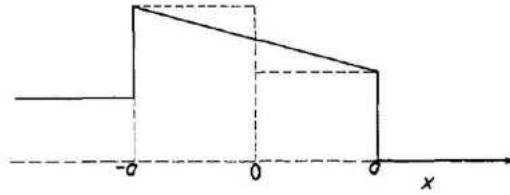


Figure 6

To the extent that we neglect the background current, the shape of the barrier is irrelevant and equation (42) is valid. This approximation will be correct if the bias eV is much smaller than the barrier height V_0 .

We are thus left with the problem of calculating $\gamma_\omega(0)$ on one side of the cut, that is of solving

$$-\frac{\hbar^2}{2m} \left[\frac{d\gamma}{dx} + \gamma^2 \right] = \begin{cases} \omega & x > a \\ \omega - V_0 & x < a. \end{cases} \quad (43)$$

The general solution of this equation is

$$\gamma_\omega(x) = \begin{cases} (-\tilde{\omega})^{1/2} \text{th} \{x(-\tilde{\omega})^{1/2} + b_1\} & x > a \\ (\tilde{V}_0 - \tilde{\omega})^{1/2} \text{th} \{x(\tilde{V}_0 - \tilde{\omega})^{1/2} + b_2\} & |x| < a \end{cases}$$

with

$$\tilde{\omega} = \frac{2m\omega}{\hbar^2} \quad \tilde{V}_0 = \frac{2mV_0}{\hbar^2}$$

Moreover γ must be continuous at $x = a$. We need one more boundary condition in order to determine the two constants b_1 and b_2 . We find it at once by noting that $\gamma(x_0)$ determines the propagator of the semi-infinite medium $x > x_0$, and thus depends only on the potential for

$x > x_0$. In particular, for $x > a$, $\gamma(x)$ is the same as for an infinite medium, with no potential barrier: in that case γ is independent of x , and equation (43) implies

$$\gamma_\omega(x) = (-\tilde{\omega})^{1/2} \quad x > a$$

(ie $b_1 = \infty$). The boundary conditions to be applied on γ and $\tilde{\gamma}$ are discussed in §3.

The continuity condition at $x = a$ determines b_2

$$(V_0 - \tilde{\omega})^{1/2} \text{th} \left\{ a \left(\tilde{V}_0 - \tilde{\omega} \right)^{1/2} + b_2 \right\} = (-\tilde{\omega})^{1/2} \quad (44)$$

which is sufficient to determine $\gamma(0)$

$$\gamma(0) = \left(\tilde{V}_0 - \tilde{\omega} \right)^{1/2} \text{th} b_2 = \left(\tilde{V}_0 - \tilde{\omega} \right)^{1/2} \frac{\left(\tilde{V}_0 - \tilde{\omega} \right)^{1/2} \text{th} \left\{ a \left(\tilde{V}_0 - \tilde{\omega} \right)^{1/2} \right\} - (-\tilde{\omega})^{1/2}}{(-\tilde{\omega})^{1/2} \text{th} \left\{ a \left(\tilde{V}_0 - \tilde{\omega} \right)^{1/2} \right\} - \left(\tilde{V}_0 - \tilde{\omega} \right)^{1/2}} \quad (45)$$

Neglecting for simplicity the applied bias eV with respect to the chemical potential μ of the right electrode ($\simeq 10$ electron volts), that is setting $\gamma_\omega(0) = \tilde{\gamma}_\omega(0)$, we obtain for the current

$$\langle J \rangle = \frac{8e}{\hbar} \int_{\mu}^{\mu+eV} \frac{d\omega}{2\pi} \frac{\omega (V_0 - \omega)}{4\omega (V_0 - \omega) + V_0^2 \text{sh}^2 \left[2a \left\{ 2m (V_0 - \omega) / \hbar^2 \right\}^{1/2} \right]} \quad (46)$$

In the usual tunnelling situation, $2a \left\{ 2m (V_0 - \omega) / \hbar^2 \right\}^{1/2} \gg 1$ (i.e. the transmission coefficient of a plane wave with wave vector $k \sim (2m\mu/\hbar^2)^{1/2}$ is small). Therefore, to lowest order, the final expression for $\langle J \rangle$ takes a very simple form

$$\langle J \rangle = \frac{8e}{\hbar} \int_{\mu}^{\mu+eV} \frac{d\omega}{2\pi} \frac{4\omega (V_0 - \omega)}{V_0^2} \exp \left[-4a \left\{ \frac{2m}{\hbar^2} (V_0 - \omega) \right\}^{1/2} \right] \quad (47)$$

Therefore, as expected, the current varies exponentially with the barrier thickness.

(!!! add plot here!)

10.4 4. Conclusion

In this paper, we have been able to extend formally the method of I to obtain an exact expression for the tunnelling current. This expression is valid whatever the precise shape of the one-electron potential in the electrodes. This extension has a conceptual advantage over the 'discrete method' of I: in I we had to make use of pseudo-Wannier localized functions, which are not known precisely in the vicinity of the I – M and I – M' contacts, and which are not a convenient representation when the periodic potential becomes small. These problems disappear completely in the 'continuous representation', where the calculation of the current reduces to that of the two functions γ and $\tilde{\gamma}$. Many-body effects can be included in this formalism as well as in the discrete one: actually, this is the main advantage of the present approach, a feature which will be exploited in another publication.

10.5 Appendix

We want to study the behavior of the semi-infinite media propagators $g_i(il)$ and $g_i(jk)$. Let us first recall the definition of $g_i(jk)(j, k > i)$: $g_i(jk)$ is the propagator of an electron from interval j to interval k , in the semi-infinite medium $p \geq i$ (i.e. the one-dimensional space has been cut between intervals $i - 1$ and i).

Let us first consider the behaviour of $g_i(il)$. Since we are dealing with a one dimensional system, all the paths going from i to l must go at least once through interval $(l - 1)$ (figure 7). Let us singularize on each path (P) the last crossing of the $(l - 1)$ interval. (P) is thus cut into two parts: (P₁) is any path starting from i and ending at $(l - 1)$ without ever going to the left of i , i.e. it is one of the paths contributing to $g_i(i, l - 1)$. In the same way, (P₂), which never goes to the left of l , is one of the paths contributing to $g_l(ll)$. Therefore we can write

$$g_i(il) = g_i(i, l - 1)Tg_l(ll) \quad (\text{A1})$$

or

$$g_i(il) - g_i(i, l - 1) = \{Tg_l(ll) - 1\} g_i(i, l - 1)_\epsilon \rightarrow_0 \epsilon_i g_i(i, l - 1) \quad (\text{A2})$$

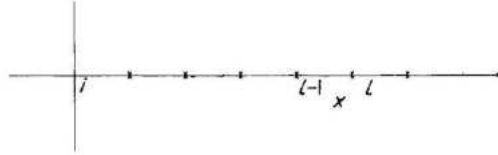


Figure 7

We see in equation (A2) that $g_i(il)$ is a continuous function of index i . Moreover, $g_i(ii)$ is the limit of $g_i(i)$ for $l \rightarrow i$, which will be of help where we want to calculate γ (i.e. $g_i(ii)$) explicitly for a given shape of the potential.

Going to the $\epsilon = 0$ limit, we obtain

$$g_i(il) \xrightarrow{\epsilon \rightarrow 0} \frac{1}{T} \exp \int_{x_i}^{x_l} \gamma(t) dt \quad (\text{A3})$$

Let us now consider the propagator $g_i(jk)(j, k \geq i + 1)$, with, for instance, $j < k$. The set of paths going from j to k without going out of the semi-infinite medium, $p \geq i$ can be separated into two groups:

- (i) The paths (P) which never go to the left of interval $i + 1$. They contribute to $g_{i+1}(jk)$.
- (ii) The paths (P') which go to the left of $i + 1$ (ie to interval i). Making the first and last $i + 1 \rightarrow i$ and $i \rightarrow i + 1$ transfers appear explicitly in the contribution of these paths to the propagator, we obtain

$$g_i(jk) = g_{i+1}(jk) + g_{i+1}(j, i + 1)Tg_i(ii)Tg_{i+1}(i + 1, k) \quad (\text{A4})$$

Let us now choose j and k to be both in the vicinity of i , so that we can expand the exponential term in equation (A3). Then

$$g_{i+1}(i + 1, k) \simeq g_{i+1}(j, i + 1) + \frac{1}{T} \quad (\text{A5})$$

and equation (A4) reads

^{12†} Equation (43) can be extended to the range $x < 0$: quite generally, $\gamma(x)$ determines the propagator for $x' > x$. When $x < -a$, $\gamma_\omega(x) = (-\tilde{\omega})^{1/2} \text{th} \{x(-\tilde{\omega})^{1/2} + b_3\}$: b_3 is not infinite, in contrast to b_1 . since in that case the propagators feel the effect of the insulating barrier.

We note that for an arbitrary potential $V(x)$ going to a given limit V_∞ when $x \rightarrow \infty$, one has $\gamma_\omega(x) = \sqrt{(\tilde{V}_\infty - \tilde{\omega})}^{1/2}$.

$$g_i(jk) \simeq g_{i+1}(j, k) + \frac{1}{T} \quad (\text{A6})$$

Setting $j = i + p (< k)$, we obtain, using equations (A5) and (A6)

$$g_i(i + p, k) \simeq \frac{p + 1}{T}$$

If $k < j$, a similar operation leads to

$$g_i(j, i + q) \simeq \frac{q + 1}{T}$$

so that, for j and k not too far from i ($\epsilon\gamma(j - i)$ and $\epsilon\gamma(k - 1) \ll 1$)

$$g_i(j, k) \simeq \{\min(j, k) - i + 1\} \frac{1}{T} \quad (\text{A7})$$

We can now deduce from equation (A7) the behaviour of $g(x, x') = \lim_{\epsilon \rightarrow 0} g_i^{(\epsilon)}(jk)$ for small but finite $x - x_0$ and $x' - x_0$. Setting $x - x_0 = (j - i)\epsilon$ and $x' - x_0 = (k - i)\epsilon$, we obtain

$$g_i(j, k) \simeq -\{\min(x - x_0, x' - x_0) + \epsilon\} \frac{2m\epsilon}{\hbar^2} \quad (\text{A8})$$

and finally

$$g(x, x') \simeq -\frac{2m}{\hbar^2} \min(x' - x_0, x - x_0) \quad (\text{A9})$$

Note that equation (A9) is but the first term of the expansion of $g(x, x')$ in powers of $(x - x_0)$ and $(x' - x_0)$, that is it is valid only when x and x' are close to the cutting point x_0 .

On the other hand, in the continuous representation, $g(x, x')$ can be written as

$$g_\omega(x, x') = \sum_\alpha \frac{\Psi_\alpha^*(x) \Psi_\alpha(x')}{\omega - E_\alpha} \quad (\text{A10})$$

where the Ψ_α are the eigenstates of the semi-infinite system $x > x_0$, with energy E_α , and

$$\Psi_\alpha(x_0) = 0 \quad (\text{A11})$$

From equations (A10) and (A11) one would at first sight infer that, when $x - x_0$ and $x' - x_0$ are both small, $g(x, x')$ behaves as $(x - x_0)(x' - x_0)$. This result is in fact not true, and the behaviour of g in this case is given by equation (A9). This paradox stems from the fact that, if the wave functions in (A10) are expanded around x_0 , one obtains the sum $\sum_z \{\Psi'_\alpha(x_0) \Psi'_\alpha(x_0)\} / (\omega - E_\alpha)$, which diverges at high energy: this is obvious for instance in the case of a free-electron semi-infinite medium, where $\Psi'_\alpha(x_0) = k$, and $E_\alpha = \hbar^2 k^2 / 2m$, and remains true in the presence of any one-electron potential. Therefore, one has to be somewhat cautious in studying the limit of expression (A10) for small $(x - x_0)$ and $(x' - x_0)$.

Let us now consider this expression for finite x' and $x - x_0 = \eta \rightarrow 0$. We can write

$$g_\omega(x_0 + \eta, x') \simeq \eta \sum_\alpha \frac{\Psi_\alpha^*(x) \Psi_\alpha(x')}{\omega - E_\alpha} \quad (\text{A12})$$

since the sum which appears in equation (A12) is now convergent-this can be checked immediately in the free electron case, where it takes the form

$$\sum_k \frac{k \sin kx'}{\omega - \hbar^2 k^2 / 2m}$$

On the other hand, we obtain from equation (A2)

$$\frac{\partial}{\partial x'} y(x_0 + \epsilon, x') = \gamma(x') g(x_0 + \epsilon, x') \quad (\text{A13})$$

Going to the $x' = x_0$ limit, and noticing that $\lim_{x' \rightarrow x_0} g(x_0 + \epsilon, x') = 1/T\epsilon$

$$\lim_{x' \rightarrow x_0+} \epsilon \sum_x \frac{\Psi_{\alpha}^{\prime*}(x_0) \Psi_{\alpha}'(x')}{\omega - E_z} = \gamma(x_0) \frac{1}{\epsilon T} \quad (\text{A14})$$

or finally

$$\gamma(x_0) = \lim_{x' \rightarrow x_0+} \lim_{x \rightarrow x_0+} -\frac{\hbar^2}{2m} \frac{\partial}{\partial x'} \frac{\partial}{\partial x} g_{\omega}(x, x') \quad (\text{A15})$$

10.6 Easy illustrative examples

10.6.1 Example 1

10.6.2 Example 2

11 Tunneling current III: Localized impurities in the barrier by Comberscot

The effect of an impurity level located in the barrier of a metal-insulator-metal tunneling junction is investigated with the help of a new formalism introduced in a preceding paper. In agreement with Hurault's results, it is found that a sharp resonance peak appears in the d^2I/dV^2 characteristic, when the impurity state energy crosses either electrode Fermi level. The width of this resonance peak is proportional to the width of the virtual bound state arising from the interaction of the two electrode continuum with the impurity state. It is stressed that the resonance effect cannot be interpreted as the opening of a new tunneling channel. The dependence of the effect on the location of the impurity state is briefly studied.

11.0.1 1. Introduction

In a recent paper (Caroli et al. 1971a), hereafter referred to as I, a new treatment of tunneling in a metal-insulator-metal junction (MIM) was displayed. This treatment, which gives an exact formal expression of the tunneling current, also enables us to deal properly with many body effects taking place in the electrodes or/and in the insulating barrier (Caroli et al. 1971 b hereafter referred to as II). In the present paper, we investigate, with the help of this method, the modifications of the tunneling characteristic of such a junction introduced by the presence in the barrier of dilute impurities. We only consider the case where each impurity has one electron bound state at an energy E lying inside the gap of the insulator, near the common Fermi level μ of the electrodes in the nonbiased junction (see figure 1).

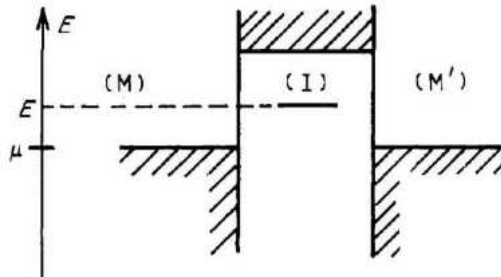


Figure 1

This problem has been studied by various authors (for a critical survey see Hurault 1969) in the past few years, and most recently by Hurault (1969 and 1971) who has considered the two cases of square well impurity potentials without and with an electron bound state, superimposed on a square barrier potential. He calculates the transmission of an electron plane wave by such a barrier, and can therefore obtain the current. The impurities are

found to influence quite strongly the tunneling characteristic in the case where they can bind an electron state: as can be expected, the presence of a finite density of states in the barrier enhances the tunneling probability, giving rise to an additional current at voltages corresponding to the energy of the bound state.

Although the essential features of this effect appear in Hurault's results, it is of some interest to rederive them in the frame of the general tunneling formalism of I. Indeed, the use of our method permits the results to appear in a much simpler and more systematic way, so that the physical aspects of the impurity effect can be discussed more easily. In particular, we shall be able to analyse the physical difference between the effects of impurities and of the electron-phonon interaction: an impurity bound electron state is often said to introduce a new channel for tunneling, in the same sense as phonon emission or absorption create a new tunneling channel. We shall show that this is in fact not true for impurities because of the absence of a continuum of electron states in the insulator that could permit the electron population of the bound states to relax. In § 2, we first solve the problem in the oversimplified case of a one dimensional model, which, however exhibits the main physical features of interest. In § 3 we generalize this model to three dimensions.

11.1 2. One dimensional model

11.1.1 2.1. Localized functions formalism

As developed in I, our method stems from the following two points:

- (a) The system is described in terms of electron wavefunctions localized on atomic sites (analogous to Wannier functions). The decomposition of the system in three parts-metal M, insulator I, metal M'-results immediately from this picture.
- (b) The current can be calculated directly at finite voltage with the help of Keldysh's generalized perturbation theory of nonequilibrium processes (1965). Since we neglect here all many body interactions, the Hamiltonian of our one dimensional system reads

$$H = \sum_{ij} T_{ij} c_i^+ c_j \quad (1)$$

c_i^+ (c_i) creates (annihilates) an electron in the state associated with the localized function Φ_i centered on site i . For the sake of simplicity, we have assumed in I that the contribution of only one energy band has to be retained in the energy range of interest (this assumption is unessential and may be easily released). We then divide the Hamiltonian into five parts

$$H = H_M + H_{M'} + H_I + H_c + V_{\text{imp}} \quad (2)$$

where

^{12†} Laboratoire Associé au Centre National de la Recherche Scientifique.

$$H_M = \sum_{\lambda, \mu} T_{\lambda\mu} c_{\lambda}^{\dagger} c_{\mu} \quad (2a)$$

$$H_{M'} = \sum_{\lambda', \mu'} T_{\lambda'\mu'} c_{\lambda'}^{\dagger} c_{\mu'} \quad (2b)$$

$$H_I = \sum_{pq} T_{pq} c_p^{\dagger} c_q \quad (2c)$$

are the Hamiltonians of the isolated electrodes and barrier (in the absence of the impurity potential) including surface effects. The unprimed (primed) greek indices label the sites of M (M'), while the roman indices ($p, q \dots$) refer to the sites in the barrier.

H_c is the transfer Hamiltonian which couples any two sites belonging to two different subparts of the system (eg M and I). The matrix elements T_{lm} of H_c between two pseudo Wannier functions decrease rapidly with $|l - m|$ (see I), so that H_c can be reasonably approximated by a first neighbour interaction between M and I, and between I and M', namely

$$H_c = T (c_{\alpha}^{\dagger} c_a + c_a^{\dagger} c_{\alpha}) + T' (c_{\alpha'}^{\dagger} c_b + c_b^{\dagger} c_{\alpha'}) \quad (3)$$

where

$$T = T_{\alpha a} = T_{a\alpha} \quad T' = T_{\alpha' b} = T_{b\alpha'}$$

α is the last site of electrode M, α' the first site of M', a and b the first and last sites of the barrier I (see figure 2).

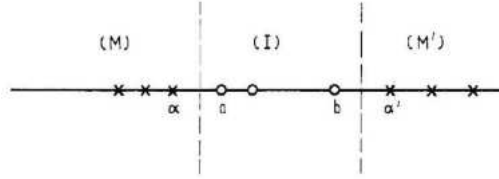


Figure 2

Finally, the fifth term V_{imp} of H is the impurity potential. This potential may have matrix elements between pseudo Wannier functions of the barrier and of the electrodes, or between two functions of one electrode, which should be included in H_c , H_M and $H_{M'}$. Because of the short range character of the metallic screening inside the electrodes, these matrix elements are certainly much smaller than the ones between two functions of the insulator. Therefore we shall neglect them and assume that V_{imp} does not extend out of the barrier. That is we neglect the renormalization of T and T' , and the changes in the local density of states (ie the Friedel oscillations) induced by the tails of the impurity potential in the metallic regions. This simplification is probably not valid for an impurity located very close to an electrode-barrier interface. However, we shall ignore this case, since the actual structure of the interface is very poorly controlled experimentally, and probably very dirty on the scale of one or two atomic layers. Finally, although our formal expression for the tunneling current is valid for any potential V_{imp} not extending out of the insulator, we will discuss its features on the simplified case where V_{imp} is localized on a single site, say site O. This approximation is by no means crucial, since the physical aspects of the impurity effect will be shown below not to depend on the detailed shape of the potential. We thus write

$$V_{\text{imp}} = v c_O^{\dagger} c_O \quad (4)$$

where v is negative, and large enough to allow for the existence of a bound state around site O.

As shown in I, in our discrete representation, and with the assumption (3) of nearest neighbour coupling between M(M') and I, the current operator reads

$$J = \frac{e}{i\hbar} T (c_a^+ c_\alpha - c_\alpha^+ c_a) \quad (5)$$

and we are thus led to compute $\langle c_a^+ c_\alpha \rangle$ and $\langle c_\alpha^+ c_a \rangle$ in order to obtain the current $\langle J \rangle$. We shall take as zeroth order system the nonconnected electrodes and barrier in the absence of the impurity, but in the presence of the bias (see I for a discussion of this procedure). It is to be noticed that the perturbing Hamiltonian

$$H' = H_c + V_{\text{imp}} \quad (6)$$

is by no means small compared to $H_M, H_{M'}$ and H_I , so that it is necessary to extend the perturbation calculation to all orders in H' .

Moreover, we want to compute the current at finite bias, that is in an essentially nonequilibrium situation. This can be dealt with in the frame of the technique of Keldysh (1965); we introduce Green matrices $\mathbf{G}$ instead of the usual Green functions

$$\mathbf{G}_{ij} = \begin{pmatrix} G_{ij}^c & G_{ij}^+ \\ G_{ij}^- & \tilde{G}_{ij}^c \end{pmatrix} \quad (7)$$

where

$$\begin{aligned} G_{ij}^c(t, t') &= -i \langle T \{ c_i(t) c_j^+(t') \} \rangle \\ \tilde{G}_{ij}^c(t, t') &= -i \langle \tilde{T} \{ c_i(t) c_j^+(t') \} \rangle \\ G_{ij}^+(t, t') &= i \langle c_j^+(t') c_i(t) \rangle \\ G_{ij}^-(t, t') &= -i \langle c_i(t) c_j^+(t') \rangle \end{aligned} \quad (8)$$

(T and $\tilde{T}$ are the chronological and antichronological operators). G^c and $\tilde{G}^c$ are the conventional causal and anticausal functions.

One must also introduce a selfenergy matrix

$$\Sigma_{ij} = \begin{pmatrix} \Sigma_{ij}^c & \Sigma_{ij}^+ \\ \Sigma_{ij}^- & \tilde{\Sigma}_{ij}^c \end{pmatrix} \quad (9)$$

The generalized Dyson equation then reads

$$\mathbf{G}_{ij} = \mathbf{g}_{ij} + \sum_{kl} \mathbf{g}_{ik} \Sigma_{kl} \mathbf{G}_{lj} \quad (10)$$

where $\mathbf{g}$ is the Green matrix of the unperturbed system. Finally, it is convenient to introduce the advanced and retarded functions (see I) †

$$\begin{aligned} G_{ij}^r &= (G^c - G^+)_{ij} = (-\tilde{G}^c + g^-)_{ij} \\ G_{ij}^a &= (G^c - G^-)_{ij} = (-\tilde{G}^c + g^+)_{ij} \\ \Sigma_{ij}^r &= (\Sigma^c - \Sigma^+)_{ij} = (-\tilde{\Sigma}^c + \Sigma^-)_{ij} \\ \Sigma_{ij}^a &= (\Sigma^c - \Sigma^-)_{ij} = (-\tilde{\Sigma}^c + \Sigma^+)_{ij} \end{aligned} \quad (11)$$

† We have rectified here a slight error of Keldysh' original paper and of I. In both cases + and superscripts were exchanged, and this resulted in a sign error in the final expression of the current. Here we take the usual definition of the advanced and retarded Green function, (in contradistinction to Keldysh and to I) and modify the definition of Σ accordingly (see equation 9). We use the same notation as in II.

The Dyson equations for $G^a, G^r, G^\pm$ are easily obtained from equation (10):

$$G_{ij}^\alpha = g_{ij}^\alpha + \sum_{kl} g_{ik}^\alpha \Sigma_{kl}^\alpha G_{lj}^\alpha \quad (\alpha = a \text{ or } r) \quad (12a)$$

$$G_{ij}^\pm = \sum_{kl} \left(\delta_{ik} + \sum_m G_{im}^r \Sigma_{mk}^r \right) g_{kl}^\pm \left(\delta_{lj} + \sum_m \Sigma_{lm}^a G_{mj}^a \right) + \sum_{kl} G_{ik}^r \Sigma_{kl}^\pm G_{lj}^a \quad (12b)$$

Note that, while G^a and G^r are related to the dynamical properties of the system, $G^\pm$ are linked with the occupation of the system states. It is clear that the current is given by

$$\begin{aligned} \langle J \rangle &= \frac{eT}{\hbar} \{ G_{ax}^+ (t, t_+) - G_{za}^+ (t, t_+) \} \\ &= \frac{eT}{\hbar} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \{ G_{a\alpha}^+(\omega) - G_{\alpha a}^+(\omega) \} \end{aligned} \quad (13)$$

11.1.2 2.2. Formal expression of the current

Since the perturbation H' is instantaneous, $\Sigma^\pm = 0, \Sigma^a = \Sigma^r = \Sigma$ (see Appendix of I). For our particular choices of H_c and V_{imp} (equations (3) and (4)), we get

$$\Sigma_{pq} = T (\delta_{px} \dot{\delta}_{qa} + \delta_{pa} \delta_{q\alpha}) + T' (\dot{\delta}_{p\alpha'} \delta_{bq} + \delta_{p \text{ b}} \delta_{q\alpha'}) + v \delta_{pO} \delta_{qO} \quad (14)$$

Using equation (12b), one verifies easily that (see II)

$$G_{a\alpha}^+ - G_{\alpha a}^+ = T (g_{a\alpha}^+ G_{aa}^- - g_{\alpha\bar{a}}^- G_{aa}^+) \quad (15)$$

We now calculate $G_{aa}^\pm$ with the help of the Dyson equation (12b). Let us recall that to zeroth order, there is no connection between the electrodes and the barrier, so that $g_{i\bar{j}}^\pm = g_{i'j}^\pm = 0$. Moreover, $g_{pq}^\pm = 0$, since to zeroth order, there is no density of states in the barrier in the energy range of interest. Equation (12b) then reduces to

$$G_{aa}^\pm = G_a^r T g_{a\alpha}^\pm T G_{aa}^a + g_{ab}^r T' g_{\alpha'x'}^\pm T' G_{ba}^a \quad (16)$$

Inserting equation (16) into (15), we obtain for the current at bias V

$$\langle J \rangle = \frac{(2\pi)^2 e T^2 T'^2}{\hbar} \int_{\mu'}^{\mu' + eV} \frac{d\omega}{2\pi} |G_{ab}^r(\omega)|^2 \rho_\alpha(\omega) \rho_{\alpha'}(\omega) \quad (17)$$

where μ' is the Fermi energy of the right electrode M' and $\mu = \mu' + eV$ the Fermi energy of $\mathbf{M}$. $\rho_\alpha(\omega)$ ($\rho_{\alpha'}(\omega)$) is the local density of states of the isolated $\mathbf{M}$ ($\mathbf{M}'$) electrode on site α (α') and is related to $g_{a\alpha}^\pm(\omega)$ at zero temperature by

$$\begin{aligned} g_{a\alpha}^+(\omega) &= 2i\pi\theta(\mu - \omega)\rho_\alpha(\omega) \\ g_{a\alpha}^-(\omega) &= -2i\pi\theta(\omega - \mu)\rho_\alpha(\omega) \end{aligned} \quad (18)$$

where θ is the ordinary step function.

Equation (17) for the current is formally identical to the expression obtained in I for a pure junction. The effect of the impurity potential therefore only appears through a renormalization of the transmission coefficient (since G_{ab} in (17) is the propagator of an electron through the barrier in the presence of both the transfer coupling H_c and the impurity potential).

It is of some interest to compare this result with the effect of the electron-phonon interaction (see II). The phonon effect can be attributed to two kinds of processes: virtual transitions (this is also said to be a change of the 'elastic current'), which renormalize the transmission coefficient, and real transitions, that is tunneling of an electron with simultaneous emission or absorption of a phonon. which give rise to the additional so-called 'inelastic current'. It is often said that the electron-phonon interaction opens a new channel for tunneling. To our understanding, this means that, due to this interaction, new tunneling processes are made possible, which correspond to real phonon emission or absorption.

In our opinion, on the contrary, a structureless impurity cannot be said to open a new channel, since it cannot induce a real transition. This appears mathematically in the form of equation (17), and is due to two facts:

- (a) The selfenergy has no $\Sigma^\pm$ component, which corresponds to the instantaneous character of the impurity potential $\dagger$. On the contrary, the phonon-induced selfenergy has $\Sigma^\pm \neq 0$, which means that phonons create a retarded effective electron-electron interaction.
- (b) We have chosen to include V_{imp} into the perturbing Hamiltonian. Therefore, to zeroth order, there is no density of states in the barrier. This is the reason for the very simple form of expression (16).

It can of course be argued that our definition of the perturbation is arbitrary, and that V_{imp} could just as well be included into the zeroth order system, in which case simplification (b) would fall (but (a) still remains true). It is of some physical interest to analyse the structure of the expression for the current which results from this alternative choice of the zeroth order system.

Let us call $\tilde{\mathbf{g}}_{ij}(\omega)$ the Green matrix of the new unperturbed junction (isolated M and M', and isolated impure insulator). The selfenergy now derives only from the transfer coupling H_c , so that the last term of equation (14) disappears. The current is then evaluated in a straightforward way with the help of equations (12b) and (15). The result one obtains for

$\langle J \rangle$ is the sum of two terms: the first one is identical with expression (17), the second one is a combination of terms which all contain products of the type

$$(\tilde{g}^r)^{-1}_{pp'} \tilde{g}^\pm_{p'p''} (\tilde{g}^a)^{-1}_{p''q}$$

with p, p', p'', q all belonging to the insulator. This term therefore results directly from the presence of a bound state in the isolated barrier at energy E. Formally, $g^\pm$ is proportional to $\delta(\omega - H_I - V_{imp})$, that is in the range of energy of interest

$$g^\pm_{p'p''}(\omega) = a^\pm_{p'p''} \dot{\delta}(\omega - E) \quad (19)$$

In the same way, $(\tilde{g}^{a,r})^{-1} = \omega - H_I - V_{imp} \pm i\delta$. Therefore,

$$\sum_{p'p''} (\tilde{g}^r)^{-1}_{pp'} \tilde{g}^\pm_{p'p''} (\tilde{g}^a)^{-1}_{p''q} \propto \sum_{p'p''} (\omega - H_I - V_{imp})_{pp'} \delta(\omega - H_I - V_{imp})_{p'p''} (\omega - H_I - V_{imp})_{p''q} \equiv 0 \quad (20)$$

¹² $\dagger$ This is true only for structureless impurities, and would not hold for instance in the case of paramagnetic centers.

and $\langle J \rangle$ effectively reduces to (17).

It is clear from this calculation that the absence of an inelastic current term comes from the δ function shape of the barrier density of states, that is from the fact that the impurity state is not connected with any continuum. This again is equivalent to saying that no real transition can take place, or, on the usual terminology of tunneling, that no new channel is opened by the presence of a nonmagnetic impurity.

11.1.3 2.3. The transmission coefficient

We shall now examine in more detail the deformation of the tunneling characteristic induced by the acceptor center, in the particular case of a potential localized on site O (equation (4)). Moreover, we shall neglect the effects of the deformation of the barrier due to the presence of a finite bias, which gives rise to the so-called 'background current'. This current varies with the applied bias much more slowly than the 'impurity current' and may be therefore neglected.

However it is clear that the presence of a finite electric field in the barrier has the effect of shifting the bound state energy. Roughly speaking, this shift is of the order of the voltage drop at the impurity site, that is. it is proportional to the applied bias.

For instance, let us take the Fermi level μ' of the right electrode as zero energy. Thus, if E is the bound state energy at zero bias, a finite bias yields a (eV) x/d shift where d is the barrier thickness and x the distance between the impurity site and the right electrode.

As we shall see later, the main effect of the impurity on the tunneling characteristic arises when the effective bound state energy $E + (eV)x/d$ crosses the Fermi level of one electrode. Assume for instance E and $\mu - \mu' = eV$ to have the same sign. Then the concerned electrode is the left one and the difference between the Fermi level of the left electrode and the effective bound state energy (the important parameter as we shall see) is

$$eV - \left(E + (eV)\frac{x}{d} \right) = eV \left(1 - \frac{x}{d} \right) - E$$

Thus the shift of the bound state energy results in practice in a magnification by a factor $(1 - x/d)^{-1}$ of the energy scale: the resonance occurs at bias $E/\{1 - (x/d)\}$ and not E . Since we will not consider impurities too close to the barrier-electrode interfaces, this magnifying factor is not very large (typically of the order of a few units) and does not yield dramatic effect. We shall omit this factor in the following calculations, since it may be included in a straightforward manner in the final result. Note that, if E and eV were of opposite sign, the magnifying factor would be $(x/d)^{-1}$ instead of $(1 - x/d)^{-1}$. We shall come back to this asymmetry at the end of the paper.

Finally, we assume the metallic densities of states $\rho_\alpha(\omega)$ and $\rho_{\alpha'}(\omega)$ to be constant on the range of variation of the impurity current, which will be seen below to be much smaller than any bandwidth or Fermi energy. We then obtain from equation (17)

$$\frac{d\langle J \rangle}{dV} = \frac{2\pi e T^2 T'^2}{\hbar} \rho_\alpha \rho_{\alpha'} |G_{ab}^r(\mu' + eV)|^2 \quad (21)$$

and all the energy dependence of $d\langle J \rangle/dV$ is contained in the 'transmission factor' $|G_{ab}^r|^2$. This factor appears in equation (21) in a somewhat nonsymmetric way, since it has to be taken at the energy $\mu' + eV$ of the Fermi level of the left electrode. This comes from our approximation of neglecting the electric field in the insulator (ie the background current term).

Using equations (12b) and (14), we obtain after some simple algebra

$$G_{ab}^r = G_{ab}^{(0)r} \Lambda_{\text{imp}}^r \quad (22)$$

where

$$G_{ab}^{(0)r} = \frac{g_{ab}}{(1 - T^2 g_{aa} g_{\alpha\alpha}^r)(1 - T'^2 g_{bb} g_{\alpha'\alpha'}^r) - T^2 T'^2 g_{ab} g_{ba} g_{\alpha\alpha}^r g_{\alpha'\alpha'}^r} \quad (23)$$

is the full retarded propagator across the barrier for the pure junction. The denominator in the right hand side of equation (23) expresses the possibility for an electron to make any number of excursions inside M or M' (see I).

The impurity effect is entirely contained in the factor

$$\Lambda_{\text{imp}}^r = \frac{1 + v g_{ab}^{-1} (g_{ao} g_{ob} - g_{oo} g_{ab})}{1 - v (g_{oo} + \lambda^r)} \quad (24)$$

where $\dagger$

$$\lambda^r = \frac{T^2 (g_{ao})^2 g_{\alpha\alpha}^r}{1 - T^2 g_{\alpha\alpha}^r g_{aa}} + \frac{T'^2 (g_{bo})^2 g_{\alpha'\alpha'}^r}{1 - T'^2 g_{\alpha'\alpha'}^r g_{bb}} \quad (25)$$

Note that we have dropped the r index on the g_{pq} (with p and q belonging to the barrier), since $\text{Im } g_{pq}^r \propto g_{pq}^\pm = 0$, and $g^r = g^a$ is real.

Finally, equation (17) becomes

$$\frac{d\langle J \rangle}{dV} = \left(\frac{d\langle J \rangle}{dV} \right)^{(0)} |\Lambda_{\text{imp}}^r|^2 \quad (26)$$

Let us now come back to expression (24): if the λ^r term were absent from the denominator, Λ_{imp}^r would have a pole for $1 - v g_{oo} = 0$. Since we assume that the impurity potential has a bound state at energy E , this equation has a real solution at $\omega = E$. The effect of λ^r is to shift the position of this pole. We have assumed that the impurity is not very close to any interface, so that $|R_O - R_a|$ and $|R_O - R_b|$ are nonzero. Since g_{pq} decreases exponentially with $|p - q|$ (the electron wavefunctions are evanescent in the insulator), $\lambda^r \ll g_{oo}$, so that the position of the shifted pole is roughly given by

$$v^{-1} - g_{oo}(\omega) = \lambda^r(E)$$

that is

$$\omega \simeq E - \lambda^r(E) \left(\frac{dg_{oo}(\omega)}{d\omega} \Big|_{\omega=E} \right)^{-1} = \omega_0 \quad (27)$$

Since $g_{\alpha\alpha}^r$ and $g_{\alpha'\alpha'}^r$ have finite imaginary parts (ie since ρ_α and $\rho_{\alpha'}$ are finite), $\text{Im } \lambda^r \neq 0$, and the main effect of λ^r is to make the pole of Λ_{imp} complex. (The real shift due to $\text{Re } \lambda^r$ can be lumped into the definition of E). In other words, when the barrier-electrodes coupling is

turned on, a connection is established between the bound state and the continuums of electron states of M and M', and a resonance occurs. The effective coupling decreases exponentially with $|R_O - R_{a,b}|$. due to the decrease of the bound state wavefunction, and so does the width of the resonance. This effect appears in i^r through the factors $(g_{Oa,b})^2$. Finally, note that both electrode-barrier contacts contribute to the broadening. If the impurity is about the middle of the barrier, the contributions of M and M' to λ^r are comparable. If the impurity is in an asymmetric position, the broadening is essentially due to the closest electrode. Finally, since the width of the resonance

^{12†} In equation (25) we have systematically neglected terms of order $(g_{ab})^2$ with respect to $(g_{aa})^2$, since their ratio is of the order of the transmission coefficient of the junction, that is it is exponentially small.

$$\Delta = \left(\text{Im } \hat{\lambda}^r \right) \left(\frac{dg_{00}(\omega)}{d\omega} \Big|_{\omega=E} \right)^{-1} \quad (28)$$

with

$$\text{Im } \hat{\lambda}^r = \pi \rho_\alpha \frac{T^2 (g_{a0})^2}{|1 - T^2 g_{\alpha\alpha}^r g_{aa}|^2} + \pi \rho_\beta \frac{T'^2 (g_{b0})^2}{|1 - T'^2 g_{\beta\beta}^r g_{bb}|^2}$$

is smaller (for $|R_0 - R_{a,b}|$ larger than about two atomic distances) than all the characteristic energies of the problem, the numerator of equation (24) is quasiconstant in the region $|\omega - E| < \Gamma$, and the peak in the conductance G of the junction due to the impurity has a Lorentzian shape. The amplitude of this peak depends on the band structure of the insulator. If we describe it in a nearest neighbour tight binding approximation (in one dimension), the numerator reduces to one, and, in the region of the resonance, the conductance becomes simple

$$G = G_0 \frac{1}{v^2 \left(\frac{dg_{00}}{d\omega} \Big|_{\omega=E} \right)^2 \left\{ (\mu' + eV - \tilde{E})^2 + \Delta^2 \right\}} \quad (29)$$

with

$$\tilde{E} = E - \left(\frac{dg_{00}}{d\omega} \Big|_{\omega=E} \right)^{-1} \text{Re } \hat{\lambda}^r$$

11.2 3. Generalization to three dimensions

As in I, we shall neglect the Moire pattern that arises from the possible difference between the periodicities of M, I and M' in the yz plane parallel to the boundaries. That is, we assume that the (yz) component $\mathbf{k}_\parallel$ of the wavevector is conserved. The transfer electrode-barrier coupling is now taken as

$$H_c = \sum_{\alpha\alpha'} T_{\alpha\alpha'} (c_\alpha^+ c_a + c_a^+ c_\alpha) + \sum_{\alpha'\alpha''} T_{\alpha'\alpha''} (c_{\alpha'}^+ c_b + c_b^+ c_{\alpha'}) \quad (30)$$

where α (α') lie in the first atomic plane of M (M'), and a(b) in the first (last) atomic plane of the insulator.

In the absence of any impurity, the junction is translationally invariant along the (yz) plane. We introduce Fourier transforms with respect to y and z , for instance

$$\begin{aligned} T(\mathbf{k}_\parallel) &= \sum_{\alpha a} T_{\alpha a} \exp \{ i \mathbf{k}_\parallel \cdot (\mathbf{R}_\alpha - \mathbf{R}_a) \} \\ G_{ij}^{(0)}(\mathbf{k}_\parallel, \omega) &= \sum_{i_\parallel j_\parallel} G_{\mathbf{R}_i, \mathbf{R}_j}^{(0)}(\omega) \exp \{ i \mathbf{k}_\parallel \cdot (\mathbf{i}_\parallel - \mathbf{j}_\parallel) \} \end{aligned} \quad (31)$$

where $G^{(0)}$ is the full propagator of the pure junction, and i and $i_\parallel$ denote respectively the x and (y, z) components of $\mathbf{R}_i$. The x component of the current then reads (see I)

$$J = \int \frac{d^2 \mathbf{k}_\parallel}{(2\pi)^2} \frac{(2\pi)^2 e}{\hbar} \frac{\{T(\mathbf{k}_\parallel)\}^2 \{T'(\mathbf{k}_\parallel)\}^2}{\int_\mu^{\mu'+eV} \frac{d\omega}{2\pi} \left| G_{ab}^{(0)r}(\mathbf{k}_\parallel, \omega) \right|^2} \rho_\alpha(\mathbf{k}_\parallel, \omega) \cdot \rho_{\alpha'}(\mathbf{k}_\parallel, \omega) \quad (32)$$

Let us now consider an impure junction with N_i impurities all located inside the barrier. As in the one dimensional case, we assume that the impurity potentials do not extend out of the insulator. The calculation of the current now proceeds exactly as in § 2, and we get

$$J = \int \frac{d^2 \mathbf{k}_{\parallel}}{(2\pi)^2} \frac{d^2 \mathbf{k}'_{\parallel}}{(2\pi)^2} \frac{(2\pi)^2 e \{T(\mathbf{k}_{\parallel})\}^2 \{T'(\mathbf{k}'_{\parallel})\}^2}{\hbar} \int_{\mu'}^{\mu'+eV} \frac{d\omega}{2\pi} \left| G_{\text{ab}}^r(\mathbf{k}_{\parallel}, \mathbf{k}'_{\parallel}, \omega) \right|^2 \cdot \rho_z(\mathbf{k}_{\parallel}, \omega) \rho_{\alpha'}(\mathbf{k}'_{\parallel}, \omega) \quad (33)$$

where

$$G_{ij}^r(\mathbf{k}_{\parallel}, \mathbf{k}'_{\parallel}, \omega) = \sum_{i_{\parallel}, j_{\parallel}} G_{\mathbf{R}_i, \mathbf{R}_j}^r(\omega) \exp \{i(\mathbf{k}_{\parallel} \cdot \mathbf{i}_{\parallel} - \mathbf{k}'_{\parallel} \cdot \mathbf{j}_{\parallel})\} \quad (34)$$

is the double (yz) Fourier transform of the full propagator of the junction, including the effect of impurities. Since the (yx) translational invariance no longer exists in the impure system, $\mathbf{k}_{\parallel}$ is no longer conserved and G depends on two independent wavevectors.

We shall now assume the impurities to be randomly distributed in the (yz) direction. In that case, once the averaging on the impurity distribution in the (yz) plane has been performed, the system recovers its translational invariance along that plane. Since $|G_{\text{ab}}^r|^2$ is the only term in equation (33) that depends on the impurity configuration, we are left with the problem of calculating

$$\langle |G_{\text{ab}}^r(\mathbf{k}_{\parallel}, \mathbf{k}'_{\parallel}, \omega)|^2 \rangle_{\text{av}}$$

We shall not make any assumption on the shape of the impurity distribution along the transverse x direction. On the one hand, it has been, up to now, impossible for technical reasons to control this distribution experimentally. On the other hand, an assumption of randomness along x would not make the problem much simpler: the broadening of the impurity bound state due to the I-M and I-M' contacts has been shown in §2 to depend on the transverse coordinate of the impurity centre, therefore transverse randomness does not restore invariance along x .

We therefore keep the transverse repartition c_p (where c_p is the impurity concentration in the p th atomic plane) as an unknown parameter. It seems that, in the present status of tunneling techniques, the impure junction experiments should be essentially aimed at getting some qualitative information on this repartition function, that is on the structure of the insulating layer.

Finally, we limit ourselves to the case of a low impurity concentration, so that we can neglect interference effects between different centres (that is we keep to first order terms in the concentration) and to local impurity potentials of type (3).

Since the advanced and retarded Green functions obey the same formal equations as in the equilibrium case, we can write

$$G_{\mathbf{R}_a, \mathbf{R}_b}^r(\omega) = G_{\mathbf{R}_a, \mathbf{R}_b}^{(0)}(\omega) + \sum_{\mathbf{R}_p} G_{\mathbf{R}_a, \mathbf{R}_p}^{(0)r}(\omega) \mathcal{T}_p^r G_{\mathbf{R}_p, \mathbf{R}_b}^{(0)r}(\omega) \quad (35)$$

where $G^{(0)}$ is the full propagator of the pure junction, and the retarded $\mathcal{T}$ matrix of an impurity located at site $\mathbf{R}_p$ (in the presence of the bias) is given by

$$\mathcal{T}_{\mathbf{R}_i, \mathbf{R}_j}^r = \mathcal{T}_p^r \delta_{\mathbf{R}_i, \mathbf{R}_j} \delta_{\mathbf{R}_j, \mathbf{R}_p} \quad (36)$$

and $\mathcal{T}_p^r = v / \{1 - v G_{\mathbf{R}_p}^{(0)\mathbf{R}_p}(\omega)\}$ does not depend on $\mathbf{p}_{\parallel}$.

Fourier transforming equation (35) in the yz plane, and taking advantage of the translational invariance of the pure junction, we obtain

$$G_{ab}^r(\mathbf{k}_{\parallel}, \mathbf{k}'_{\parallel}, \omega) = G_{ab}^{(0)r}(\mathbf{k}_{\parallel}) \delta_{\mathbf{k}, \mathbf{k}'} + \sum_{\mathbf{R}_p} \exp \{i(\mathbf{k}_{\parallel} - \mathbf{k}'_{\parallel}) \cdot \mathbf{p}_{\parallel}\} G_{ap}^{(0)r}(\mathbf{k}_{\parallel}) \mathcal{T}_p^r G_{pb}^{(0)r}(\mathbf{k}'_{\parallel}) \quad (37)$$

and, to first order in the concentration c_p in the p th atomic plane

$$\begin{aligned} \langle |G_{ab}^r(\mathbf{k}_{\parallel}, \mathbf{k}'_{\parallel}, \omega)|^2 \rangle_{av} &= \left| G_{ab}^{(0)r}(\mathbf{k}_{\parallel}, \omega) \right|^2 \delta_{\mathbf{k}_1 \cdot \mathbf{k}'} + \delta_{\mathbf{k}_{\parallel}, \mathbf{k}'_{\parallel}} \left\{ G_{ab}^{(a)}(\mathbf{k}_{\parallel}) \sum_p c_p G_{ap}^{(0)a}(\mathbf{k}_{\parallel}, \omega) \right. \\ &\quad \times \mathcal{T}_p^a G_{pb}^{(0)a}(\mathbf{k}_{\parallel}, \omega) + \text{h.c.} \left. \right\} + \sum_p c_p \left| \mathcal{T}_p^r G_{ap}^{(0)r}(\mathbf{k}_{\parallel}, \omega) G_{pb}^{(0)r}(\mathbf{k}'_{\parallel}, \omega) \right|^2 \end{aligned} \quad (38)$$

The first term of the rhs of equation (38) is simply the transmission coefficient of the pure junction, and the impurity effect appears through the $\mathcal{T}$ matrices $\mathcal{T}_p$. Especially, $\mathcal{T}_p$ contains the broadening effect due to metal-insulator transfers that has been described in §2. Since we have seen that the width of the resulting resonance is always very small compared to the other characteristic energies of the problem, in the vicinity of the (shifted) bound state energy, $\mathcal{T}_p$ can be approximated (with notations analogous to those of § 2) by

$$\begin{aligned} \mathcal{T}_p^r(\omega) &\simeq \frac{\alpha_p}{\omega - E_p + i\Delta_p} \\ \alpha_p &\sim \left[\frac{dg_{pp}(\omega)}{d\omega} \Big|_{\omega=E_p} \right]^{-1} \sim \epsilon^2 \end{aligned} \quad (39)$$

where ϵ is a characteristic energy of the order of the insulator gap or typical bandwidth, therefore is in the electron-volt range.

On the other hand, in order of magnitude,

$$G_{ap} G_{pb} \sim G_{ab} G_{pp}$$

so that, in the resonance region, the third term of equation (38) is larger than the second one by a factor of the order of

$$\mathcal{T}_p G_{pp} \sim \frac{\epsilon}{\Delta_p} \gg 1$$

so that the latter can be neglected. We thus obtain

$$\langle J \rangle_{av} = J^{(0)} + \frac{(2\pi)^2 e}{\hbar} \int_{\mu'}^{\mu' + eV} \frac{d\omega}{2\pi} \sum_p c_p |\mathcal{T}_p^T(\omega)|^2 U_{ap}(\omega) U'_{pb}(\omega) \quad (40)$$

with

$$U_{ap}(\omega) = \int \frac{d^2 \mathbf{k}_{\parallel}}{(2\pi)^2} \{T(\mathbf{k}_{\parallel})\}^2 \left| G_{ap}^{(0)r}(\mathbf{k}_{\parallel}, \omega) \right|^2 \rho_a(\mathbf{k}_{\parallel}, \omega)$$

and a similar definition (with $T \rightarrow T'$, $\rho_{\alpha} \rightarrow \rho_{\alpha'}$) for U'_{pb} . Since $J^{(0)}$, U and U' vary slowly on the scale of a Δ_p , the structure in the tunneling characteristic comes essentially from the $\mathcal{T}$ matrix factor in equation (40), and, for energies close to the bound state level E , the conductance may be written

$$\frac{d\langle J \rangle_{av}}{dV} = \sum_p c_p |\mathcal{T}_p^r(\mu' + eV)|^2 A_p \quad (41)$$

where the A_p are roughly constant. Moreover, since $\mu' + eV$ lies in the gap of the insulator, the product $U_{ap}U'_{pb}$ does not show any important oscillations (these are important only in the metallic state, due to the rapidly oscillating character of the electron wavefunctions) and the p dependence of A_p can safely be neglected.

If the impurities were located in one single atomic plane, say O , $d\langle J \rangle_{av}/dV$ would be a Lorentzian curve:

$$\frac{d\langle J \rangle_{av}}{dV} = \alpha c \frac{\epsilon^2}{(\omega - \tilde{E}_0)^2 + \Delta_0^2}$$

The shift $(\tilde{E}_0 - E)$ and the width Δ_0 of the resonance arise from the same physical effect as in the one dimensional case: an electron goes from the impurity site to either electrode, where it experiences the finite density of states of the metallic continuum, then comes back to the impurity. Of course, in the three dimensional situation, the electron transfer process can take place at any point of the interface. However, as soon as the impurity is not very close to a boundary, due to the exponential decrease of the insulator propagators, paths like OB (see figure 3) contribute much less than the shortest paths like OA , and the shift and width are essentially the same as in the one dimensional problem.

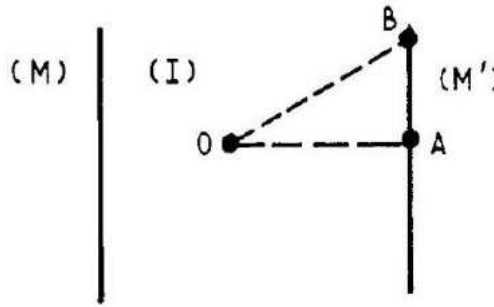


Figure 3

In actual cases of impure junctions, the conductance, as given by equation (41) is a superposition of Lorentzian curves of different centres and widths, corresponding to the contributions of the different atomic planes. This superposition has the effect of broadening the structure; but, in fact these Lorentzian curves are all quite sharp, except for the impurities that are very close to an interface. Thus the resulting broadening is negligible.

Nevertheless, there is in the system a second source of broadening: we have noted in §2 that, due to the presence of a finite electric field in the barrier, each bound state energy is shifted by an amount of the order of the voltage drop at the corresponding impurity site. Therefore it is to be expected that, if there is a uniform repartition of impurities across the depth of the barrier, the structure should be almost completely blurred out, and observation of the impurity effect should be possible only if the acceptor centres are located in a few neighbouring atomic planes in the depth of the insulator.

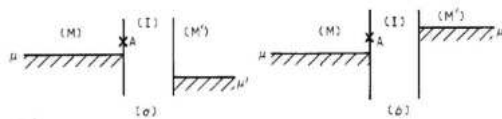


Figure 4. In (b), the energy of the impurity level A belongs to the energy range of tunnelling, which is never the case in (a).

It is of some interest to note that impurities lying near an interface yield a contribution very asymmetric with respect to the bias. For example, in direct bias (see figure 4(a)), impurity A

never provides an increase in the tunneling current. But in reverse bias, there is a contribution from impurity A (see figure 4(b)). On the other hand, an impurity lying in the middle of the barrier provides a nearly symmetric contribution.

So it seems that the impurity effect can be of use mainly as a tool in the study of the insulating layer of a MIM junction, since its presence or absence gives indications on the location of a given impurity species.

12 Tunneling current IV: Electron-phonon interaction effects by Caroli et. al.

by C Caroli, R Combescot, P Nozieres And D Saint-James

Electron-phonon interaction effects in tunnelling junctions are considered in the framework of the new formalism proposed in a preceding paper. This formalism is simplified and then used to study electron-phonon interaction in metal-insulator-metal junctions. The interactions inside the metallic electrodes are shown to yield a complicated mixture of dispersive and dissipative lineshapes in the second derivative d^2I/dV^2 . Such a structure originates in a destructive interfering, which also leads to a strong reduction in the magnitude of the effect. On the other hand, the interactions in the insulator yield dissipative lineshapes, due to the opening of a new inelastic tunnelling channel. Such a contribution is reinforced by the absence of destructive interference inside the insulator. It is proposed that the observed structures arising from electrode phonons are due predominantly to the penetration of metallic phonons inside the insulator, rather than to selfenergies localized in the metal itself.

12.0.1 1. Introduction

In a recent article (Caroli et al 1971, hereafter referred to as I), a new formulation of tunnelling phenomena through insulating barriers was proposed. This approach, based on the nonequilibrium perturbation formalism of Keldysh, does not rely on the transfer hamiltonian approximation, and provides an exact expression of the tunnelling current. The purpose of this paper is to present a somewhat simplified version of the theory and to discuss the influence of the electron-phonon interaction, both in the electrodes and in the barrier. The interplay of inelastic and selfenergy effects is thus clearly displayed, in a way that is far more compact and simple than usual.

We use the same model as in I. The momentum parallel to the barrier is assumed to be conserved: for a given $k_{\parallel}$, the problem is one dimensional. We describe the propagation of electrons in terms of 'pseudo-Wannier functions' localized on each atomic site, and we assume that only one energy band is involved. These assumptions are not essential, and can be released in order to treat, for instance, the Schottky barrier; this will be done in future publications: here, we stick for simplicity to a model as elementary as possible. The one electron hamiltonian is thus

$$H_1 = \sum_{ij} T_{ij} c_i^+ c_j \quad (1)$$

where c_i^+ , c_i are the electron creation and destruction operators on site i . In addition, we assume that the electrons are coupled to lattice vibrations (in fact to any collective mode in the system). If we disregard the polarization of phonons, the coupling hamiltonian is

¹²† Laboratoire Associé au Centre National de la Recherche Scientifique.

$$H_2 = \sum_{ijl} V_{ijl} c_i^\dagger c_j \phi_l + \text{cc} \quad (2)$$

where ϕ_l is the field operator of the phonon on site l . These lattice modes include the phonons in the electrodes and in the barrier (possibly perturbed by surface effects), as well as localized modes around impurities (OH bonds in the insulator, etc). The matrix element V_{ijl} is certainly very short ranged in the distance (i, j) ; on the other hand, the range (i, l) , although small in the electrodes when all interactions are screened, may be quite large in the insulator (especially if the vibration carries an electric dipole). Nevertheless, we assume V_{ikl} to be local, and we replace equation 2 by

$$H_2 = \sum_i V_i c_i^\dagger c_i \phi_i + \text{cc}$$

Again, this is done only to simplify the exposition, and is by no means essential. We wish to calculate the current through the barrier when a DC bias is applied. For that purpose, we divide the system into two parts: the left half whose last site is labelled α , the right half whose first site is labelled a . For simplicity (Caroli et al 1971), we assume that the only term of equation 1 which couples the two halves involves nearest neighbours: we write it as

$$H_c = T (c_x^\dagger c_a + c_a^\dagger c_x) \quad (3)$$

(T being real). The current across the barrier is then described by the operator

$$J = \frac{ie}{\hbar} T (c_x^\dagger c_a - c_a^\dagger c_x) \quad (4)$$

In order to calculate $\langle J \rangle$, we bias the junction at time $t = -\infty$, and we establish H_c adiabatically. The final result does not depend on where the partition has been made if the calculation is carried out to all orders. The actual choice is then a matter of convenience: for instance, the partition will be made at the contact between the left electrode and the insulator in a MIM junction, or between the metal and the semiconductor in Schottky barriers.

12.0.2 2. Perturbation formalism

In view of the irreversible character of the tunnelling current, the usual perturbation theory does not apply, and we must rely on the more general formulation of Keldysh (the principle of which was laid down by Schwinger 1961). In its original form, this formalism contains some errors in sign; moreover the definition of the various propagators differs from the one commonly used. In the present paper, we shall return to the usual notation, which is slightly different from that found in Keldysh (1965) and Caroli et al (1971).

We introduce two kinds of propagators:

(i) Spectral densities $G^\pm$, defined as

$$\begin{aligned} G_{ij}^+(t, t') &= i \langle c_j^\dagger(t') c_i(t) \rangle \\ G_{ij}^-(t, t') &= -i \langle c_i(t) c_j^\dagger(t') \rangle \end{aligned}$$

which carry the information on the occupation of particle and hole states

(ii) Retarded and advanced functions $\dagger$, G^r and G^a

$$\begin{aligned} G_{ij}^r(t, t') &= -i \theta(t - t') \langle [c_i(t), c_j^\dagger(t')]_+ \rangle \\ G_{ij}^a(t, t') &= i \theta(t' - t) \langle [c_i(t), c_j^\dagger(t')]_+ \rangle \end{aligned}$$

the Fourier transforms of which are analytic respectively above and below the real energy axis. G^r and G^a contain only information on the distribution of available states, with no reference to their occupation. The phonon propagators are defined in similar fashion, for example

$$\begin{aligned} D_{ij}^+(t, t') &= -i \langle \phi_j(t') \phi_i(t) \rangle = D_{ji}^-(t', t) \\ D_{ij}^r(t, t') &= -i \theta(t - t') \langle [\phi_i(t), \phi_j(t')] - \rangle \\ D_{ij}^a(t, t') &= i \theta(t' - t) \langle [\phi_i(t), \phi_j(t')] \dots \rangle \end{aligned}$$

As shown by Keldysh, one can construct a perturbation 'expansion' of these quantities, which remains valid out of equilibrium. We refer to his paper for a detailed discussion. Here, we only quote a few simple recipes that we shall use again and again:

(i) G^r (G^a) obeys a separate Dyson equation, which involves only an appropriately defined retarded selfenergy Σ^r (Σ^a). Thus

$$G^r = G_0^r + g_0^r \Sigma^r G^r$$

(ii) On the other hand, there is no separate Dyson equation for G^+ and G^- , which remain coupled to G^r and G^a . As shown by Keldysh, the following relation holds

$$G^\pm = (1 + g^r \Sigma^r) G_0^\pm (1 + \Sigma^a G^a) + g^r \Sigma^\pm G^a \quad (5a)$$

where $\Sigma^\pm$ is an off diagonal component of the selfenergy (which will be specified later [‡]). The result (5a) may be cast in the form of a very simple recipe: in order to obtain $G^\pm$, one picks up successively any factor (either G_0 or Σ) in the usual Dyson expansion $G_0 + g_0 \Sigma G_0 + \dots$ and one takes its $\pm$ component. All the factors that lie on the left of that one are taken as retarded, all those on the right being advanced. It follows that the $\pm$ component of a product AB is:

$$(AB)^\pm = A^t B^a + A^r B^\pm \quad (5b)$$

In contrast, all the factors that enter in the Dyson expansion of G^r (G^a) are retarded (advanced), in such a way that

$$(AB)^r = A^r B^r$$

This recipe can be applied automatically, and will be widely used throughout this paper. It now remains to define Σ^r , Σ^a and $\Sigma^\pm$. If the perturbation is a one body time independent potential T_{ij} , the corresponding selfenergies are (see Caroli et al)

$$\begin{aligned} {}^T \Sigma_{ij}^r &= {}^r \Sigma_{ij}^a = T_{ij} \\ {}^T \Sigma_{ij}^+ &= {}^T \Sigma_{ij}^- = 0 \end{aligned}$$

An explicit calculation of the selfenergy ${}^{\text{ph}} \Sigma$ arising from the electron-phonon interaction would require the Keldysh algebra in its full form. Here, we shall restrict ourselves to the Migdal approximation, in which only the diagrams of second order in the electron-phonon coupling are retained. It then follows that

^{12†} In the notation of Keldysh (and of Caroli et al) the definitions of G^r and G'' are permuted.

[‡] Our definition of $\Sigma^\pm$ is opposite to that of Keldysh: more precisely, the component of $\Sigma(t, t')$ going from one branch of the time contour to the other that is introduced in his paper is equal to $-\Sigma^\pm(t, t')$.

$$\begin{aligned}
{}^{\text{ph}}\Sigma_{ij}^{\pm}(\omega) &= iV_i V_j \int \frac{d\omega'}{2\pi} D_{ij}^{\pm}(\omega - \omega') G_{ij}^{\pm}(\omega') \\
{}^{\text{ph}}\Sigma_{ij}^{\text{r}}(\omega) &= iV_i V_j \int \frac{d\omega'}{2\pi} \left\{ D_{ij}^{\text{r}}(\omega - \omega') G_{ij}^{-}(\omega') + D_{ij}^{+}(\omega - \omega') G_{ij}^{\text{r}}(\omega') \right\}
\end{aligned} \tag{6}$$

(We note that while ${}^{\text{ph}}\Sigma^{\pm}$ is a convolution of D^{+} and G^{+} , ${}^{\text{ph}}\Sigma^{\text{r}}$ is not a convolution of D^{r} and G^{r}). The demonstration of all these results may be found in Keldysh's paper: here, we only quote them, as they provide the algebra on which our theory is based.

From equation 4, we see that the average current through the junction is

$$\langle J \rangle = \frac{e}{\hbar} \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} T \left\{ G_{az}^{+}(\omega) - G_{aa}^{+\gamma(\omega)} \right\} \tag{7}$$

Our calculation is thus reduced to that of G^{+} . It is interesting to note that time reversal symmetry implies that all the propagators are symmetric: $G_{ij} = G_{ji}$ (reflection invariance is not required). In the absence of bias, the system is time reversal invariant, and the current (7) vanishes as expected. On the other hand, the applied bias destroys the time reversal invariance (because of the irreversibility): the G are no longer symmetric, which is shown in the existence of a nonzero current given by equation 7.

We want to treat H_c and H_2 as perturbations. The complete propagator G will thus be obtained by inserting on the zeroth order g_0 any number of either the phonon selfenergy ${}^{\text{ph}}\Sigma$ or the propagator transfer matrix element T . We shall denote by g the propagator obtained by inserting only phonon selfenergies on the g_0 : the full G is then achieved by inserting transfer elements T on the g . We note that g is not the propagator of the isolated electrode: inside the phonon selfenergy (6), there appears an electron propagator which may cross the partition; in fact, ${}^{\text{ph}}\Sigma_{ij}$ does depend on T . In defining g , we exclude diagrams where the electron crosses the barrier alone, but we retain those in which the electron makes an excursion across the partition while a virtual phonon is emitted. We shall see that this point is essential in describing the influence of localized barrier modes; here, we just emphasize the physical meaning of g , which is not as obvious as one might think.

We now come to the central approximation of our model: we assume that the phonon selfenergy ${}^{\text{ph}}\Sigma_{ij}$ cannot cross the partition (i lying on one side, and j on the other). This means that the electron propagator G_{ij} entering in ${}^{\text{ph}}\Sigma_{ij}$ only includes even numbers of T vertices. This is clearly valid for an Einstein model of phonons, since D_{ij} is then local (zero except when $i = j$). More generally, we expect ${}^{\text{ph}}\Sigma_{ij}$ to have a short range: one may always choose the partition in such a way that for a given process i and j lie on the same side: for instance, if we want to study virtual phonons near the insulator left electrode contact, we simply locate the partition near the right electrode. Such a trick always works, except possibly if the phonon is emitted in one electrode and reabsorbed in the other (the fact that the phonon goes all the way across the insulator is not a problem in itself). This assumption is needed in order to retain a simple algebra; while all the preceding approximations could easily be released at little expense in complication, the theory becomes very heavy if we allow ${}^{\text{ph}}\Sigma_{ij}$ to cross the dividing line. The perturbation expansion still works, but all the elegance of the method is lost (essentially because the distinction

between right and left is somewhat blurred by the phonons). At best, one can calculate G to second order in the phonon coupling V_i , that yields results which look complicated and are not very rewarding. Consequently, we do not consider further this complication, and we assume everywhere that phonons cannot cross the partition.

The only way for a propagator to cross the barrier is then via the transfer matrix element T . If we isolate the first isolated (ie not belonging to a Σ blob) T vertex which occurs, starting

from either the left or the right of a given diagram. we see that

$$G_{za}^r = g_{\alpha a}^r T G_{aa}^r = G_z^r T g_{aa}^r \quad (8)$$

If we make use of the recipe (5b), we see that

$$\begin{aligned} G_{xa}^+ &= g_{\alpha\alpha}^+ T G_{aa}^a + g_{x\alpha}^r T G_{aa}^+ \\ G_{a\alpha}^+ &= G_{aa}^r T g_{x\alpha}^+ + g_{aa}^+ T g_{x\alpha}^a \end{aligned} \quad (9)$$

We carry these expressions into equation 7, and we note that $G^r - G^a = G^- - G^+$. We thus obtain the simple result

$$\langle J \rangle = \frac{e}{\hbar} \int_{-\infty}^{\infty} T^2 \{ g_{\alpha\alpha}^+(\omega) G_{aa}^-(\omega) - g_{\alpha\alpha}^-(\omega) G_{aa}^+(\omega) \} \frac{d\omega}{2\pi} \quad (10)$$

G_{aa} itself may be written as an expansion in powers of T and g which has the structure

$$G_{aa} = g_{aa} + g_{aa} T g_{\alpha x} T g_{aa} + \dots$$

The $\pm$ component is then obtained from the recipe (5), and is equal to

$$G_{aa}^{\pm} = \Lambda^r (g_{aa}^{\pm} + g_{aa}^r T g_{\alpha a}^{\pm} T g_{aa}^a) \Lambda^a \quad (11)$$

where we have set

$$\Lambda = \frac{1}{1 - T^2 g_{aa} g_{\alpha x}} \quad (12)$$

The second term in the bracket of equation 11 gives a contribution to equation 10 which vanishes—a result which was to be expected since the current must depend on the occupation of electron states on both sides of the barrier. The current $\langle J \rangle$ reduces to

$$\langle J \rangle = \frac{e}{\hbar} \int |T \Lambda^r(\omega)|^2 [g_{x\alpha}^+(\omega) g_{aa}^-(\omega) - g_{x\alpha}^-(\omega) g_{aa}^+(\omega)] \frac{d\omega}{2\pi} \quad (13)$$

Let us emphasize that the result (13) is exact, to all orders in T (and of course in the bias). In the absence of phonon coupling, the propagators $g_{\alpha x}$ and g_{aa} correspond to two strictly isolated regions: if $T = 0$, there is no way to go from left to right. Under such conditions, the two sides are separately in thermodynamic equilibrium, respectively with chemical potentials μ_L and μ_R . Let us define densities of states at sites α and a by the relations

$$\begin{aligned} \rho_x &= -\frac{1}{\pi} \text{Im } g_{\alpha\alpha}^r \\ \rho_a &= -\frac{1}{\pi} \text{Im } g_{aa}^r \end{aligned} \quad (14)$$

The spectral densities then obey the relations

$$\begin{aligned} g_{\alpha x}^+(\omega) &= 2i\pi \rho_x(\omega) f_L(\omega) & g_{aa}^+(\omega) &= 2i\pi \rho_a(\omega) f_R(\omega) \\ g_{\alpha x}^-(\omega) &= 2i\pi \rho_x(\omega) \{f_L(\omega) - 1\} & g_{aa}^-(\omega) &= 2i\pi \rho_a(\omega) \{f_R(\omega) - 1\} \end{aligned} \quad (15)$$

where f_L and f_R are the Fermi factors respectively on the left and right hand side of the barrier. If we carry these expressions into equation 13, we find

$$\langle J \rangle = 2\pi \frac{e}{\hbar} \int d\omega |T \Lambda^r|^2 \rho_x(\omega) \rho_a(\omega) \{f_L(\omega) - f_R(\omega)\} \quad (16)$$

We recover the usual result of the transfer hamiltonian method (Duke 1969) with however one major difference: the effective 'transfer matrix element' $T\Lambda^r$ is energy dependent (Appelbaum and Brinkman 1969). Note that we have an explicit expression for that matrix element, in contrast to the usual approaches in which it remains more or less phenomenological. In the limit of zero temperature, the factor $(f_L - f_R)$ reduces to a step function between μ_L and μ_R : the integration over ω is thus restricted to a narrow range.

While equation 13 is exact, the result (16) is not strictly valid in the presence of an electron-phonon interaction. The reason for that breakdown is that inside a selfenergy $^{\text{ph}}\Sigma$, the electron may cross the partition; if a selfenergy $^{\text{ph}}\Sigma$ is inserted in a graph of say $g_{\alpha x}$, the final result samples the distributions on the two sides, and depends on μ_R as well as on μ_L . As a consequence, $g_{\alpha x}$ and g_{aa} no longer describe the propagation in two isolated regions which would be separately in equilibrium. The relations (15) do not apply, and equation 16 is thus incorrect. One feels however that this difficulty should not be dramatic in real tunnelling junctions, in which the transparency of the barrier is very small: in order for an electron to sample both μ_L and μ_R , it must cross the insulating layer, and the probability that it does so more than once is negligible. In order to clarify that point, we are led to study in more detail the nature of the junction.

In this paper we shall only consider the case of an MIM junction. Let α represent the last site of the left electrode, a the first site of the insulator; similarly, b denotes the last site of the insulator, facing the first site β of the right electrode. The latter contact is described by a transfer matrix element T' . From now on, g represents the propagator in the absence of any explicit vertex T or T' .

Let us consider the propagator $G_{ij}^{\pm}(\omega)$ between two points of the insulator. If T and T' were zero, the propagator would necessarily remain inside the barrier; since the range of ω of interest (mostly between μ_L and μ_R) lies in the forbidden band of the insulator, there is no density of states there, and $G_{ij}^{\pm}$ should then vanish. The only contributions to $G_{ij}^{\pm}$ result from excursions of the electron into either of the electrodes, where it can sample the finite density of states $g_{\alpha\alpha}^{\pm}(\omega)$ and $g_{\beta\beta}^{\pm}(\omega)$. That can occur either directly (independently of the phonons), or inside a phonon selfenergy (the excursion then occurs while a virtual phonon is emitted). The two processes are depicted in figure 1. From recipe (5), it follows that $G_{aa}^{\pm}$ contains three contributions

$$G_{aa}^{\pm} = G_{aa}^r T g_{\alpha\alpha}^{\pm} T G_{aa}^a + g_{ab}^r T' g_{\beta\beta}^{\pm} T' G_{ba}^a + g_{ai}^{\text{rph}} \Sigma_{ij}^{\pm} G_{ja}^a \quad (17)$$

where i, j are any sites in the insulator. The last term in equation 17 gives rise to the inelastic structure due to barrier modes, and will be dealt with later. If we ignore it for the moment, we see that only the second term in equation 17 contributes to equation 10 : the current through the junction may be written as

$$\langle J \rangle = \frac{e}{\hbar} \int \frac{d\omega}{2\pi} |U^r|^2 (g_{\alpha\alpha}^+ g_{\beta\beta}^- - g_{\alpha\alpha}^- g_{\beta\beta}^+) \quad (18)$$

in which the effective coupling matrix element between the two electrodes is now

$$U^r = T T' G_{ab}^r \quad (19)$$

The propagator G_{ab} decreases exponentially with the width d of the insulating layer—a statement of the physical fact that the electron wavefunction is evanescent in that region. It follows that U^r is very small, $|U^r|^2$ being essentially the 'transparency' of the barrier.

^{12†} We note that relations such as (15) are precisely those which break down when the system is not in thermodynamic equilibrium: that is why one needs the full formalism of Keldysh.

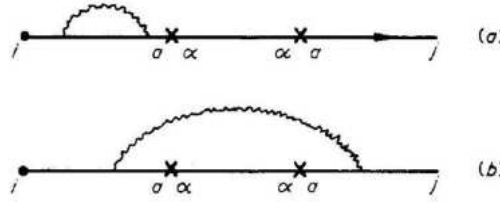


Figure 1. An electron in the insulator samples g_x^+ by going into the left electrode, either directly (case *a*) or inside a phonon selfenergy (case *b*).

Strictly speaking, the aforementioned difficulty remains: $g_{\alpha\alpha}^\pm$ and $g_{\beta\beta}^\pm$ are not the propagators of isolated electrodes in thermodynamic equilibrium, since inside a selfenergy $^{\text{ph}}\Sigma$ the electron can propagate anywhere in the system. For instance, $g_{\alpha\alpha}^\pm$ depends on the chemical potential μ_R . Such an effect however, is completely negligible for real junctions, since it implies that the electron makes an additional round trip across the insulator, from α to the right electrode and back. The corresponding contribution to $g_{\alpha\alpha}^\pm$ involves a factor $\sim |TT'G_{ab}|^2$; since equation 18 already contains a factor $|U^r|^2$, such corrections to the current are of second order in the barrier transparency, and thus utterly negligible. We can safely assume that $g_{\alpha\alpha}$ and $g_{\beta\beta}$ correspond to a thermal equilibrium situation, with chemical potentials respectively equal to μ_I and μ_R . That does not mean that the phonon selfenergy $^{\text{ph}}\Sigma$ is not affected by the penetration of the electron inside the insulator: indeed, it is! Simply, because of the evanescent nature of the propagation, this penetration remains localized in the immediate vicinity of the electrodes, and does not extend all the way across the barrier: the selfenergy $^{\text{ph}}\Sigma$ in the electrodes cannot sample simultaneously μ_L and μ_R (as shown in §4, this is not true if the selfenergy $^{\text{ph}}\Sigma$ is localized in the insulator).

Within the above approximation, the relations (15) are valid, and we may cast equation 18 in the form

$$\langle J \rangle = \frac{2\pi e}{\hbar} \int d\omega |U^r(\omega)|^2 \rho_\alpha(\omega) \rho_\beta(\omega) \{f_L(\omega) - f_R(\omega)\} \quad (20)$$

We recover the result of I. Equation 20 is identical to the expression obtained by the transfer hamiltonian technique-but for the fact that we have an explicit energy dependent expression for U^r . We are in a position to calculate systematically the effect of various renormalization processes on the transfer matrix element as well as on the densities of states.

It was shown in I that G_{ab}^r is modified by the presence of the electrodes. Let g_{ab} denote the barrier propagator in the absence of coupling T and T' (but possibly with phonon selfenergies $^{\text{ph}}\Sigma$). We have

$$G_{ab}^r = \frac{g_{ab}^r}{(1 - T^2 g_{\alpha\alpha}^r g_{aa}^r) (1 - T'^2 g_{\beta\beta}^r g_{bb}^r) - T^2 T'^2 g_{ab}^r g_{ba}^r g_{\alpha\alpha}^r g_{\beta\beta}^r} \quad (21)$$

Once again, $g_{ab} \ll g_{aa}$ if the barrier transparency is small: we can safely neglect the last term in the denominator of equation 21. The result (18) may then be cast in the form

$$\langle J \rangle = \frac{e}{\hbar} \int \frac{d\omega}{2\pi} |\tilde{U}^r|^2 (\tilde{g}_{\alpha\alpha}^+ \tilde{g}_{\beta\beta}^- - \tilde{g}_{\alpha\alpha}^- \tilde{g}_{\beta\beta}^+) \quad (22)$$

where we have defined

$$\begin{aligned} \tilde{U}^r &= TT' g_{ab}^r \\ \tilde{g}_{x\alpha}^\pm &= \frac{g_{xz}^\pm}{|1 - T^2 g_{x\alpha}^r g_{aa}^r|^2} \end{aligned} \quad (23)$$

(with a similar definition for $\tilde{g}_{\beta\beta}$). We note that the electrodes act to renormalize the transfer matrix element U ; in the form (22), that effect is absorbed into an 'effective' spectral density $\tilde{g}$.

As discussed before, $g_{\alpha\alpha}^\pm$ and $g_{\beta\beta}^\pm$ obey approximately the 'equilibrium relations' (15). Moreover, at this stage we do not take into account possible phonon selfenergies inside the barrier (we have ignored the last term in equation 17): g_{ab}^r is thus the propagator of the insulating layer alone, which is real in the energy range of interest (there is no density of states inside the gap of an insulator). We may then write equation 22 in the simpler, somewhat more explicit form

$$\langle J \rangle = \frac{2\pi e}{\hbar} \int d\omega \left| \tilde{U}^r(\omega) \right|^2 \tilde{\rho}_\alpha(\omega) \tilde{\rho}_\beta(\omega) \{f_L(\omega) - f_R(\omega)\} \quad (24)$$

where the effective densities of states $\tilde{\rho}_\alpha$ and $\tilde{\rho}_\beta$ are defined according to

$$\tilde{\rho}_\alpha = \frac{\rho_\alpha}{|1 - T^2 g_{xx}^r g_{aa}|^2} = -\frac{1}{\pi T^2 g_{aa}} \text{Im} \left[\frac{1}{1 - T^2 g_{\alpha\alpha}^r g_{aa}} \right] \quad (25)$$

(we have dropped the superscript on g_{aa} , since $g_{aa}^r = g_{aa}^a = g_{aa}$). Here again, $\tilde{\rho}_z$ lumps together the real density of states and the renormalization of U due to excursions in the electrodes, an effect which is usually neglected.

It is interesting to consider in some detail the frequency dependence of the various factors in equation 24. The matrix element $\tilde{U}^r(\omega)$ is characteristic of the insulator, and must be calculated in the presence of the applied bias (ie of the electric field inside the barrier); indeed, the deformation of the barrier with bias gives rise to the so called 'background current' in the tunnelling characteristics. We expect this effect to be negligible if the bias remains much smaller than the barrier height V_B ; the frequency dependence of $\tilde{U}(\omega)$ is then small in the range of interest (between μ_L and μ_R), and the exact potential profile in the insulator is mostly irrelevant. Here we consider only this simplified case, and we replace $|U^r|^2$ by a constant in equation 24: it must be borne in mind that such an approximation may break down if the bias is too large (for instance in Schottky barriers). On the contrary, phonon selfenergies in the electrodes will give rise to an appreciable frequency dependence of $g_{\alpha\alpha}^r$ and $g_{\beta\beta}^r$, and thus of $\tilde{\rho}_\alpha(\omega)$ and $\tilde{\rho}_\beta(\omega)$; this is precisely what is observed experimentally. On each side of the barrier, such a structure 'follows' the corresponding chemical potential (μ_L for $\tilde{\rho}_\alpha$, μ_R for $\tilde{\rho}_\beta$), so that strictly speaking the densities of states in equation 24 depend on bias. This is easily taken into account by

shifting the origin of ω ; $\tilde{\rho}_\alpha$ and $\tilde{\rho}_\beta$ are then written as functions of respectively $(\omega - \mu_L)$ and $(\omega - \mu_R)$. In the limit of zero temperature, equation 24 then reduces to

$$\langle J \rangle = \frac{2\pi e}{\hbar} |U^r|^2 \int_{\mu_R}^{\mu_L} d\omega \tilde{\rho}_\alpha(\omega - \mu_L) \tilde{\rho}_\beta(\omega - \mu_R) \quad (26)$$

Equation 26 is the starting point of our whole discussion of selfenergy effects in the electrodes.

12.0.3 3. Electron-phonon interaction in the electrodes of an MIM junction

In a metal, the electron-phonon interaction is correctly described within the Migdal approximation, and the equation 6 for Σ is valid. (Hereafter, we drop the index ph on Σ , which will only refer to the phonon selfenergy). It is not necessary to calculate separately Σ^r and $\Sigma^\pm$; to the extent that each electrode is nearly at equilibrium with a chemical potential μ equal to μ_L or μ_R , both quantities can be deduced directly from the single causal selfenergy Σ^c , which

in turn is obtained by ordinary perturbation theory. In a uniform system, Σ_{ij} depends only on the distance $(i - j)$, and may be characterized by its Fourier transform $\Sigma(k, \omega)$. As shown by Migdal, $\Sigma(k, \omega)$ is basically small (of the order of the Debye energy ω_D); it depends weakly on ϵ_k (on a scale $\sim \mu$), but very much on ω (on a scale $\sim \omega_D$). Qualitatively, these features persist in a nonuniform structure. The slow dependence on k means that in real space Σ_{ij} has a small range, of atomic size. Physically, the selfenergy describes the emission of a virtual phonon together with an electron fluctuation; because the sound velocity is much smaller than the Fermi velocity, the phonon does not have time to move very far during the lifetime of the electron fluctuation; Σ_{ij} is thus short ranged, mostly because of the phonon factor D_{ij} in equation 6. It is not a bad approximation to assume that Σ is local (ie zero unless $i = j$). On the other hand, $\Sigma_{ij}(\omega)$ depends rapidly on the frequency ω . In a metal, the bare density of electron states is practically constant on a scale ω_D : one may then show that $\text{Re } \Sigma^r$ is an odd function of $(\omega - \mu)$, while $\text{Im } \Sigma^r$ is even (the value of Σ at $\omega = \mu$ is of no physical interest, as it may be absorbed in the definition of the chemical potential). As shown by Migdal, the derivative $d(\text{Im } \Sigma^r)/d\omega$ is proportional to the phonon density of states at frequency $(\omega - \mu)$, while $d(\text{Re } \Sigma^r)/d\omega$ is essentially its Hilbert transform. For instance, in the case of an Einstein phonon spectrum with frequency ω_0 , one has

$$\begin{aligned} \frac{d}{d\omega} (\text{Im } \Sigma^r) &\sim \delta(|\omega - \mu| - \omega_0) \text{sgn}(\omega - \mu) \\ \frac{d}{d\omega} (\text{Re } \Sigma^r) &\sim \frac{2\omega_0}{\omega_0^2 - (\omega - \mu)^2} \end{aligned}$$

$d(\text{Im } \Sigma^r)/d\omega$ displays a characteristic dissipative shape, while $d(\text{Re } \Sigma^r)/d\omega$ has a dispersive behaviour. For real phonon spectra, these structures are broadened. It appears difficult to go beyond these simple qualitative remarks, since both the phonons and the electrons are affected by the metal-insulator contact. In calculating Σ_{ij} , the phonon propagator D_{ij} depends on the acoustic matching between metal and insulator (it probably involves also surface phonons at the interface); similarly, the electron propagator G_{ij} is affected by the reflection of the wavefunction on the insulating barrier (this is shown on a specific example in the Appendix). It is reasonable to assume that any peak in the phonon density of states will show up in $d(\text{Im } \Sigma^r)/d\omega$, but one cannot honestly make any reliable quantitative calculation, at least for the electrodes that we are presently considering.

Let us now consider to what extent Σ affects the propagator g_{xx}^r . In an uniform medium, we would write

$$g_{\alpha\alpha}^r = \sum_k \frac{1}{\epsilon_k - \omega - \Sigma^r(k, \omega)} \quad (27)$$

If we neglect the dependence of Σ^r on k , the selfenergy acts only to shift the frequency, and we obtain

$$g_{x\alpha}^r(\omega) = g_{\alpha\alpha}^{r0} \{\omega + \Sigma^r(\omega)\} \quad (28)$$

where $g_{\alpha\alpha}^0$ is the propagator in the absence of electron-phonon interaction. g_{xx}^0 varies on a scale of order μ , so that the selfenergy correction to equation 28 is of relative magnitude ω_D/μ , that is, very small. Put another way, the electron-phonon interaction in the electrodes does not affect the tunnel current within Migdal's approximation—a well known result. If we retain corrections of order ω_D/μ , we must also take into account the k dependence of Σ^r .

As such, the above discussion does not apply to structures which are not translationally invariant. To the extent, however, that Σ is small, the following lowest order expansion $\dagger$ is sufficient

$$\begin{aligned}
g_{xx}^r &= g_{x\alpha}^{r0} + \delta g_{\alpha x}^r \\
\delta g_{x\alpha}^r &= \sum_{\lambda, \mu} g_{x\lambda}^{r0} \Sigma_{\lambda\mu}^r g_{\mu x}^{ro}
\end{aligned} \tag{29}$$

$g_{xx}^\pm$ is then given by equation 15. In this section, we assume that λ and μ lie in the metallic regions: the unperturbed propagator $g_{\alpha\lambda}^{r0}$ is then an oscillatory function of the distance $(\alpha - \lambda)$, with a wavevector of order k_F . The summation over λ and μ in equation 29 is therefore subject to a strong destructive interference, the whole sum being of the order of a single term; since $\Sigma \sim \omega_D$, $g_{x\alpha}^0 \sim 1/\mu$, the correction $\delta g_{\alpha x}$ is of order ω_D/μ smaller than g_{xx}^0 . This is another statement of Migdal's theorem, which here appears as a consequence of the atomic range of the summations in equation 29. We shall see that this interference effect is essential in controlling the shape of the $I - V$ characteristics.

The phonon renormalization of the current appears in the effective densities of states $\tilde{\rho}_\alpha$ and $\tilde{\rho}_\beta$ of equation 24. In analogy with equation 29, we may write (to first order)

$$\begin{aligned}
\tilde{\rho}_\alpha &= \tilde{\rho}_\alpha^0 + \delta\tilde{\rho}_\alpha \\
\tilde{\rho}_\beta &= \tilde{\rho}_\beta^0 + \delta\tilde{\rho}_\beta
\end{aligned} \tag{30}$$

The corrections $\delta\tilde{\rho}_\alpha$ and $\delta\tilde{\rho}_\beta$ are obtained from equation 25

$$\begin{aligned}
\delta\tilde{\rho}_\alpha &= -\frac{1}{\pi} \text{Im} \frac{\delta g_{\alpha x}^r}{(1 - T^2 g_{ad} G_{x\alpha}^{r0})^2} \\
\delta\tilde{\rho}_\beta &= -\frac{1}{\pi} \text{Im} \frac{\delta g_{\beta\beta}^r}{(1 - T'^2 g_{bb} g_{\beta\beta}^{ro})^2}
\end{aligned} \tag{31}$$

Within our linear approximation, we retain only first order corrections in $\tilde{\rho}_\alpha \tilde{\rho}_\beta$, that is

$$\tilde{\rho}_\alpha^0 \tilde{\rho}_\beta^0 + \tilde{\rho}_\alpha^0 \delta\tilde{\rho}_\beta + \tilde{\rho}_\beta^0 \delta\tilde{\rho}_\alpha$$

The phonon corrections due to the two electrodes are thus additive; the conductance $G(V) = d\langle J \rangle / dV$ may be written as

$$G = G_0 + \delta G_I + \delta G_R \tag{32}$$

where δG_I and δG_R arise respectively from phonons in the left and right electrodes. (We saw in §2 that the renormalization of $\tilde{\rho}_\alpha$ and $\tilde{\rho}_\beta$ were totally decoupled).

In practice, phonon structures appear for biases $\mu_L - \mu_R = eV$ of the order of the phonon energy ω_D , which for ordinary MIM junctions is much smaller than both the barrier height and the Fermi energy. We may therefore neglect the frequency dependence of $\tilde{\rho}_\alpha^0, \tilde{\rho}_\beta^0$, as well as that of the transfer matrix element $\tilde{U}^r$. In the limit of zero temperature, we use the simpler form (26), which yields at once

$$\begin{aligned}
G^0 &= \frac{2\pi e^2}{\hbar} |U^r|^2 \tilde{\rho}_\alpha^0 \tilde{\rho}_\beta^0 \\
\delta G_I(V) &= \frac{2\pi e^2}{\hbar} |U^r|^2 \tilde{\rho}_\beta^0 \delta\tilde{\rho}_\alpha(-eV) \\
\delta G_R(V) &= \frac{2\pi e^2}{\hbar} |U^r|^2 \tilde{\rho}_\alpha^0 \delta\tilde{\rho}_\beta(+eV)
\end{aligned} \tag{33}$$

¹² The summation in equation 29 is one dimensional, in a direction perpendicular to the barrier; an expression of the form (29) is valid for every value of $k_\parallel$. The problem of summing over $k_\parallel$ is briefly discussed later in this section.

A measurement of the conductance $G(V)$ thus provides direct information on $\delta\tilde{\rho}_\alpha$ and $\delta\tilde{\rho}_\beta$. (We note that equation 33 is not valid if U^r . $\tilde{\rho}_\alpha^0$ or $\tilde{\rho}_\beta^0$ vary appreciably over a scale eV . as is the case. for instance. in Schottky barriers; one must then return to the general expression 24).

Let us consider for example phonon effects in the left electrode, contributing to δG_L . We have seen that $\tilde{\rho}_\alpha$ lumps together two different physical factors: the density of states ρ_α of the isolated electrode, and the effect of the electrode on the transfer matrix element U^r (due to excursions of the tunnelling electron into the metallic region). Both factors are affected by the phonons, and contribute to $\delta\tilde{\rho}_\alpha$. We note that the renormalization of U is energy dependent, and thus modifies the shape of the $G(V)$ characteristics. Such an effect would be entirely missed in the transfer hamiltonian formulation, which considers U^r as time independent from the outset. On the contrary, it is included very naturally in our theory, and in fact exactly. For practical purposes, we treat $\delta\tilde{\rho}_\alpha$ as a whole, given by equation 31 . The corresponding denominator $(1 - T^2 g_{ag} g_{\alpha x}^{r0})^2$ is energy independent in the range of interest; let A and θ be its modulus and argument: $\delta\tilde{\rho}_\alpha$ may be written as

$$\delta\tilde{\rho}_\alpha = -\frac{1}{\pi A} (\cos \theta \operatorname{Im} \delta g_{\alpha x}^r - \sin \theta \operatorname{Re} \delta g_{\alpha x}^r) \quad (34)$$

The value of θ depends on the detailed band structure of the metal M and insulator I, as well as on the matching of the wavefunctions at the boundary. Since $T, 1/g_{aa}$ and $1/g_{\alpha x}^{r0}$ are all of the same order of magnitude (typically a bandwidth), θ can take about any value, depending on the junction considered. The conductance δG_L will thus contain an unpredictable mixture of $\operatorname{Im} \delta g_{\alpha x}^r$ and $\operatorname{Re} \delta g_{\alpha x}^r$ (in contrast to ordinary theories which corresponds to $\theta = 0$); the coefficient of $\operatorname{Im} \delta g_{\alpha x}^r$ may even be negative!

$\delta g_{\alpha x}^r$ is itself given by equation 29. Let us for simplicity assume that $\Sigma_{\lambda\mu}$ is local: we have

$$\delta g_{\alpha x}^r = \sum_{\lambda} (g_{\alpha\lambda}^{r0}) \Sigma_{\lambda\lambda}^r \quad (35)$$

(remember that time reversal symmetry of the isolated electrode implies $g_{\alpha,i}^{r0} = g_{\lambda\alpha}^{r0}$. We know that in a metal $g_{x\lambda}^{r0}$ is oscillatory, with a phase which varies rapidly on atomic distances. The real and imaginary parts of $\delta g_{\alpha x}^r$ therefore contain a combination of $\operatorname{Im} \Sigma_{\lambda\lambda}^r$ and $\operatorname{Re} \Sigma_{\lambda\lambda}^r$, in a proportion that is unpredictable unless one has a precise model of the junction. This is another aspect of the destructive interference mentioned earlier. Altogether, the conductance $\Delta G_L(V)$ appears as a combination of $\operatorname{Re} \Sigma_{\lambda\lambda}^r(-eV)$ and $\operatorname{Im} \Sigma_{\lambda\lambda}^r(-eV)$. For MIM junctions, the frequency dependence of $g_{\alpha\lambda}^{r0}$ in equation 35 can be neglected: the imaginary and real parts of the selfenergy thus give contributions to ΔG_L that are respectively even and odd in the applied bias V . The coefficients of these two terms should be comparable in magnitude; their exact value is very much model dependent, and cannot be calculated reliably. Note that the coefficient of $\operatorname{Im} \Sigma_{\lambda\lambda}^r$ may well be negative: that would correspond to a decrease in the conductance, in contrast to the inelastic tunnelling model which always gives an increase of G (due to the opening of a new 'tunnelling channel'). †

In conclusion, the shape of the structures in $G(V)$ due to selfenergy corrections in the electrodes is practically impossible to predict, as it would require an exact description of the interface, both for the electrons and for the phonons. One thing is sure: the effect is small, because of the strong destructive interference due to the oscillation of $g_{\alpha i}^{r0}$. Experimentally, the situation looks appreciably different: indeed, the structure in $G(V)$ arising from electrode phonons does contain even and odd parts; however the even part is usually dominant, and also is positive (corresponding to an increase in G). Formally, that would correspond to a small value of θ in equation 34 . It is clear that the above discussion cannot account for these facts: a likely explanation will be given later in this section.

Up to now, the partition of the system into two halves was made at the contact between the left electrode and the insulator, between sites α and a . Such a choice, although natural, is not very appropriate in order to study selfenergy effects in the left electrode close to the boundary: excursions into the insulator must be treated explicitly, which somehow complicates the calculation. An alternate possibility is to locate the partition well inside the insulator, in such a way that the propagator from the I – M contact to the partition be exponentially small (because of evanescent wavefunctions). With such a choice, what happens in or near the metal M will essentially ignore the existence of a right hand side of the partition. In order to avoid confusion with the previous treatment, we denote by γ and c the sites on either side of the new partition, and by $\mathcal{G}$ the one sided propagator (previously called g): our choice corresponds to $\mathcal{G}_{xy}$ very small. The current is given by the general expression (13), with the obvious replacements $\alpha, a \rightarrow \gamma, c, g \rightarrow \mathcal{G}$.

The spectral densities $\mathcal{G}_{\gamma\gamma}^{\pm}$ necessarily involve the penetration of the electron into the left electrode, where it can sample the finite density of states near $\omega = \mu_L$. The corresponding corrections due to left electrode phonons will thus be sizeable, since the electron has to enter the metal M anyway in order to contribute to $\mathcal{G}_{\gamma\gamma}^{\pm}$ (the fact that this factor is evanescent gives rise to the small transparency of the barrier). On the contrary, the matrix element Λ^r would be nonzero even for a pure insulator: since it involves retarded propagators, it does not require any finite density of states near energy ω (as did $\mathcal{G}^{\pm}$). To the extent that the partition is far from the electrodes, we can neglect the influence of the latter on Λ^r : that would involve an additional round trip from γ to α and back, yielding an extra evanescent factor. Λ^r is thus a pure 'insulator' quantity, unaffected by the electrodes, and especially by their selfenergies (note that in such a case $\Lambda^r = \Lambda$ is real). In principle, the spectral densities $\mathcal{G}_{cc}^{\pm}$ could be affected by the left electrode via the selfenergies near the partition (see § 2): for the time being, we neglect selfenergies in the insulator, and such corrections do not exist. We conclude that selfenergy effects in the left electrode only appear in $\mathcal{G}_{\gamma\gamma}^{\pm}$. To lowest order in the electron-phonon coupling, we may write

$$\begin{aligned}\mathcal{G}_{\gamma\gamma}^r &= \mathcal{G}_{\gamma\gamma}^{r0} + \delta\mathcal{G}_{\gamma\gamma}^r \\ \delta\mathcal{G}_{\gamma\gamma}^r &= \sum_{\lambda\mu} \mathcal{G}_{\gamma\lambda}^{r0} \Sigma_{\lambda\mu}^r \mathcal{G}_{\mu\gamma}^{r0}\end{aligned}\quad (36)$$

For an MIM junction in which Λ and the bare density of states ρ_{β}^0 are weakly ω dependent, the extra conductance $\Delta G_L(V)$ is given by

$$\Delta G_L = \left(\frac{e^2}{2\pi\hbar} T^2 \Lambda^2 \text{Im} \mathcal{G}_{cc}^{r0} \right) \text{Im} \delta\mathcal{G}_{\gamma\gamma}^r(-eV) \quad (37)$$

(at zero temperature). We know that $\text{Im} \mathcal{G}_{cc}^{r0} < 0$ (see equation 14): thus within a real, positive constant, ΔG_L is proportional to $-\text{Im} \delta\mathcal{G}_{\gamma\gamma}^r$. The result (37) is equivalent to equation 33. Indeed, the propagator $\mathcal{G}$ for the left electrode + half barrier is easily expressed in terms of the corresponding g for the metal M and insulator I alone; simple algebra shows that

$$\mathcal{G}_{\gamma\lambda}^r = \frac{g_{\gamma a} T g_{\alpha\lambda}^r}{1 - T^2 g_{aa} g_{\alpha\lambda}^r}$$

The first order correction $\delta\mathcal{G}_{\gamma\gamma}^r$ is thus

$$\delta\mathcal{G}_{\gamma\gamma}^r = (g_{\gamma a} T)^2 \frac{g_{\alpha\lambda}^r \Sigma_{\lambda g}^r g_{\mu\alpha}^r}{(1 - T^2 g_{aa} g_{\alpha\alpha}^r)^2} \quad (38)$$

On comparing equation 38 with 29 and 31, we see that $\text{Im} \delta\mathcal{G}_{\gamma\gamma}^r$ is proportional to $\delta\tilde{\rho}_{\alpha}$. By moving the partition away from the metal M, we avoid the need to distinguish corrections to

the density of states and to the matrix element U^r : the two effects are lumped together in $\delta\mathcal{G}_{\gamma\nu}$ (which is hardly surprising since we now let the electron move freely from M to I).

The advantage of equations 36 and 37 is that they allow a simple comparison of selfenergy effects in the electrodes and in the barrier. Let us for instance assume that $\Sigma_{\lambda,\mu}^r$ is local; its real part is odd in V , its imaginary part is even (see above). Within a positive constant, $\Delta G_L^{(V)}$ is proportional to

$$\sum_{\lambda} \text{Im} \left\{ (\mathcal{G}_{\gamma\lambda}^{r0})^2 \Sigma_{\lambda\lambda}^r(-eV) \right\} \quad (39)$$

Everything will depend on the argument of $\mathcal{G}_{vi\lambda}^{r0}$! A specific example is analysed in the Appendix: here we only give a qualitative discussion. When λ lies in the metal, $\mathcal{G}_{\gamma\lambda}^{r0}$ is oscillatory: we find again the destructive interference described before; the final result involves a more or less arbitrary combination of $\text{Re} \Sigma_{\lambda\lambda}$ and $\text{Im} \Sigma_{\lambda\lambda}$, the shape of the phonon structure being unpredictable. If on the other hand λ lies in the insulator, $\mathcal{G}_{\gamma\lambda}$ is real: in that case, the current is proportional to $-\text{Im} \Sigma_{\lambda\lambda}$. The corresponding conduction $\Delta\mathcal{G}_L(V)$ is even and positive, in agreement with the relevant experimental facts. The physical reason for that quite different behaviour is obvious: because the evanescent wavefunctions have a constant phase, the destructive interference characteristic of metallic regions disappears in the insulator; only the dissipative processes, proportional to $\text{Im} \Sigma_{\lambda i}$, remain in the conductance. The situation is thus much simpler than in the electrodes.

The problem of selfenergy effects inside the barrier will be treated more systematically in §4. The only purpose of the present discussion is to give a qualitative feeling for the essential physical effect, namely interference in the metal. In this way, we can understand the discrepancy between our previous results and the experimental facts concerning structure due to the electrode phonons. The latter are mostly due to selfenergy corrections inside the barrier, due to phonons which have penetrated from the metal into the insulator. (Such a penetration is easy since usually the oxides are harder than the metals, and thus have higher phonon frequencies). The structure of the density of states of metallic phonons extends into the insulator, and contributes to Σ_{ij} even when i and j belong to the barrier. Even though the phonon amplitudes may be reduced in going through the interface, such an effect is competitive in magnitude, since the interference is constructive, instead of destructive, in the summation of equation 39. (While the 'pure' electrode effect involved only the first few layers of the metal, the 'indirect' contribution extends throughout the insulating layer.) In conclusion, we believe that the electrode phonons manifest themselves through their tails in the insulator more strongly than by their selfenergies in the metal itself. That explains why the observed phonon structures in $G(V)$ are mostly even and positive, the odd part being usually noticeably smaller.

The preceding discussion raises a new problem: over what distance does the junction 'feel' the phonons in the electrode. For instance, if the left electrode is a sandwich, made of a layer of metal M_1 , of thickness a , covered with metal M_2 , what is the minimum a for which the phonons of M_2 are seen in the tunnelling characteristics? The selfenergy $\Sigma_{\lambda\mu}$ is certainly negligible when the distance $(\lambda - \mu)$ exceeds a few atomic spacings. Long range action can thus occur only in two ways; (i) through the long range of the factor $(g_{\alpha\lambda}^{r0})^2$ in equation 35, and (ii) through the penetration of phonons over long distances, as discussed in the preceding paragraph. In equation 35, the sites λ, μ are close to the barrier, but the selfenergy feels the phonon density of states of remote regions. We shall discuss the two effects separately.

We have seen that in the metal, $g_{\alpha\lambda}^{r0}$ oscillates with the position λ . It is shown in the Appendix that for a continuous model

$$g_{\alpha\lambda}^{r0}(\omega, k_{\parallel}) \sim \exp(ik_0 z) \quad (40)$$

where z is the distance to MI interface, and k_0 the z component of the Fermi wavevector

$$k_0 = (k_F^2 - k_{\parallel}^2)^{1/2} \quad (41)$$

To the extent that the exponential (40) is undamped, the range of z in equation 35 may be quite large for a given $k_{\parallel}$ (in principle, it is limited only by the phonon and electron mean free paths). Of course, the destructive interference limits the amplitude of the effect, yet it is in principle possible to feel phonon selfenergies from large distances. Such a conclusion, however, breaks down when we perform the sum over $k_{\parallel}$: since k_0 varies between 0 and k_F , the various terms of the summation interfere destructively, and the range of λ in equation 35 is considerably reduced. The actual range is determined by the dependence of all other factors on $k_{\parallel}$: if the leading terms correspond to a width Δk_0 , the spatial range is of order $1/\Delta k_0$.

In practice, the selfenergy $\Sigma_{\lambda\lambda}$ depends weakly on $k_{\parallel}$. (by virtue of Migdal's approximation); its variation is completely negligible as compared to that of the transfer matrix element U^r in equation 33. For a qualitative estimate, we can use the WKB approximation, which yields

$$|U^r|^2 \sim \exp \left\{ -2d \left(\frac{2m}{\hbar^2} V_B - k_0^2 \right)^{1/2} \right\} \quad (42)$$

d is the insulator thickness, V_B the barrier height measured from the bottom of the conduction band. As expected $|U^r|^2$ is sharply peaked near $k_{\parallel} = 0$, that is near $k_0 = k_F$. In practice, V_B and $\mu = \hbar^2 k_F^2 / 2m$ are of the same order of magnitude. The transfer probability $|U^r|^2$ then becomes negligible if $\Delta k_0 = k_F - k_0$ exceeds $1/d$. The range of the summation over λ in equation 35 should be of the order of the barrier thickness: the junction can only feel selfenergy corrections that are located within about 50Å from the insulator.

Actually, this already short range is probably overestimated, as it assumes the conservation of $k_{\parallel}$, that is well defined plane parallel layers of the different materials. In real junctions, the thickness of the insulator will vary from place to place, the current flowing mostly in 'weak spots' where the insulator is narrow. The conservation of $k_{\parallel}$ is then relaxed, the uncertainty $\Delta k_{\parallel}$ being of the order of the inverse diameter of the weak spots, b . If $b < d$, the decrease of $|U^r|^2$ with $k_{\parallel}$ is controlled by this effect, and the range of the selfenergy corrections is b rather than d . Moreover, any thickness fluctuation of the layer $\mathbf{M}_1$ will tend to blur the phase of $g_{\alpha,\lambda}^{\text{ro}}$ in equation 35, and thus to reduce the selfenergy correction. Such an effect will be dominant if the thickness fluctuations exceed a few atomic layers-at any rate a distance far too short to be attained experimentally.

We are thus left with the other mechanism, in which the selfenergy near the barrier is affected by phonons from remote metallic regions; in the sandwich described before that means that the phonons of M_2 can penetrate close to the insulator. A quantitative discussion would require a detailed analysis of the acoustic matching in such a multilayer structure, which has been attempted by Léger (1971). Léger points out that the penetration is possible if the M_1 material is harder than M_2 (higher phonon frequencies); if M_1 is softer than M_2 , most of the phonons of M_2 are totally reflected, and do not reach the insulator. This qualitative picture is substantiated by experiment, which confirms that the range of the λ summation in equation 35 is very small. (Note that a similar argument was used to support the predominance of barrier effects over electrode effects).

To conclude this section, it may be useful to discuss the terminology usually employed to analyse experimental data. It is customary to separate out the so called 'elastic' and 'inelastic' parts of the phonon structure, their exact definition varying somewhat from paper to paper. In our opinion, the correct definition is obtained by using equations 18 and 29, and by expanding $\delta g_{\alpha\alpha}^+$ in the form

$$\delta g_{\alpha\alpha}^{\pm} = \sum_{\lambda\mu} \left(g_{\alpha\lambda}^{\pm 0} \Sigma_{\lambda\mu}^a g_{\mu\alpha}^{a0} + g_{\alpha\lambda}^{r0} \Sigma_{\lambda\mu}^r g_{\mu\alpha}^{0\pm} + g_{\alpha\lambda}^{r0} \Sigma_{\lambda\mu}^{\pm} g_{\mu\alpha}^{a0} \right) \quad (43)$$

The first two terms in the bracket of equation 38 involve the bare spectral density $g^{0\pm}$; they describe an elastic process in which no real phonon is emitted: the phonons act only to renormalize the various propagators. The last term, on the other hand, involves $\Sigma_{i,\mu}^+ \sim G_{\lambda\mu}^+ D_{i\mu}^+$: we shall see in the following section that such terms contain the spectral densities of electrons and phonons; they may be considered as inelastic terms, describing processes in which a real phonon is emitted (the energy conservation involves the phonon energy). In this respect, our formulation may thus be reconciled with the more usual descriptions. It should be recognized, however, that such a distinction is in a way a

matter of semantics. It depends on whether one wants to consider the tunnelling of bare particles or quasiparticles. It is in fact simpler to lump the two contributions into a single term, as we did in this paper (anyhow, the 'elastic' part contains a combination of $\text{Re } \Sigma$ and $\text{Im } \Sigma$ which is comparable to the 'inelastic' term, and cannot be separated from it). In addition, it should be stressed that the phonon emission in 'inelastic' processes occurs before or after the tunnelling has taken place: one can hardly say that it opens a new 'channel'. We therefore believe that the decomposition (43) is of no physical interest, and tends instead to confuse the issue.

12.0.4 4. Electron-phonon interaction in the barrier

We now consider selfenergies Σ_{ij} for which both i and j are located in the insulating barrier [†]. Σ_{ij} may arise from phonons of the insulating matrix itself, or from localized modes due to impurities, for instance from the internal vibrations of adsorbed molecules (OH bonds, etc . . .). We shall not attempt a specific calculation of Σ_{ij} . We note only that its range ($i - j$) may be appreciable, not so much because of the phonon factor D_{ij} (which is very localized) as for the long range of the electron-phonon interaction (which is not screened as in a metal) [‡].

The current across the junction is obtained from equations 10 and 17 . Within a first order calculation, the electrode and barrier effects are simply additive; in this section we shall ignore the electrode selfenergies: $g_{\alpha\alpha}$ and $g_{\beta\beta}$ are the zeroth order propagators, with no electron-phonon coupling. The second term of equation 17 thus involves the bare spectral densities of the two electrodes, ρ_{α}^0 and ρ_{β}^0 ; we may interpret it as describing an elastic tunnelling process, in which the barrier phonons act only to renormalize the transfer matrix element $TT'G_{ab}^r$. The last term of equation 17, on the other hand, involves $\Sigma_{ij}^{\pm}$: we shall now show that it corresponds to an inelastic process, in which a real phonon is emitted during the tunnelling event. The corresponding contribution to the current may be written as

$$\langle J_{\text{in}} \rangle = \sum_{ij} \frac{e}{\hbar} \int \frac{d\omega}{2\pi} T^2 T'^2 G_{ai}^r(\omega) G_{ja}^a(\omega) \{ g_{\alpha\alpha}^+(\omega) \Sigma_{ij}^-(\omega) - g_{\alpha\alpha}^-(\omega) \Sigma_{ij}^+(\omega) \} \quad (44)$$

(the summation over i and j extends over the whole barrier). At first sight, it may look surprising that the two electrodes do not enter symmetrically in the expression of $\langle J \rangle$: in fact, this difficulty will disappear as soon as we replace $\Sigma_{ij}^{\pm}$ by its explicit form. Let us for simplicity work to second order in the electron-phonon coupling. We denote as G^0 the value of G in the absence of phonon effects[§]. Since Σ appears explicitly in equation 44, the factors G_{aj}^r and G_{ja}^a may be replaced by their zeroth order values: moreover, $\Sigma_{ij}^{\pm}$ is given by equation 6^{||}, in which $G_{ij}^{\pm}$ may be replaced by $G_{ij}^{\pm 0}$. Equation 44 thus becomes

^{12†} The intermediate case in which i lies in the insulator, j in an electrode-say M -was explicitly excluded in §1. If required, it is easily treated by moving the partition away, as done at the end of §3.

$$\langle J_{\text{in}} \rangle = \sum_{ij} \frac{ie}{\hbar} V_i V_j \iint \frac{d\omega}{2\pi} \frac{d\omega'}{2\pi} T^2 G_{ai}^{\text{r}0}(\omega) G_{ja}^{\text{a}}(\omega) \quad (45)$$

$$\times \{ g_{\alpha\alpha}^+(\omega) G_{ij}^{-0}(\omega') D_{ij}^-(\omega - \omega') - g_{\alpha\alpha}^-(\omega) G_{ij}^{+0}(\omega') D_{ij}^+(\omega - \omega') \}$$

As discussed in § 2, $G_{ij}^{\pm 0}$ would vanish if the electron did not make excursions in the electrodes, where it can sample a finite density of states in the energy range of interest. In analogy with equation 17, we may write (see equation 5):

$$G_{ij}^{\pm 0} = G_{ia}^{\text{r}0} T g_{\alpha\alpha}^{\pm} T G_{aj}^{\text{a}0} + g_{ib}^{\text{r}0} T' g_{\beta\beta}^{\pm} T' G_{bj}^{\text{a}0} \quad (46)$$

We now proceed to simplify these results.

When inserted into equation 45, the first term of equation 46 yields a contribution which involves a factor

$$G_{ai}^{\text{r}0}(\omega) G_{ja}^{\text{a}0}(\omega) G_{ia}^{\text{r}}(\omega') G_{aj}^{\text{a}0}(\omega') \{ g_{\alpha\alpha}^+(\omega) g_{\alpha\alpha}^-(\omega') D_{ij}^-(\omega - \omega') - g_{\alpha\alpha}^-(\omega) g_{\alpha\alpha}^+(\omega') D_{ij}^+(\omega - \omega') \} \quad (47)$$

We notice that $G^{\text{r}0}$ and $G^{\text{a}0}$ are symmetric: $G_{ai}^{\text{r}0} = G_{ia}^{\text{r}0}$; the factor in front of the bracket in equation 47 is therefore invariant if $\omega \leftrightarrow \omega'$. We moreover assume that the phonons are in thermodynamic equilibrium, which implies ‡

$$D_{ij}^+(\omega) = D_{ij}^-(-\omega) \quad (48)$$

It is then clear that the contribution (47) vanishes-as it should since it involves spectral densities in the left electrode only. We are left with the second term of equation 46, which may be cast in the form

$$\langle J_{\text{tn}} \rangle = \sum_{ij} \frac{2\pi e}{\hbar} \iint d\omega d\omega' W_{ij}(\omega, \omega') F \quad (49)$$

where we have set

$$W_{ij}(\omega, \omega') = T^2 T'^2 V_i V_j G_{ai}^{\text{r}0}(\omega) G_{ib}^{\text{r}0}(\omega') G_{bj}^{\text{a}}(\omega') G_{ja}^{\text{a}0}(\omega) \quad (50)$$

$$F = \frac{i}{(2\pi)^3} \{ g_{\alpha\alpha}^+(\omega) g_{\beta\beta}^-(\omega') D_{ij}^-(\omega - \omega') - g_{\alpha\alpha}^-(\omega) G_{\beta\beta}^+(\omega') D_{ij}^+(\omega - \omega') \}$$

$W_{ij}(\omega, \omega')$ replaces the former squared matrix element $|U^{\text{r}}|^2$ for the phonon assisted process under consideration. It describes a fairly simple physical situation: the electron goes from a to i with energy ω , interacts with a virtual phonon, and then proceeds to b with energy ω' ; the return trip is similar, except that the phonon is reabsorbed at a different site j (one must let

‡ The propagator G_{ij} in the equation 6 for the selfenergy does not control the range of Σ_{ij} , since it is real and thus not subject to any interference. Its role is mainly to contribute to the exponential factor characterizing the transparency of the barrier, as shown later in this section.

§ Remember that G is the full propagator, while g is the propagator when the matrix elements T and T' are set equal to zero.

|| For simplicity, we consider only a local electron-phonon vertex. The generalization to an arbitrary matrix element V_{ijl} (see equation 2) is straightforward.

the scattering amplitudes interfere before one calculates the transition probability). The other factor in equation 49, F , contains all the information

on the density and occupation of electron and phonon states: we now study its structure in more detail.

Let us define a phonon spectral density

$$\sigma_{ij}(\omega) = \sum_n A_n^*(i) A_n(j) \delta(\omega - \omega_n) \quad (51)$$

where $A_n(j)$ is the amplitude of a phonon eigenstate n , with energy ω_n , at the crystal site j . Time reversal symmetry implies that σ_{ij} is symmetric in the interchange $i \leftrightarrow j$. In thermodynamical equilibrium, the propagators $D_{ij}^\pm$ may be related to σ_{ij} , in much the same way as equation 15 relates $g_{\alpha x}^\pm$ to the density of states ρ_x . From the definition of the various phonon propagators, it follows that for $\omega > 0$, one has [†]

$$\begin{aligned} D_{ij}^+(\omega) &= D_{ij}^-(-\omega) = -2i\pi\sigma_{ij}(\omega)N(\omega) \\ D_{ij}^+(-\omega) &= D_{ij}^-(\omega) = -2i\pi\sigma_{ij}(\omega)\{N(\omega) + 1\} \end{aligned} \quad (52)$$

where $N(\omega)$ is the statistical factor for bosons

$$N(\omega) = \frac{1}{\exp \beta\omega - 1} \quad (53)$$

Using equations 15 and 52, we see that for $\omega > \omega'$, F takes the form

$$\begin{aligned} F &= \rho_\chi(\omega)\rho_\beta(\omega')\sigma_{ij}(\omega - \omega')[f_L(\omega)\{1 - f_R(\omega')\}\{1 + N(\omega - \omega')\} \\ &\quad - \{1 - f_L(\omega)\}f_R(\omega')N(\omega - \omega')] \end{aligned} \quad (54a)$$

while in the opposite case $\omega < \omega'$ one has

$$\begin{aligned} F &= \rho_g(\omega)\rho_\beta(\omega')\sigma_{ij}(\omega' - \omega)[f_L(\omega)\{1 - f_R(\omega')\}N(\omega' - \omega) \\ &\quad - \{1 - f_L(\omega)\}f_R(\omega')\{1 + N(\omega' - \omega)\}] \end{aligned} \quad (54b)$$

The physical interpretation of these results is now obvious; for instance, the first term in the bracket of equation 54a describes tunnelling from a filled, left state, with energy ω , to an empty right state with lower energy ω' , while a phonon of energy $(\omega - \omega')$ is emitted; in the first term of equation 54b, $\omega' > \omega$, and the phonon is absorbed instead of emitted. The current (49) thus corresponds to a true inelastic process, in which a real phonon is involved (the phonon energy enters the energy conservation relation). The relations (54a.b) display clearly the statistical factors associated with each possible process. The result (49) is just what one would guess by using the golden rule of perturbation theory. We note that the two electrodes enter symmetrically in equations 49,50 and 54 for the current, which takes care of the objection raised earlier.

In the limit of zero temperature, the phonon population $N(\omega)$ vanishes, and the inelastic current takes the simple form

^{12†} This property is obvious if we express G^{r0} as an iteration of unperturbed propagators g^{r0} separated by vertices T and T' . g^{r0} is an equilibrium quantity which is symmetric because of time reversal invariance: the symmetry of G^{r0} follows at once.

[‡] From the definition of the propagator, it follows that $D_{ij}^+(\omega) = D_{ji}^-(-\omega)$. Time reversal invariance provides the further relation $D_{ij}^+(\omega) = D_{ji}^+(\omega)$, from which equation 48 follows.

(as in equation 26, we here measure ρ_x and ρ_β from their respective chemical potentials). We note that equation 55 always corresponds to an increase of conductance, in contrast with the result found in the electrodes. We can say that the phonons open a 'new channel' for tunnelling. In practice, we are interested in biases eV of the order of the phonon

energy, which for MIM junctions are usually smaller than both the barrier height and Fermi energies: as in §3, we may neglect the energy dependence of W_{ij} , ρ_x and ρ_β . The conductance G_{in} is then given by

$$\frac{dG_{\text{in}}}{dV} = \frac{d^2 J_{\text{in}}}{dV^2} = \sum_{ij} \frac{2\pi e^3}{\hbar} W_{ij} \rho_\alpha \rho_\beta \sigma_{ij}(|eV|) \text{sgn } V \quad (56)$$

The second derivative of the current reflects the density of states of phonons, weighted by the matrix element W_{ij} . The corresponding lineshape in dG/dV is thus typically dissipative, much simpler than in §3.

One may wonder to what extent the expression (56) depends on the position of sites i and j inside the barrier. The density of states σ_{ij} and the matrix elements V_i, V_j may be affected by the proximity of metallic electrodes, but their variation should be moderate. The remaining factors in equation 50 involve the combination of electron propagators

$$G_{ai}^{r0}(\omega) G_{ib}^{r0}(\omega') \quad (57)$$

Since G^r is an exponentially decreasing function of positions in the barrier, such a product should be relatively independent of i . (To the extent that one neglects the frequency dependence of G^{r0} , one may show that equation 57 is strictly independent of i if the insulator is described in a tight binding model with nearest neighbour interactions). We thus conclude qualitatively that the efficiency of a barrier selfenergy is roughly independent of its location in the insulating layer. Note that if the electron factors do not depend on the difference $(i - j)$, the corresponding summation in equation 56 will favour phonons whose momentum has a small component perpendicular to the barrier -either at the center of the Brillouin zone or nearly parallel to the interface.

The preceding analysis could be extended to higher order in the electron-phonon interaction. One effect would be the replacement of G^0 by renormalized propagators G in the matrix element W_{ij} . Moreover, there would appear a collection of new terms, arising partly from new graphs in Σ^+ , partly from the expansion of the factors G^+ inside Σ^+ . Such terms describe processes in which several phonons are emitted in the inelastic tunnelling event. The analysis is complicated and somewhat academic: that is why we did not attempt it. Let us simply emphasize that, if required, such multiphonon effects can be studied with our formulation of the problem.

Having discussed in detail the inelastic current $\langle J_{\text{in}} \rangle$, we now return to the phonon correction $\langle \delta J_{\text{el}} \rangle$ to the elastic part of the current. From equation 18, we see that

$$\langle \delta J_{\text{el}} \rangle = \frac{e}{\hbar} \int \frac{d\omega}{2\pi} \delta |U^r|^2 \left(g_{\alpha z}^+ g_{\beta \beta}^- - g_{\alpha a}^- g_{\beta \beta}^+ \right) \quad (58)$$

To lowest order in the electron-phonon interaction. the correction $\delta |U^r|^2$ is equal to

$$\delta |U^r|^2 = \sum_{ij} T^2 T^2 \left(G_{ai}^{r0} \Sigma_{ij}^r G_{ib}^{r0} G_{ba}^{a0} + \text{cc} \right) \quad (59)$$

- We note that $D_{ij}^+ - D_{ij}^- = \nu_{ij} - D_{ij}^r = 2i\pi\sigma_{ij}||\omega| \rangle \text{sgn}(1)$, a relation which may serve to define σ_{ij} , and which is the counterpart for Bose excitations of the relation (14) valid for fermions.

Σ_{ij} is in turn given by equation 6. If i and j lie well inside the insulator, the functions $G_{ij}^{\pm}$ are negligible compared with G_{ij}^r , since they require an excursion in the electrodes, which implies a small evanescent factor. In that case the expression (6) reduces to

$$\Sigma_{ij}^r(\omega) \simeq iV_iV_j \int \frac{d\omega'}{2\pi} D_{ij}^+(\omega - \omega') G_{ij}^r(\omega') \quad (60)$$

$G_{ij}^r(\omega')$ is a slowly varying function of ω' (which varies at most on a scale of the order of the barrier height). It follows that $\Sigma_{ij}^r(\omega)$ is practically independent of ω : the current $\langle \delta J_{el} \rangle$ does not give rise to any structure in the characteristics $G(V)$. Experimentally, the barrier phonons are only seen in the inelastic current $\langle J_{in} \rangle$. The situation is quite different for the electrode phonons, which give comparable contributions to the elastic and inelastic currents. In that case, the 'elastic' part arises from the term $D_{ij}^r G_{ij}^-$ in equation 6: because of the discontinuity of G_{ij}^- at the Fermi level, the structure of D_{ij}^r is reflected in the selfenergy Σ_{ij}^r (which is a convolution of D^r and G^-): the correction $\langle \delta J_{ei} \rangle$ thus retains a structure as a function of bias. Inside the barrier, $G^{-\infty}$ is instead negligible, and this term disappears.

The preceding discussion breaks down if i and j lie too close to one electrode. In that case, G_{ij}^- is no longer negligible, and $\langle \delta J_{el} \rangle$ is of the same order as $\langle J_{in} \rangle$. The conductance $G(V)$ will be a mixture of the dissipative part (56) (which is positive and even) with a reactive term involving $\text{Im } \Sigma$ and $\text{Re } \Sigma$. This is another statement of the results obtained at the end of §3: if in equation 39 the site λ lies far from the electrodes, $\mathcal{G}^r(\mathbf{j})$ is real, and the current involves only $\text{Im } \Sigma$ (as in equation 44): if on the other hand i gets too close to an electrode, the phase of $\mathcal{G}_{i,\lambda}^{r0}$ begins to change, and the current acquires a term proportional to $\text{Re } \Sigma$. The transition from 'barrier' behaviour to 'electrode' behaviour is thus smooth: typically it extends over an extinction length of the evanescent wavefunctions.

In conclusion, we can say that in MIM junctions, selfenergy corrections to the conductance are larger in the barrier than in the electrodes, essentially because there is no destructive interference in the former case. In the barrier, one has a purely dissipative effect, yielding an even, positive correction to the conductance G , which may be interpreted as the opening of a new tunnelling channel. In the electrodes, instead, the effect is reduced by the destructive interference, which at the same time introduces a dispersive, odd part in $G(V)$; the even part may even become negative. Usually, this unpredictable behaviour is masked by barrier effects, even for the electrode phonons which find it more advantageous to penetrate the barrier, rather than interacting with the electrons in the metal itself. We have seen that the distinction between electrode and barrier structure was not sharp: one goes smoothly from one to the other. Finally, it should be emphasized that many of these conclusions are only valid for MIM junctions. While the general formalism remains correct for Schottky barriers, the assumptions of constant U , ρ , etc are no longer correct; moreover, the deformation of the barrier with the applied bias V becomes essential; finally, the interference effects in the electrodes become much weaker, as the wavelength of carriers at the Fermi level is not so different from the barrier thickness: the distinction between electrode and barrier processes is far less sharp. A detailed, quantitative analysis of phonon structures in Schottky barriers remains to be done. This problem is presently under study, using our formulation which provides the natural language in which to treat such problems.

12.0.5 Appendix

In §3, we introduced a propagator $\mathcal{G}^0$ for the region comprising one electrode and half the barrier, in the absence of any electron-phonon interaction. In this Appendix, we wish to illustrate the general behaviour of $\mathcal{G}^r$ by considering a specific example, namely a square barrier potential for a continuous model. The one particle hamiltonian is

$$H = -\frac{\hbar^2}{2m}\nabla^2 + V(x)$$

where the potential $V(x)$ is depicted in figure 2. The Green function $\mathcal{G}^{r0}(x, y)$ obeys the differential equation

$$\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x) - \omega\right)\mathcal{G}^{r0}(x, y; \omega) = \delta(x - y) \quad (\text{A.1})$$

Equation A. 1 must be solved with the boundary condition that $\mathcal{G}^{r0} = 0$ if x or $y = 0$ (since the wavefunctions vanish at the edge of a semi-infinite medium). $\mathcal{G}^o$ may be calculated by the methods of Schwinger (1961); for a change, we calculate it directly by solving equation A.1.

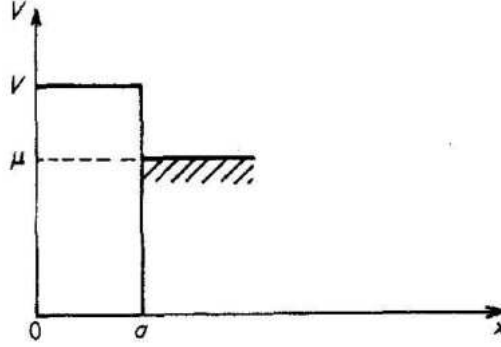


Figure 2. The model potential barrier described in the Appendix.

For $x < a$, the general solution of equation A. 1 is

$$\mathcal{G}^{r0}(x, y) = A_{\pm} \exp(Kx) + B_{\pm} \exp(-Kx) \quad (\text{A.2})$$

with $\hbar K = \{2m(V - \omega)\}^{1/2}$, the coefficients $A_{\pm}, B_{\pm}$ refer respectively to $x > y$ and $x < y$. For $x > a$, the general solution is

$$\mathcal{G}^{r0}(x, y) = \alpha \exp(iqx) + \beta \exp(-iqx) \quad (\text{A.3})$$

with $\hbar q = (2m\omega)^{1/2}$. We know that $\mathcal{G}^{r0}$ must be analytic when continued into the upper half ω plane: that implies immediately $\beta = 0$ (conversely, α would vanish for $\mathcal{G}^a$). The remaining coefficients are then determined by the conditions (i) $\mathcal{G}^{r0} = 0$ for $x = 0$, (ii) $\mathcal{G}^{r0}$ is continuous at $x = y$, while $\partial\mathcal{G}^0/\partial x$ has a discontinuity $-2m/\hbar^2$, (iii) $\mathcal{G}^{r0}$ and $\partial\mathcal{G}^{r0}/\partial x$ are continuous at $x = a$. The algebra is straightforward, and yields

$$\mathcal{G}^{r0}(x, y) = \begin{cases} -\frac{2m \sinh Kx}{\hbar^2} \frac{KD}{KD} [(K + iq) \exp\{K(y - 1)\} \\ + (K - iq) \exp\{K(a - y)\}] & x < y \\ -\frac{2m \sinh Ky}{\hbar^2} \frac{KD}{KD} [(K + iq) \exp\{K(x - a)\} \\ + (K - iq) \exp\{K(a - x)\}] & y < x < a \\ -\frac{2m \sinh Ky}{\hbar^2} \frac{\sin}{KD} 2K \exp\{iq(x - a)\} & a < x \end{cases} \quad (\text{A.4})$$

in which the denominator D is given by

$$D = (K + iq) \exp(-Ka) + (K - iq) \exp(Ka) \quad (\text{A.5})$$

In our case, $\exp(-Ka) \ll 1$ (the barrier transparency is small), so that the first term of equation A. 5 is negligible.

Let us assume $y \ll a$. We then see that for $x > a$, $\mathcal{G}^{r0}$ takes the form

$$\mathcal{G}^{r0}(x, y) = -\frac{2m}{\hbar^2} 2 \sinh Ky \frac{\exp(-Ka)}{K - iq} \exp\{iq(x - a)\} \quad (\text{A.6})$$

Equation A. 6 has all the features we expect: it is an oscillatory function of x , as discussed in §3; it is exponentially small, the factor $\exp(-Ka)$ describing the evanescence of the wavefunctions from the partition to the metal-insulator contact. The factor $(K - iq)$ in the denominator accounts for the impedance matching of the wavefunctions at the I-M contact; note that this factor contributes an extra phase to $\mathcal{G}$; a feature that will affect the shape of phonon structures. Usually, $q \gg K$, and that extra phase is $-\pi/2$, contributing an additional minus sign to the quantity (39).

If on the other hand $K(a - x) \geq 1$ (x well inside the insulator), the first term in the bracket of equation A. 4 is negligible, and we have

$$\mathcal{G}^{r0}(x, y) \simeq -\frac{2m}{\hbar^2} \frac{\sinh Ky}{K} \exp(-Kx) \quad (\text{A.7})$$

$\mathcal{G}^{r0}$ is then real, as claimed in §3. That accounts for the absence of interference, and for the fact that the corresponding contribution to the conductance is even and positive (ie purely dissipative). If finally $0 < K(a - x) \lesssim 1$ (electron in the insulator, but close to the metal), the first term in the bracket of equation A. 4 becomes important, and the phase of $\mathcal{G}^{r0}$ departs from zero; one thus goes smoothly from the barrier behaviour ($\mathcal{G}^{r0}$ real) to the electrode behaviour ($\mathcal{G}^{r0}$ oscillatory).

12.1 Easy illustrative examples

12.1.1 Example 1

12.1.2 Example 2

Part III

Examples and Solved Problems

13 Illustrative examples

13.1 Superconductive ring

13.1.1 Three-terminal ring

Model 0, kind of ring

Model 0

Consider

$$H_0 = \tau_z \left(\frac{(-i\partial_\vartheta + \tau_z \phi)^2}{2mR^2} - \mu \right) \quad (13.1)$$

The coupling to the superconducting electrodes is provided by the effective (energy-dependent) potential

$$V(\vartheta) = \delta(\vartheta)\Sigma_R(z) + \delta(\vartheta - \delta\vartheta)\Sigma_L(z), \quad (13.2)$$

$$\Sigma_{R/L}(z) = -\Gamma \frac{z - \Delta[e^{\pm i\varphi/2}\tau_+ + e^{\mp i\varphi/2}\tau_-]}{\sqrt{\Delta^2 - z^2}} \quad (13.3)$$

The density of states of the system is expressed as

$$\rho(\omega) = \rho^{(0)}(\omega) - \frac{1}{\pi} \frac{\partial}{\partial \omega} \log \det[1 - \hat{G}\hat{\Sigma}], \quad (13.4)$$

$$\hat{G} = \begin{pmatrix} g(0) & g(-\delta\vartheta) \\ g(\delta\vartheta) & g(0) \end{pmatrix}, \quad \hat{\Sigma} = \begin{pmatrix} \Sigma_R(z) & 0 \\ 0 & \Sigma_L(z) \end{pmatrix}. \quad (13.5)$$

One has

$$\bar{g}(\vartheta) = \frac{1}{2\pi} \sum_n \frac{e^{in\vartheta}}{z - \tau_z \left(\frac{(n + \tau_z \phi)^2}{2mR^2} - \mu \right)} \quad (13.6)$$

$$= \frac{mR^2\tau_z}{2\pi M(z)} \left(\sum_n \frac{e^{in\vartheta}}{n + \tau_z \phi + M(z)} - \sum_n \frac{e^{in\vartheta}}{n + \tau_z \phi - M(z)} \right) \quad (13.7)$$

$$= \frac{i\tau_z}{\omega_F} \left(\frac{e^{-in_F\vartheta} e^{-i\tau_z z\vartheta/\omega_F}}{1 - e^{-2\pi i\tau_z(\phi+z/\omega_F)}} - \frac{e^{in_F\vartheta} e^{i\tau_z z\vartheta/\omega_F}}{1 - e^{-2\pi i\tau_z(\phi-z/\omega_F)}} \right) e^{-i\tau_z \phi\vartheta}, \quad \vartheta \in [0, 2\pi]. \quad (13.8)$$

$$M(\tau_z z) = \sqrt{2m(\mu + \tau_z z)} R \simeq n_F + \tau_z z / \omega_F, \quad (13.9)$$

$$g(-\delta\vartheta) = \bar{g}(2\pi - \delta\vartheta) \quad (13.10)$$

where we used

$$\sum_n \frac{e^{in\vartheta}}{n + \tau_z \phi \pm M(z)} = \frac{2\pi i e^{-i(\tau_z \phi \pm M(z))\vartheta}}{1 - e^{-2\pi i(\tau_z \phi \pm M(z))}}, \quad \vartheta \in [0, 2\pi]$$

It follows that

$$\hat{G} = \begin{pmatrix} \bar{g}(0) & \bar{g}(2\pi - \delta\vartheta) \\ \bar{g}(\delta\vartheta) & \bar{g}(0) \end{pmatrix} = \frac{1}{\omega_F} \begin{pmatrix} f(0, \bar{z}/\bar{\omega}_F, \phi) & f(2\pi - \delta\vartheta, \bar{z}/\bar{\omega}_F, \phi) \\ f(\delta\vartheta, \bar{z}/\bar{\omega}_F, \phi) & f(0, \bar{z}/\bar{\omega}_F, \phi) \end{pmatrix}, \quad (13.11)$$

$$f(\delta\vartheta, \bar{z}/\bar{\omega}_F, \phi) = i\tau_z \left(\frac{e^{-in_F\delta\vartheta} e^{-i\tau_z \bar{z}\delta\vartheta/\bar{\omega}_F}}{1 - e^{-2\pi i\tau_z(\phi + \bar{z}/\bar{\omega}_F)}} - \frac{e^{in_F\delta\vartheta} e^{i\tau_z \bar{z}\delta\vartheta/\bar{\omega}_F}}{1 - e^{-2\pi i\tau_z(\phi - \bar{z}/\bar{\omega}_F)}} \right) e^{-i\tau_z\phi\vartheta}, \quad (13.12)$$

$$1 - \tilde{\Gamma} \begin{pmatrix} f(0, \bar{z}/\bar{\omega}_F, \phi) & f(2\pi - \delta\vartheta, \bar{z}/\bar{\omega}_F, \phi) \\ f(\delta\vartheta, \bar{z}/\bar{\omega}_F, \phi) & f(0, \bar{z}/\bar{\omega}_F, \phi) \end{pmatrix} \begin{pmatrix} \frac{e^{i\varphi/2}\tau_+ + e^{-i\varphi/2}\tau_- - \bar{z}}{\sqrt{1-\bar{z}^2}} & 0 \\ 0 & \frac{e^{-i\varphi/2}\tau_+ + e^{i\varphi/2}\tau_- - z}{\sqrt{1-\bar{z}^2}} \end{pmatrix}, \quad \tilde{\Gamma} = \Gamma/\omega_F \quad (13.13)$$

Andreev bound states of a single contact

$$f(0, \bar{z}/\bar{\omega}_F, \phi) = \frac{\sin(2\pi\bar{z}/\bar{\omega}_F)}{\cos(2\pi\phi) - \cos(2\pi\bar{z}/\bar{\omega}_F)} \quad (13.14)$$

Short ring approximation

In the case of large ω_F we can ignore the z -dependence of $\hat{G}$. One has

$$f(\delta\vartheta, 0, \phi) = -i \frac{\sin(n_F\delta\vartheta)}{\sin(\pi\phi)} e^{-i\tau_z\phi(\delta\vartheta-\pi)}, \quad f(2\pi - \delta\vartheta, 0, \phi) = i \frac{\sin(n_F\delta\vartheta)}{\sin(\pi\phi)} e^{i\tau_z\phi(\delta\vartheta-\pi)} \quad (13.15)$$

Model 1, ring

We consider the surface states of 3DTI. Let us chose $A_x(y) = -\frac{B_z}{2}y$, $A_y(x) = \frac{B_z}{2}x$, then

$$H = v\sigma_x p_x + ve\sigma_x A_x(y) + v\sigma_y p_y + ve\sigma_y A_y(x) + \frac{g\mu_B}{2} \mathbf{B} \cdot \boldsymbol{\sigma} \quad (13.16)$$

$$= v(\sigma_x \cos \varphi + \sigma_y \sin \varphi) p_r - \frac{v}{r} (\sin \varphi \sigma_x - \cos \varphi \sigma_y) p_\varphi + \frac{g\mu_B}{2} \mathbf{B} \cdot \boldsymbol{\sigma} - \frac{veB_z r}{2} (\sigma_x \sin \varphi - \sigma_y \cos \varphi), \quad (13.17)$$

here

$$\sigma_x \cos \varphi + \sigma_y \sin \varphi = e^{-i\varphi\sigma_z/2} \sigma_x e^{i\varphi\sigma_z/2}, \quad (13.18)$$

$$\sigma_x \sin \varphi - \sigma_y \cos \varphi = -e^{-i\varphi\sigma_z/2} \sigma_y e^{i\varphi\sigma_z/2}. \quad (13.19)$$

Define

$$\tilde{H} = \sqrt{r} e^{i\varphi\sigma_z/2} H e^{-i\varphi\sigma_z/2} \frac{1}{\sqrt{r}} \quad (13.20)$$

$$= v\sigma_x p_r + \frac{v}{r} \sigma_y p_\varphi + \frac{g\mu_B}{2} e^{i\varphi\sigma_z/2} \mathbf{B} \cdot \boldsymbol{\sigma} e^{-i\varphi\sigma_z/2} - \frac{veB_z r}{2} \quad (13.21)$$

Now the wave functions have to satisfy anti-periodic boundary conditions

$$\tilde{\Psi}(r, \varphi + 2\pi) = -\tilde{\Psi}(r, \varphi), \quad (13.22)$$

that is because

$$\Psi(r, \varphi + 2\pi) = \Psi(r, \varphi), \quad \Psi(r, \varphi) := e^{-i\varphi\sigma_z/2} \frac{1}{\sqrt{r}} \tilde{\Psi}(r, \varphi) \quad (13.23)$$

and respect the normalization

$$\int_0^{2\pi} d\varphi \int_{R_-}^{R_+} r dr |\Psi(r, \varphi)|^2 = \int_0^{2\pi} d\varphi \int_{R_-}^{R_+} dr |\tilde{\Psi}(r, \varphi)|^2 = 1. \quad (13.24)$$

Here R_+ and R_- are the radii of the ring. Now we assume the ring to be very thin $W := R_+ - R_- \ll [\text{other length scales}]$. In this connection, we assume $\tilde{\Psi}(r, \varphi)$ to be independent of r and replace it with its average value at $r = \frac{R_+ + R_-}{2}$:

$$\tilde{\Psi}(r, \varphi) \simeq \frac{1}{\sqrt{W}} \tilde{\psi}(\varphi), \quad (13.25)$$

such that

$$\int_0^{2\pi} d\varphi |\tilde{\psi}(\varphi)|^2 = 1, \quad \tilde{\psi}(\varphi + 2\pi) = -\tilde{\psi}(\varphi). \quad (13.26)$$

The approximate Hamiltonian for $\tilde{\psi}(\varphi)$ is

$$\tilde{h} = \Omega_F \left(\sigma_y p_\varphi - \frac{\Phi}{\Phi_0} \right) + \frac{g\mu_B}{2} e^{i\varphi\sigma_z/2} \mathbf{B} \cdot \boldsymbol{\sigma} e^{-i\varphi\sigma_z/2}, \quad (13.27)$$

where

$$\frac{\Phi}{\Phi_0} = \frac{eB_z R^2}{2}, \quad \Omega_F = \frac{v}{R}. \quad (13.28)$$

Case $B_x = B_y = 0$

First way

Let us introduce electron-hole states τ_z as a new space which is in tensor product with the previous states.

$$\tilde{h} = \tau_z \Omega_F \left(\tau_z \sigma_y p_\varphi - \frac{\Phi}{\Phi_0} \tau_z \right) + \frac{g\mu_B}{2} B_z \sigma_z \quad (13.29)$$

This is a 4x4 matrix.

Let us denote by overbar $\bar{G}$ a Green's function for only $\varphi \in [0, 2\pi]$,

Let us use the formula $G(\mathbf{x}, \mathbf{x}') = \sum_n \frac{\psi_n^*(\mathbf{x}') \psi_n(\mathbf{x})}{\lambda_n - \lambda}$, where $\psi_n(\mathbf{x})$ is a normalized eigenfunction of $\tilde{h}$. For this model $\psi(x) = \frac{1}{\sqrt{2\pi}} e^{i(n+1/2)\varphi}$ (???? what exactly is this formula???)

Let us denote $\Delta_Z := \frac{g\mu_B B_z}{2}$. The Green's function will be

$$\bar{G}(z, \varphi) = \frac{1}{2\pi} \sum_n \frac{e^{in\varphi+i\varphi/2}}{n+1/2 - \tau_z \Omega_F \left(\tau_z \sigma_y p_\varphi - \frac{\Phi}{\Phi_0} \tau_z \right) - \frac{g\mu_B}{2} B_z \sigma_z} = \quad (13.30)$$

$$= - \frac{e^{i\varphi/2}}{2\pi} \sum_n \frac{e^{in\varphi}}{\tau_z \sigma_y \left(-n - 1/2 + \tau_z \Omega_F \left(\tau_z \sigma_y p_\varphi - \frac{\Phi}{\Phi_0} \tau_z \right) + \Delta_Z \sigma_z \right)} \tau_z \sigma_y \quad (13.31)$$

$$= - \frac{e^{i\varphi/2}}{2\pi} \sum_n \frac{e^{in\varphi}}{\tau_z \sigma_y (-n - 1/2) + \Omega_F \tau_z p_\varphi - \tau_z \sigma_y \Omega_F \frac{\Phi}{\Phi_0} + \Delta_Z \tau_z \sigma_y \sigma_z} \tau_z \sigma_y \quad (13.32)$$

$$= - \frac{e^{i\varphi/2}}{2\pi} \sum_n \frac{e^{in\varphi}}{\tau_z \sigma_y (-n - 1/2) + \Omega_F \tau_z p_\varphi - \tau_z \sigma_y \Omega_F \frac{\Phi}{\Phi_0} + i\Delta_Z \tau_z \sigma_x} \tau_z \sigma_y \quad (13.33)$$

$$(13.34)$$

(???? what to do with p_φ ????)

Now, to make formulas shorter, let us introduce $R := [z + \Omega_F \frac{\Phi}{\Phi_0}]$, and we continue the transformation:

$$\bar{G}(z, \varphi) = - \frac{e^{i\varphi/2}}{2\pi \Omega_F} \sum_n \frac{e^{in\varphi}}{n+1/2 + i\frac{\Delta_z}{\Omega_F} \tau_z \sigma_x - \frac{R}{\Omega_F} \tau_z \sigma_y} \tau_z \sigma_y = \quad (13.35)$$

$$(13.36)$$

$$\left(\sum_n \frac{e^{in\vartheta}}{n+A} = \frac{2\pi i e^{-iA\vartheta}}{1 - e^{-2\pi i A}} \right) \quad (13.37)$$

$$(13.38)$$

$$= - \frac{i e^{i\varphi/2}}{\Omega_F} \frac{\exp \left(-i \left(-\frac{R}{\Omega_F} \tau_z \sigma_y + i \frac{\Delta_z}{\Omega_F} \tau_z \sigma_x \right) \varphi - i\varphi/2 \right)}{1 - e^{i\pi} \exp \left(-2\pi i \left(-\frac{R}{\Omega_F} \tau_z \sigma_y + i \frac{\Delta_z}{\Omega_F} \tau_z \sigma_x \right) \right)} \tau_z \sigma_y = \quad (13.39)$$

$$(13.40)$$

$$= - \frac{i}{\Omega_F} \frac{\exp \left(i \left(\frac{R}{\Omega_F} \tau_z \sigma_y - i \frac{\Delta_z}{\Omega_F} \tau_z \sigma_x \right) \varphi \right)}{1 + \exp \left(2\pi i \left(\frac{R}{\Omega_F} \tau_z \sigma_y - i \frac{\Delta_z}{\Omega_F} \tau_z \sigma_x \right) \right)} \tau_z \sigma_y \quad (13.41)$$

$$(13.42)$$

$$= - \frac{i}{\Omega_F} \frac{\exp \left(i \tau_z (R \sigma_y - i \Delta_Z \sigma_x) (\varphi - \pi) / \Omega_F \right)}{\exp \left(-\pi i \left(\frac{R}{\Omega_F} \tau_z \sigma_y - i \frac{\Delta_z}{\Omega_F} \tau_z \sigma_x \right) \right) + \exp \left(\pi i \left(\frac{R}{\Omega_F} \tau_z \sigma_y - i \frac{\Delta_z}{\Omega_F} \tau_z \sigma_x \right) \right)} \tau_z \sigma_y. \quad (13.43)$$

Now, let us transform with $e^{ia(\hat{n} \cdot \vec{\sigma})} = I \cos a + i(\hat{n} \cdot \vec{\sigma}) \sin a$, where for this case there will be

$$n := \left(\frac{R}{\sqrt{R^2 - \Delta_Z^2}}, \frac{i\Delta_Z}{\sqrt{R^2 - \Delta_Z^2}} \right), \quad a := \pi \sqrt{R^2 - \Delta_Z^2} / \Omega_F.$$

$$\exp\left(-\pi i\left(\frac{R}{\Omega_F}\tau_z\sigma_y - i\frac{\Delta_z}{\Omega_F}\tau_z\sigma_x\right)\right) + \exp\left(\pi i\left(\frac{R}{\Omega_F}\tau_z\sigma_y - i\frac{\Delta_z}{\Omega_F}\tau_z\sigma_x\right)\right) \quad (13.44)$$

$$= \exp\left(-\pi i\frac{\sqrt{R^2 - \Delta_Z^2}}{\Omega_F}\left(\frac{R}{\sqrt{R^2 - \Delta_Z^2}}\tau_z\sigma_y - i\frac{\Delta_z}{\sqrt{R^2 - \Delta_Z^2}}\tau_z\sigma_x\right)\right) + \quad (13.45)$$

$$+ \exp\left(\pi i\frac{\sqrt{R^2 - \Delta_Z^2}}{\Omega_F}\left(\frac{R}{\sqrt{R^2 - \Delta_Z^2}}\tau_z\sigma_y - i\frac{\Delta_z}{\sqrt{R^2 - \Delta_Z^2}}\tau_z\sigma_x\right)\right) = \quad (13.46)$$

$$= \cos(a)\hat{I} - i(n \cdot \sigma)\sin(a) + \cos(a)\hat{I} + i(n \cdot \sigma)\sin(a) = \quad (13.47)$$

$$= 2\cos(\pi\sqrt{R^2 - \Delta_Z^2}/\Omega_F) \quad (13.48)$$

We can also check in Wolfram that we have an identity matrix in the denominator:

```

1 Simplify[
2 MatrixExp[-I Pi (A ArrayFlatten[
3 TensorProduct[PauliMatrix[3], PauliMatrix[2]]] - I B ArrayFlatten[
4   TensorProduct[PauliMatrix[3], PauliMatrix[1]]])]
5 +
6 MatrixExp[ I Pi (A ArrayFlatten[ TensorProduct[PauliMatrix[3], PauliMatrix[2]]] - I B
   ArrayFlatten[TensorProduct[PauliMatrix[3], PauliMatrix[1]]])]
] // MatrixForm

```

Finally, we have

$$\bar{G}(z, \varphi) = -\frac{i}{2\Omega_F} \frac{\exp(i\tau_z(R\sigma_y - i\Delta_Z\sigma_x)(\varphi - \pi)/\Omega_F)}{\cos(\pi\sqrt{R^2 - \Delta_Z^2}/\Omega_F)} \tau_z\sigma_y. \quad (13.49)$$

And the whole Green's function for any angle is

$$G(z, \varphi) = \sum_n \Theta(\varphi - 2\pi n) \Theta(2\pi[n + 1] - \varphi) \bar{G}(z, \varphi - 2\pi n). \quad (13.50)$$

Let's plug angles $0, \frac{2\pi}{3}, \frac{4\pi}{3}$ to the total Green's function:

$$\mathcal{G} = \begin{pmatrix} G(z, 0) & G(z, 2\pi/3) & G(z, 4\pi/3) \\ G(z, -2\pi/3) & G(z, 0) & G(z, 2\pi/3) \\ G(z, -4\pi/3) & G(z, -2\pi/3) & G(z, 0) \end{pmatrix} \quad (13.51)$$

$$= -\frac{i}{2\Omega_F} \frac{1}{\cos(\pi\sqrt{[z + \Omega_F \frac{\Phi}{\Phi_0}]^2 - \Delta_Z^2}/\Omega_F)} \quad (13.52)$$

$$\times \begin{pmatrix} e^{-i\tau_z([z + \Omega_F \frac{\Phi}{\Phi_0}]\sigma_y - i\Delta_Z\sigma_x)\pi/\Omega_F} & e^{-i\tau_z([z + \Omega_F \frac{\Phi}{\Phi_0}]\sigma_y - i\Delta_Z\sigma_x)\pi/3\Omega_F} & e^{i\tau_z([z + \Omega_F \frac{\Phi}{\Phi_0}]\sigma_y - i\Delta_Z\sigma_x)\pi/3\Omega_F} \\ e^{i\tau_z([z + \Omega_F \frac{\Phi}{\Phi_0}]\sigma_y - i\Delta_Z\sigma_x)\pi/3\Omega_F} & e^{-i\tau_z([z + \Omega_F \frac{\Phi}{\Phi_0}]\sigma_y - i\Delta_Z\sigma_x)\pi/\Omega_F} & e^{-i\tau_z([z + \Omega_F \frac{\Phi}{\Phi_0}]\sigma_y - i\Delta_Z\sigma_x)\pi/3\Omega_F} \\ e^{-i\tau_z([z + \Omega_F \frac{\Phi}{\Phi_0}]\sigma_y - i\Delta_Z\sigma_x)\pi/3\Omega_F} & e^{i\tau_z([z + \Omega_F \frac{\Phi}{\Phi_0}]\sigma_y - i\Delta_Z\sigma_x)\pi/3\Omega_F} & e^{-i\tau_z([z + \Omega_F \frac{\Phi}{\Phi_0}]\sigma_y - i\Delta_Z\sigma_x)\pi/\Omega_F} \end{pmatrix} \tau_z\sigma_y. \quad (13.53)$$

Now we couple the superconductors via

$$\Sigma_{SC}(\omega, \{\varphi_j\}) = -\Gamma_S \begin{pmatrix} \frac{\omega - \Delta_0 e^{i\varphi_1\tau_z/2} \tau_x e^{-i\varphi_1\tau_z/2}}{\sqrt{\Delta_0^2 - \omega^2}} & 0 & 0 \\ 0 & \frac{\omega - \Delta_0 e^{i\varphi_2\tau_z/2} \tau_x e^{-i\varphi_2\tau_z/2}}{\sqrt{\Delta_0^2 - \omega^2}} & 0 \\ 0 & 0 & \frac{\omega - \Delta_0 e^{i\varphi_3\tau_z/2} \tau_x e^{-i\varphi_3\tau_z/2}}{\sqrt{\Delta_0^2 - \omega^2}} \end{pmatrix}. \quad (13.54)$$

Calculate the spectrum from $\det[d] = 0$, where $d = 1 - \mathcal{G}\Sigma_{SC}$. Consider $\varphi_3 = 0$ and vary φ_1 and φ_2 . What happens when $\varphi_1 = \varphi_2$ and $\varphi_1 = -\varphi_2$?

$$B_x = B_y = 0$$

$$\tilde{h} = \tau_z \Omega_F \left(\tau_z \sigma_y p_\varphi - \frac{\Phi}{\Phi_0} \tau_z \right) + \frac{g\mu_B}{2} B_z \sigma_z \quad (13.55)$$

$$\bar{G}(z, \varphi) = -\frac{e^{i\varphi/2}}{2\pi\Omega_F} \sum_n \frac{e^{in\varphi}}{n+1/2 + i\frac{g\mu_B B_z}{2\Omega_F} \tau_z \sigma_x - \frac{[z+\Omega_F \frac{\Phi}{\Phi_0}]}{\Omega_F} \tau_z \sigma_y} \tau_z \sigma_y \quad (13.56)$$

$$= -\frac{1}{\Omega_F} \frac{i e^{i\left(\frac{[z+\Omega_F \frac{\Phi}{\Phi_0}]}{\Omega_F} \tau_z \sigma_y - i\frac{g\mu_B B_z}{2\Omega_F} \tau_z \sigma_x\right)\varphi}}{1 + e^{2\pi i\left(\frac{[z+\Omega_F \frac{\Phi}{\Phi_0}]}{\Omega_F} \tau_z \sigma_y - i\frac{g\mu_B B_z}{2\Omega_F} \tau_z \sigma_x\right)}} \tau_z \sigma_y \quad (13.57)$$

$$= -\frac{i}{2\Omega_F} \frac{e^{i\tau_z([z+\Omega_F \frac{\Phi}{\Phi_0}]\sigma_y - i\Delta_Z \sigma_x)(\varphi-\pi)/\Omega_F}}{\cos(\pi\sqrt{[z+\Omega_F \frac{\Phi}{\Phi_0}]^2 - \Delta_Z^2/\Omega_F})} \tau_z \sigma_y \quad (13.58)$$

$$= \frac{1}{2\Omega_F^2} \frac{([z+\Omega_F \frac{\Phi}{\Phi_0}] + \Delta_Z \sigma_z) \sin(A(\varphi-\pi)) - i\Omega_F A \tau_z \sigma_y \cos(A(\varphi-\pi)/\Omega_F)}{A \cos(\pi A)} \quad (13.59)$$

where $\Delta_Z = \frac{g\mu_B B_z}{2}$. Also, φ here is an angle, don't confuse it with phase in Σ (change notation later). Here the overbar implies that this object is defined for $\varphi \in [0, 2\pi]$, such that

$$G(z, \varphi) = \sum_n \Theta(\varphi - 2\pi n) \Theta(2\pi[n+1] - \varphi) \bar{G}(z, \varphi - 2\pi n). \quad (13.60)$$

One has

$$\mathcal{G} = \begin{pmatrix} G(z, 0) & G(z, 2\pi/3) & G(z, 4\pi/3) \\ G(z, -2\pi/3) & G(z, 0) & G(z, 2\pi/3) \\ G(z, -4\pi/3) & G(z, -2\pi/3) & G(z, 0) \end{pmatrix} \quad (13.61)$$

$$= -\frac{i}{2\Omega_F} \frac{1}{\cos(\pi\sqrt{[z+\Omega_F \frac{\Phi}{\Phi_0}]^2 - \Delta_Z^2/\Omega_F})} \times \begin{pmatrix} e^{-i\tau_z([z+\Omega_F \frac{\Phi}{\Phi_0}]\sigma_y - i\Delta_Z \sigma_x)\pi/\Omega_F} & e^{-i\tau_z([z+\Omega_F \frac{\Phi}{\Phi_0}]\sigma_y - i\Delta_Z \sigma_x)\pi/3\Omega_F} & e^{i\tau_z([z+\Omega_F \frac{\Phi}{\Phi_0}]\sigma_y - i\Delta_Z \sigma_x)\pi/3\Omega_F} \\ e^{i\tau_z([z+\Omega_F \frac{\Phi}{\Phi_0}]\sigma_y - i\Delta_Z \sigma_x)\pi/3\Omega_F} & e^{-i\tau_z([z+\Omega_F \frac{\Phi}{\Phi_0}]\sigma_y - i\Delta_Z \sigma_x)\pi/\Omega_F} & e^{-i\tau_z([z+\Omega_F \frac{\Phi}{\Phi_0}]\sigma_y - i\Delta_Z \sigma_x)\pi/3\Omega_F} \\ e^{-i\tau_z([z+\Omega_F \frac{\Phi}{\Phi_0}]\sigma_y - i\Delta_Z \sigma_x)\pi/3\Omega_F} & e^{i\tau_z([z+\Omega_F \frac{\Phi}{\Phi_0}]\sigma_y - i\Delta_Z \sigma_x)\pi/3\Omega_F} & e^{-i\tau_z([z+\Omega_F \frac{\Phi}{\Phi_0}]\sigma_y - i\Delta_Z \sigma_x)\pi/\Omega_F} \end{pmatrix} \tau_z \sigma_y. \quad (13.63)$$

Now we couple the superconductors via

$$\Sigma_{SC}(\omega, \{\varphi_j\}) = -\Gamma_S \begin{pmatrix} \frac{\omega - \Delta_0 e^{i\varphi_1 \tau_z/2} \tau_x e^{-i\varphi_1 \tau_z/2}}{\sqrt{\Delta_0^2 - \omega^2}} & 0 & 0 \\ 0 & \frac{\omega - \Delta_0 e^{i\varphi_2 \tau_z/2} \tau_x e^{-i\varphi_2 \tau_z/2}}{\sqrt{\Delta_0^2 - \omega^2}} & 0 \\ 0 & 0 & \frac{\omega - \Delta_0 e^{i\varphi_3 \tau_z/2} \tau_x e^{-i\varphi_3 \tau_z/2}}{\sqrt{\Delta_0^2 - \omega^2}} \end{pmatrix}. \quad (13.64)$$

Let us consider case $\Phi=0$ and in another model $\Delta_z \neq 0$. We will have

$$\bar{G}(z, \varphi) = -\frac{i}{2\Omega_F} \frac{e^{i\tau_z(z\sigma_y - i\Delta_Z \sigma_x)(\varphi-\pi)/\Omega_F}}{\cos(\pi\sqrt{z^2 - \Delta_Z^2/\Omega_F})} \tau_z \sigma_y. \quad (13.65)$$

Calculate the spectrum from $\det[d] = 0$, where $d = 1 - \mathcal{G}\Sigma_{SC}$. Consider $\varphi_3 = 0$ and vary φ_1 and φ_2 . What happens when $\varphi_1 = \varphi_2$ and $\varphi_1 = -\varphi_2$?

$$B_x = B_y = 0$$

$$\tilde{h} = \Omega_F \left(\sigma_y p_\varphi - \frac{\Phi}{\Phi_0} \right) + \frac{g\mu_B}{2} B_z \sigma_z \quad (13.66)$$

$$G(z, \varphi) = \frac{1}{2\pi} \sum_n \frac{e^{i(n+1/2)\varphi}}{z + \Omega_F \frac{\Phi}{\Phi_0} - \Omega_F \sigma_y (n + 1/2) - \frac{g\mu_B}{2} B_z \sigma_z} \quad (13.67)$$

$$= \frac{1}{2\pi} \left(z + \Omega_F \frac{\Phi}{\Phi_0} + \Omega_F \sigma_y p_\varphi + \frac{g\mu_B}{2} B_z \sigma_z \right) \sum_n \frac{e^{i(n+1/2)\varphi}}{\left(z + \Omega_F \frac{\Phi}{\Phi_0} \right)^2 - \Omega_F^2 (n + 1/2)^2 - \frac{g^2 \mu_B^2}{4} B_z^2} = \quad (13.68)$$

$$= \frac{1}{2\pi} \frac{\left(z + \Omega_F \frac{\Phi}{\Phi_0} + \Omega_F \sigma_y p_\varphi + \frac{g\mu_B}{2} B_z \sigma_z \right) e^{i\varphi/2}}{\Omega_F^2} \sum_n \frac{e^{in\varphi}}{\frac{\left(z + \Omega_F \frac{\Phi}{\Phi_0} \right)^2 - \frac{g^2 \mu_B^2}{4} B_z^2}{\Omega_F^2} - (n + 1/2)^2} = \quad (13.69)$$

$$= -\frac{1}{2\pi} \frac{\left(z + \Omega_F \frac{\Phi}{\Phi_0} + \Omega_F \sigma_y p_\varphi + \frac{g\mu_B}{2} B_z \sigma_z \right) e^{i\varphi/2}}{\Omega_F^2} \sum_n \frac{e^{in\varphi}}{\left(n - A + \frac{1}{2} \right) \left(n + A + \frac{1}{2} \right)}. \quad (13.70)$$

So let us focus on the sum

$$S(A, \varphi) := \sum_n \frac{e^{in\varphi}}{(n + A_1)(n + A_2)}, \quad \varphi \in \{0, 2\pi\}, \quad (13.71)$$

where

$$A_1 := -A + 1/2; \quad A_2 := A + 1/2; \quad A := \sqrt{\frac{\left(z + \Omega_F \frac{\Phi}{\Phi_0} \right)^2 - \frac{g^2 \mu_B^2}{4} B_z^2}{\Omega_F^2}}.$$

I don't assume that the argument of the square root is positive, later for spectral density we will consider a case of imaginary A . From this, we know that $A_1 + A_2 = 1$; $A_2 - A_1 = 2A$.

Now, let us use the formula from Matsubara summations:

$$\sum_n \frac{e^{in\varphi}}{(n + B_1)(n + B_2)} = \begin{cases} -\oint_{\gamma} \frac{e^{i\varphi\zeta}}{(\zeta + B_1)(\zeta + B_2)} \frac{d\zeta}{e^{2\pi i\zeta} - 1} & \varphi > 0 \\ -\oint_{\gamma} \frac{e^{i\varphi\zeta}}{(\zeta + B_1)(\zeta + B_2)} \frac{e^{2\pi i\zeta} d\zeta}{e^{2\pi i\zeta} - 1} & \varphi < 0, \end{cases} \quad (13.72)$$

where γ is a set of two contours around poles on points $-B_1$ and $-B_2$. We will have

$$\sum_n \frac{e^{in\varphi}}{(n+A_1)(n+A_2)} \Big|_{\varphi>0} = -2\pi i \left[\frac{e^{-i\varphi A_1}}{(-A_1+A_2)(e^{-2\pi i A_1}-1)} + \frac{e^{-i\varphi A_2}}{(-A_2+A_1)(e^{-2\pi i A_2}-1)} \right] = \quad (13.73)$$

$$= \frac{-2\pi i}{2A} \left[\frac{e^{i\varphi A} e^{-i\varphi/2}}{(e^{2\pi i A} e^{-2\pi i/2}-1)} - \frac{e^{-i\varphi A} e^{-i\varphi/2}}{(e^{-2\pi i A} e^{-2\pi i/2}-1)} \right] = \quad (13.74)$$

$$= \frac{\pi i e^{-i\varphi/2}}{A} \left[\frac{e^{i\varphi A}}{e^{2\pi i A}+1} - \frac{e^{-i\varphi A}}{e^{-2\pi i A}+1} \right] = \quad (13.75)$$

$$= \frac{\pi i e^{-i\varphi/2}}{A} \cdot 2i \operatorname{Im} \left[\frac{e^{i\varphi A}}{e^{2\pi i A}+1} \right] = \quad (13.76)$$

$$= -\frac{\pi e^{-i\varphi/2}}{A} \cdot 2 \operatorname{Im} \left[e^{-\pi i A} \frac{e^{i\varphi A/2}}{(e^{\pi i A}+e^{-\pi i A})/2} \right] = \quad (13.77)$$

$$= -\frac{\pi e^{-i\varphi/2}}{A} \frac{\operatorname{Im} e^{i(\varphi-\pi)A}}{\cos(\pi A)} = \quad (13.78)$$

$$= -\frac{\pi e^{-i\varphi/2}}{A} \frac{\sin((\varphi-\pi)A)}{\cos(\pi A)}. \quad (13.79)$$

The result is:

$$G(z, \varphi) = -\frac{1}{2\pi} \frac{\left(z + \Omega_F \frac{\Phi}{\Phi_0} + \Omega_F \sigma_y p_\varphi + \frac{g\mu_B}{2} B_z \sigma_z \right) e^{i\varphi/2}}{\Omega_F^2} \left(-\frac{\pi}{A} \right) e^{-i\varphi/2} \frac{\sin((\varphi-\pi)A)}{\cos(\pi A)} = \quad (13.80)$$

$$= \frac{\left(z + \Omega_F \frac{\Phi}{\Phi_0} + \Omega_F \sigma_y p_\varphi + \frac{g\mu_B}{2} B_z \sigma_z \right) \sin((\varphi-\pi)A)}{2A\Omega_F^2 \cos(\pi A)} \quad \begin{array}{l} \text{act on sum} \\ p_\varphi = -i \frac{\partial}{\partial \varphi} \end{array} \quad (13.81)$$

$$= \frac{\left(z + \Omega_F \frac{\Phi}{\Phi_0} + \frac{g\mu_B}{2} B_z \sigma_z \right) \sin((\varphi-\pi)A) - i\Omega_F \sigma_y A \cos((\varphi-\pi)A)}{2A\Omega_F^2 \cos(\pi A)} \quad (13.82)$$

where $A := \sqrt{\frac{(z + \Omega_F \frac{\Phi}{\Phi_0})^2 - \frac{g^2 \mu_B^2}{4} B_z^2}{\Omega_F^2}}$, $\Phi_3 := \frac{\Phi}{\Phi_0} = \frac{eB_z R^2}{2}$, $\Omega_F = \frac{v}{R}$.

Let us simplify this equation:

$$A := \sqrt{\frac{(z + \Omega_F \frac{\Phi}{\Phi_0})^2 - \frac{g^2 \mu_B^2}{4} B_z^2}{\Omega_F^2}} = \sqrt{\frac{(z + \Omega_F \frac{\Phi}{\Phi_0})^2 - \frac{g^2 \mu_B^2}{4} \frac{e^2 B_z^2 R^4}{2^2}}{\Omega_F^2}} = \frac{\sqrt{(z + \Omega_F \Phi_3)^2 - S^2 \Phi_3^2}}{\Omega_F},$$

$$S := \frac{g\mu_B}{eR^2}; \quad \frac{g\mu_B}{2} B_z = \frac{g\mu_B}{eR^2} \frac{eB_z R^2}{2} = S\Phi_3.$$

$$G_{11/22}(z, \varphi) = \frac{(z + \Omega_F \Phi_3 \pm S\Phi_3) \sin \left((\varphi - \pi) \sqrt{(z + \Omega_F \Phi_3)^2 - S^2 \Phi_3^2} / \Omega_F \right)}{2\sqrt{(z + \Omega_F \Phi_3)^2 - S^2 \Phi_3^2} \Omega_F \cos \left(\pi \sqrt{(z + \Omega_F \Phi_3)^2 - S^2 \Phi_3^2} / \Omega_F \right)} \quad (13.83)$$

$$\operatorname{tr} G(z, \varphi) = \frac{(z + \Omega_F \Phi_3) \sin \left((\varphi - \pi) \sqrt{(z + \Omega_F \Phi_3)^2 - S^2 \Phi_3^2} / \Omega_F \right)}{\sqrt{(z + \Omega_F \Phi_3)^2 - S^2 \Phi_3^2} \Omega_F \cos \left(\pi \sqrt{(z + \Omega_F \Phi_3)^2 - S^2 \Phi_3^2} / \Omega_F \right)}, \quad (13.84)$$

and we have here two parameters by a setup: Ω_F and S , and two arguments: z and φ .

(The problem is that for $\Omega=0$ we don't have spectrum, but from Hamiltonian even with $\Omega=0$ the spectrum is fine)

Now, let us plot the function:

(see the notebook)

The spectral density (????)

Let us check the spectral density:

$$\rho(E) = -\frac{1}{\pi} \text{Im tr } G(E) = 0 \quad (13.85)$$

On another side,

$$\tilde{h} |\psi_E\rangle \equiv \Omega_F \left(\sigma_y p_\varphi - \frac{\Phi}{\Phi_0} \right) + \frac{g\mu_B}{2} B_z \sigma_z = E |\psi_E\rangle \quad (13.86)$$

$$\|\tilde{h} - \mathbb{1}E\| = 0$$

$$\det \begin{pmatrix} \frac{g\mu_B}{2} B_z - E - \Omega_F \frac{\Phi}{\Phi_0} & -i\Omega_F p_\varphi \\ i\Omega_F p_\varphi & -\frac{g\mu_B}{2} B_z - E - \Omega_F \frac{\Phi}{\Phi_0} \end{pmatrix} = 0$$

$$(E + \Omega_F \frac{\Phi}{\Phi_0})^2 - (\frac{g\mu_B}{2} B_z)^2 - \Omega_F^2 p_\varphi^2 = 0$$

$$E = -\Omega_F \frac{\Phi}{\Phi_0} \pm \sqrt{(\frac{g\mu_B}{2} B_z)^2 + \Omega_F^2 p_\varphi^2} =$$

$$= -\Omega_F \frac{\Phi}{\Phi_0} \pm \sqrt{(\frac{g\mu_B}{2} B_z)^2 + \Omega_F^2 \left(n + \frac{1}{2}\right)^2} \equiv$$

$$\equiv -\Omega_F \Phi_3 \pm \sqrt{S^2 \Phi_3^2 + \Omega_F^2 \left(n + \frac{1}{2}\right)^2}$$

(?? why from the Hamiltonian spectrum always exit, but from Green's function - only under some conditions (square root is imaginary)???)

13.1.2 Two-termial SC ring

(есть графики, напишу)

13.1.3 Two-termial SC ring with $H = p^2/2m + \dots$

$$G = \frac{1}{2\pi} \sum_n \frac{1}{z - \tau_z \left(\left(\frac{2\pi n}{L} \right)^2 \frac{1}{2m} + \alpha \frac{2\pi n}{L} \sigma_x - \mu \right)}$$

14 Examples for practice

(???)

Part IV

Other Methods

15 Boundary Green's function approach for spinful single-channel and multichannel Majorana nanowires By M. Alvarado, A. Iks, A. Zazunov, R. Egger, A. Levy Yeyati

The boundary-Green's function (bGF) approach has been established as a powerful theoretical technique for computing the transport properties of tunnel-coupled hybrid nanowire devices. Such nanowires may exhibit topologically nontrivial superconducting phases with Majorana bound states at their boundaries. We introduce a general method for computing the bGF of spinful multichannel lattice models for such Majorana nanowires, where the bGF is expressed in terms of the roots of a secular polynomial evaluated in complex momentum space. In many cases, those roots, and thus the bGF, can be accurately described by simple analytical expressions, while otherwise our approach allows for the numerically efficient evaluation of bGFs. We show that from the behavior of the roots, many physical quantities of key interest can be inferred, e.g., the value of bulk topological invariants, the energy dependence of the local density of states, or the spatial decay of subgap excitations. We apply the method to single- and two-channel nanowires of symmetry class D or DIII. In addition, we study the spectral properties of multiterminal Josephson junctions made out of such Majorana nanowires.

15.1 Concluding remarks

In the present work, we have generalized the boundary Green's function approach of Refs. [26, 61] to quasi-1D spinful models of Majorana nanowires. For single-channel class D and class DIII wire models, we have obtained an analytical understanding of the behavior of the roots of the corresponding secular polynomial in complex momentum space. This advance helps physical intuition and allows for a practical and numerically efficient method for computing the bGF, and thereby also physical observables. The method has also been extended to spinful multichannel models, where it appears to allow for more efficient numerical bGF calculations than the alternative recursive technique [60, 77]. Let us remark that the computational complexity of the method is only limited by the ability to evaluate the roots of a polynomial. Typically, the numerical demands are therefore much smaller than those for a recursive calculation of the bGF.

Given the efficient construction of the bGF put forward in this work, one can now apply the general bGF approach [26] to study the transport properties of many different hybrid devices composed of Majorana nanowires and/or conventional metals or superconductor electrodes. In Sec. 15.7, we have provided two examples for such devices, namely phase-biased trijunctions of TS wires and TRITOPS-TS junctions. In both cases, we have carried out an analysis of the subgap Andreev (or Majorana) state dispersion at zero temperature.

We believe that this approach offers many interesting perspectives for future research. In particular, one can study nonequilibrium transport properties away from the linear-response regime, and one can also include electron-electron or electron-phonon effects, at least on a perturbative level. We are confident that the results of our work can also be helpful for the interpretation of transport experiments carried out on hybrid devices containing nanowires with topologically nontrivial superconducting phases.

15.2 Introduction

The interest in proximitized nanostructures where topological superconductor phases could be engineered is continuing to grow [1, 2, 3, 4, 5, 6, 7, 8]. In particular, the case of one-dimensional (1D) semiconducting hybrid nanowires with strong Rashba spin-orbit interaction has been intensely studied as a potential route towards the generation of Majorana bound states (MBSs) [9, 10, 11, 12, 13, 14, 15, 16, 17]. Such states are of high interest for topological quantum information processing applications [4]. While a phase with broken time-reversal symmetry (class D) can be expected for the cited nanowire experiments because of the presence of a magnetic Zeeman field (we use the abbreviation ‘TS’ for such topological superconductors below), a time-reversal invariant topological superconductor (TRITOPS) phase has been predicted from related wire constructions [18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35]. The TRITOPS phase has symmetry class DIII and is still awaiting experimental tests. The interest in hybrid nanowires goes well beyond the generation of topological phases. For instance, recent microwave spectroscopy experiments have investigated the role of spin-orbit coupling effects on the formation of Andreev bound states [36].

The physics of devices made from different types of nanowires coupled by tunneling contacts has been explored by a variety of theoretical models and techniques [1, 2, 3, 5]. On one hand, minimal models restrict the Hilbert space to include only subgap bound states. This key simplification then allows for analytical progress (see, e.g., Refs. [37, 38] for early contributions). On the other hand, microscopic models aim for a more detailed understanding of how material properties can influence transport observables (see, e.g., Refs. [39, 40, 41, 42, 43, 44, 45, 46]). Recent works along this line have studied the electrostatic potential profile along the nanowire [47, 48, 49] and the effects of disorder on the phase diagram [50, 51]. However, the solution of such microscopic models requires information about many model parameter values and generally can be obtained only by performing a detailed numerical analysis. In this context, theoretical approaches of *intermediate* complexity are of high interest. Such a framework allows one to describe transport properties by taking into account both subgap and continuum states while keeping the algebra sufficiently simple so as to permit analytical progress. The scattering matrix formalism is a widely known representative for this type of approach (see, e.g., Refs. [52, 53, 54, 55, 56, 57, 58]). The present paper will employ the complementary boundary Green’s-function (bGF) method [26, 60, 61, 62, 63, 64], which is particularly useful for analyzing nonequilibrium transport properties in different types of hybrid nanojunctions. The bGF approach also allows one to examine other electronic properties such as the tunneling density of states (DoSs) or the bulk-boundary correspondence expected for topological phases [8, 66, 67]. Furthermore, electron-phonon and/or electron-electron interaction effects can in principle also be taken into account.

In the present paper, we extend and generalize the bGF approach for 1D or quasi-1D proximitized nanowires, which has been introduced in Refs. [26, 60, 61, 62, 63], along several directions. First, we demonstrate that a bGF construction in terms of the roots of a secular equation extended to complex momenta (as discussed in Ref. [26] for the Kitaev chain model) can be generalized to arbitrary spinful multichannel (i.e., quasi-1D) nanowires with topologically nontrivial superconducting phases. In particular, by studying the evolution of the roots in the complex momentum plane under the variation of model parameters, one can readily detect topological transitions, determine bulk topological invariants, or compute the local density of states as a function of energy for translationally invariant cases. In addition, the same roots determine the bGF and thereby give access to the transport properties of devices made from tunnel-coupled (semi-infinite or finite-length) nanowires. In particular, their knowledge also gives access to the spatial decay profile of Majorana states.

Below we investigate the roots and the corresponding bGFs for two widely used spinful single-channel nanowire models harboring topologically nontrivial phases. First, we study TS

wires with broken time-reversal invariance using the model by Lutchyn *et al.* [68] and by Oreg *et al.* [69]. Second, we consider TRITOPS wires using the model of Zhang *et al.* [21]. Quasi-1D multichannel models in class D or class DIII are then constructed by coupling several wires of the respective symmetry class by tunnel couplings. We show that also such multichannel models can be efficiently tackled by our bGF method. As application, we will discuss the Josephson current-phase relation both for a multiterminal junction composed of three tunnel-coupled TS wires and for a TRITOPS-TS Josephson junction.

The remainder of this article is organized as follows. In Sec. 15.3, we describe a general formalism for analyzing 1D or quasi-1D lattice models of proximitized nanowires, where we only assume that the hopping amplitudes in the corresponding tight-binding model are of finite range. We show that the real-space bulk Green's function (GF) adopts a compact expression in terms of the roots of the secular polynomial of the bulk Hamiltonian extended into complex momentum space. We also show how the boundary GF can be obtained from the bulk GF by solving a Dyson equation, and we discuss general properties of the corresponding roots. In Sec. 15.4, we consider a discretized version of the single-channel class D model of Refs. [68, 69]. We introduce a simple ansatz for the respective roots in the trivial and in the topological phase. This ansatz allows us to obtain analytical insights about the bulk spectral density and the spatial variation of MBSs. In Sec. 15.5, we extend the analysis to a two-channel model describing two coupled class D wires, where we can study spin-orbit interaction effects in multichannel nanowires [70]. The phase diagram and the spectral density of this model show a richer behavior than in the single-channel case. In Sec. 15.6, we apply our methods to single- and multichannel models for TRITOPS wires. Finally, in Sec. 15.7, we study the Josephson effect and the formation of Andreev bound states in phase-biased multiterminal TS junctions and for TRITOPS-TS junctions. We finally offer some conclusions in Sec. 15.1. Technical details have been delegated to two appendices. We often use units with $\hbar = 1$ and focus on the zero-temperature limit throughout.

15.3 Boundary Green's function

A central aim of the present work is to construct the bGF for different hybrid nanowire models which are described by a bulk Hamiltonian of the form

$$H_{\text{bulk}} = \frac{1}{2} \sum_k \hat{\Psi}_k^\dagger \hat{\mathcal{H}}(k) \hat{\Psi}_k, \quad (15.1)$$

corresponding to an infinitely long and translationally invariant (quasi-)1D chain with lattice spacing a . Here, $\hat{\mathcal{H}}(k)$ is an $N \times N$ Bogoliubov-de Gennes (BdG) Hamiltonian in reciprocal space, and the $\hat{\Psi}_k$ are fermionic Nambu spinor fields. Specific examples for these spinor fields will be given in the subsequent sections. The number N may include the Nambu index, the spin degree of freedom, and channel indices for multichannel models. Using $\hat{\mathcal{H}}(k+2\pi/a) = \hat{\mathcal{H}}(k)$, the BdG Hamiltonian can be expanded in a Fourier series, $\hat{\mathcal{H}}(k) = \sum_n \hat{\mathcal{V}}_n e^{inka}$, where Hermiticity implies $\hat{\mathcal{V}}_{-n} = \hat{\mathcal{V}}_n^\dagger$. For simplicity, we here consider only models with nearest-neighbor hopping, $\hat{\mathcal{V}}_n = 0$ for $|n| > 1$, but the generalization to arbitrary finite-range hopping amplitudes is straightforward.

The retarded bulk GF of the infinite chain is defined as

$$\hat{G}^R(k, \omega) = [\omega + i0^+ - \hat{\mathcal{H}}(k)]^{-1}, \quad (15.2)$$

where the $N \times N$ matrix structure is indicated by the hat notation. In real space representation, the GF has the components (j and j' are lattice site indices)

$$\hat{G}_{jj'}^R(\omega) = \frac{a}{2\pi} \int_{-\pi/a}^{\pi/a} dk e^{i(j-j')ka} \hat{G}^R(k, \omega). \quad (15.3)$$

By the identification $z = e^{ika}$, this integral is converted into a complex contour integral,

$$\hat{G}_{jj'}^R(\omega) = \frac{1}{2\pi i} \oint_{|z|=1} \frac{dz}{z} z^{j-j'} \hat{G}^R(z, \omega). \quad (15.4)$$

Introducing the roots $z_n(\omega)$ of the secular polynomial in the complex- z plane,

$$P(z, \omega) = \det [\omega - \hat{\mathcal{H}}(z)] = \frac{1}{z^N} \prod_{n=1}^{2N} [z - z_n(\omega)], \quad (15.5)$$

the contour integral (15.4) can be written as a sum over the residues of all roots inside the unit circle:

$$\hat{G}_{jj'}^R(\omega) = \sum_{|z_n| < 1} \frac{z_n^{j-j'} \hat{A}(z_n, \omega)}{\prod_{m \neq n} (z_n - z_m)}, \quad (15.6)$$

where $\hat{A}(z, \omega)$ is the cofactor matrix of $[\omega - \hat{\mathcal{H}}(z)]z$. For notational simplicity, we omit the superscript ‘ R ’ in retarded GFs from now on.

Given the real-space components of the bulk GF in Eq. (15.6), we next employ Ref. [26] (see also Ref. [71]) to derive the bGF characterizing a *semi-infinite* nanowire. To that effect, we add an impurity potential ϵ localized at lattice site $j = 0$. Taking the limit $\epsilon \rightarrow \infty$, the infinite chain is cut into disconnected semi-infinite chains with $j < -1$ (left side, L) and $j > 1$ (right side, R). Using the Dyson equation, the local GF components of the cut nanowire follow as [26]

$$\hat{\mathcal{G}}_{jj}(\omega) = \hat{G}_{jj}(\omega) - \hat{G}_{j0}(\omega) [\hat{G}_{00}(\omega)]^{-1} \hat{G}_{0j}(\omega). \quad (15.7)$$

The bGF for the left and right semi-infinite chain, respectively, are with Eq. (15.7) given by

$$\hat{\mathcal{G}}_L(\omega) = \hat{\mathcal{G}}_{-1,-1}(\omega), \quad \hat{\mathcal{G}}_R(\omega) = \hat{\mathcal{G}}_{11}(\omega). \quad (15.8)$$

We note that by proceeding along the lines of Refs. [60, 9], one can also compute reflection matrices from the corresponding bGF,

$$\hat{r}_{L/R} = \lim_{\omega \rightarrow 0} \frac{\mathbb{1} - i\hat{\mathcal{V}}_{\pm 1}^\dagger \hat{\mathcal{G}}_{L/R}(\omega) \hat{\mathcal{V}}_{\pm 1}}{\mathbb{1} + i\hat{\mathcal{V}}_{\pm 1}^\dagger \hat{\mathcal{G}}_{L/R}(\omega) \hat{\mathcal{V}}_{\pm 1}}. \quad (15.9)$$

This relation allows one to express topological invariants of the bulk Hamiltonian [66, 67] in terms of bGFs.

The roots $z_n(\omega)$ play an important role in what follows. In particular, their knowledge allows us to construct both the bulk and the boundary GFs. In simple cases, this can be done analytically, and otherwise this route offers an efficient numerical scheme. The roots can also provide detailed information about the decay of subgap states localized at the boundaries of semi-infinite wires, and they allow one to compute topological invariants of the bulk system. Let us therefore summarize some general properties of these roots:

- (i) Hermiticity of the BdG Hamiltonian implies that every root $z_n(\omega)$ is accompanied by a root $1/z_n^*(\omega)$, where ‘ $*$ ’ denotes complex conjugation.
- (ii) Electron-hole symmetry of the BdG Hamiltonian implies that $z_n(\omega) = z_n^*(-\omega)$. In the presence of an additional symmetry $\hat{\mathcal{H}}(k) = \hat{U}\hat{\mathcal{H}}(-k)\hat{U}^\dagger$ with a unitary matrix $\hat{U}$, for every root $z_n(\omega)$, also $z_n^*(\omega)$ must be a root.
- (iii) As a consequence of (i) and (ii), $\prod_{n=1}^{2N} z_n(\omega) = 1$.

- (iv) Topological phase transitions can occur once a pair of zero-energy roots hits the unit circle, $|z_n(0)| = 1$, which corresponds to the closing and reopening of a gap in the bulk spectrum.
- (v) Equations (15.6) and (15.7) imply that subgap bound states (with energy E) localized near the boundary of a semi-infinite wire decay into the bulk in a manner controlled by $\max(|z_n(E)| < 1)$.

We illustrate the usefulness of these properties in the following sections for different models of proximitized (quasi-)1D nanowires.

15.4 Spinful single-channel hybrid nanowires

As first example, we consider the spinful single-channel model of Refs. [68, 69] for a proximitized semiconductor nanowire. This model has been extensively studied as prototype for 1D wires harboring a TS phase with broken time-reversal invariance. We use the Nambu bispinor $\hat{\Psi}_k^T = (c_{k\uparrow}, c_{k\downarrow}, c_{-k\downarrow}^\dagger, -c_{-k\uparrow}^\dagger)$, i.e., $N = 4$ in Eq. (15.1). Here, $c_{k\sigma}$ is a fermionic annihilation operator for momentum k and spin $\sigma = \uparrow, \downarrow$, and the bulk BdG Hamiltonian in Eq. (15.1) takes the form

$$\mathcal{H}(k) = \epsilon_k \sigma_0 \tau_z + V_x \sigma_x \tau_0 + \alpha_k \sigma_z \tau_z + \Delta \sigma_0 \tau_x, \quad (15.10)$$

where $\sigma_{x,y,z}$ and $\tau_{x,y,z}$ are Pauli matrices in spin and Nambu (electron-hole) space, respectively, with the identity matrices σ_0 and τ_0 . Regularizing the continuum model of Refs. [68, 69] by imposing a finite lattice spacing a , the kinetic energy $\epsilon_k = 2t[1 - \cos(ka)] - \mu$ includes the chemical potential μ and the nearest-neighbor hopping amplitude t . Furthermore, V_x encapsulates a magnetic Zeeman field oriented along the wire axis, $\alpha_k = \alpha \sin(ka)$ describes the spin-orbit interaction, and Δ refers to the proximity-induced on-site pairing amplitude. The bulk dispersion relation, $E = E_{k,\pm} \geq 0$, then follows from [68, 69]

$$E_{k,\pm}^2 = \Delta^2 + \alpha_k^2 + V_x^2 + \epsilon_k^2 \pm 2\sqrt{\Delta^2 V_x^2 + (\alpha_k^2 + V_x^2) \epsilon_k^2}. \quad (15.11)$$

This model exhibits a topological transition at $V_x = V_c = \sqrt{\Delta^2 + \mu^2}$, where the TS phase is realized for $V_x > V_c$.

Although it is not essential for the subsequent discussion, the parameters t and α can be assigned values appropriate for InAs nanowires [26]. To that end, we put $t = \hbar^2/(2m^*a^2)$, where m^* is the effective mass, and $\alpha = \hbar u/a$, where u is the spin-orbit parameter [69]. This parameter depends on material properties and can be tuned by an external electric field. Putting $a = 10$ nm and using typical InAs material parameters, we estimate $t \approx 10$ meV and $\alpha \approx 4$ meV [26]. On the other hand, a proximity gap of order $\Delta \approx 0.2$ meV represents the case of a nanowire in good contact with a superconducting Al layer. (We will use this value in the figures below unless noted otherwise.) The only remaining free variables are then given by V_x and μ .

Using Eq. (15.5) and $z = e^{ika}$, the roots $z_n(\omega)$ for this model satisfy the condition

$$2\Delta^2 [\tilde{\alpha}^2(z) + \epsilon^2(z) - V_x^2 - \omega^2] + 2\tilde{\alpha}^2(z) [V_x^2 - \omega^2 - \epsilon^2(z)] + \tilde{\alpha}^4(z) + V_x^4 + \Delta^4 + [\omega^2 - \epsilon^2(z)]^2 - 2V_x^2 [\omega^2 + \epsilon^2(z)] = 0, \quad (15.12)$$

with the functions

$$\tilde{\alpha}(z) = -i\alpha(z - z^{-1})/2, \quad \epsilon(z) = -t(z + z^{-1} - 2) - \mu. \quad (15.13)$$

Equation (15.12) can be written as

$$\sum_{n=1}^4 C_n(\omega) \left(z^n + \frac{1}{z^n} \right) + C_0(\omega) = 0, \quad (15.14)$$

where the real coefficients $C_n(\omega)$ are given in Appendix 15.8. Clearly, Eq. (15.14) is consistent with the general properties (i) and (ii) listed in Sec. 15.3. Alternatively, Eq. (15.14) can be expressed as an eighth-order polynomial equation:

$$\sum_{m=0}^8 a_m(\omega) z^m = 0, \quad (15.15)$$

where the coefficients a_m are trivially related to the C_n and we can impose the normalization conditions $a_0 = a_8 = 1$.

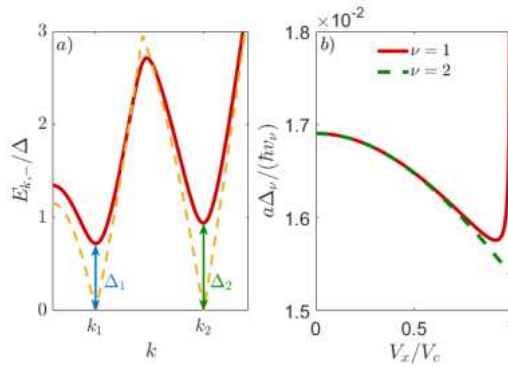


Figure 70: Bulk dispersion relation of the spinful single-channel Majorana wire model [68, 69]. (a) $E_{k,-}$ vs k [see Eq. (15.11)] for the topologically trivial regime $V_x < V_c$ (solid red curve), indicating the two pairing gaps Δ_1 and Δ_2 at $k = k_1$ and $k = k_2$, respectively [see Eq. (15.16)]. We use $\mu = 5$ meV, $\Delta = 2$ meV, and $V_x = 0.5V_c$. All other parameters are specified in the main text. The dashed yellow curve is for $\Delta = 0$. (b) Evolution of the two gaps (normalized to the velocities $v_{1,2}$) vs Zeeman parameter V_x for $\Delta = 0.2$ meV and $\mu = 2$ meV. We note that Δ_1 and v_1 simultaneously vanish as $V_x \rightarrow V_c$.

The resulting roots z_n can be grouped into two different classes associated with the two pairing gaps Δ_1 and Δ_2 in the bulk spectrum [68, 69] [see Fig. 95(a)]. In the limit $\Delta \rightarrow 0$, these gaps Δ_1 and Δ_2 will also vanish. For $\Delta = 0$, we find from Eq. (15.12) that the zero-frequency roots $z_n(\omega = 0)$ simplify to $e^{\pm ik_1 a}$ and $e^{\pm ik_2 a}$, with

$$k_{1,2} \simeq \cos^{-1} \left(\frac{2t(2t - \mu)}{\alpha^2 + 4t^2} \pm \frac{\sqrt{V_x^2(\alpha^2 + 4t^2) + \alpha^4 + 4t\mu\alpha^2 - \alpha^2\mu^2}}{\alpha^2 + 4t^2} \right). \quad (15.16)$$

At these momenta, the dispersion relation becomes gapless for $\Delta = 0$ [see Fig. 95(a)]. We observe from Eq. (15.16) that k_1 (corresponding to the + sign) becomes purely imaginary for $V_x > V_c$. We will then first discuss the topologically trivial regime $V_x < V_c$.

Figure 95(a) shows that the low-energy physics will be dominated by the regions with $|k| \approx k_1$ and $|k| \approx k_2$. The pairing gaps $\Delta_{1,2} = |E_{k_{1,2},-}|$ then follow by substituting $k_{1,2}$ into the bulk dispersion relation (15.11). In particular, we find that Δ_1 closes and reopens when ramping V_x through the topological transition at $V_x = V_c$. An approximate expression for the roots is obtained by linearizing the $\Delta = 0$ dispersion relation in Eq. (15.11) for electrons and holes near $k = k_1$ and $k = k_2$. Defining the respective velocities as $v_{\nu=1,2} = |\partial_k E_{k=k_{\nu},-}|_{\Delta=0}$,

the effective low-energy Nambu Hamiltonian valid near the respective momentum k_ν can be written as

$$\mathcal{H}_{\text{eff},\nu=1,2}(k) \simeq \begin{pmatrix} v_\nu(k - k_\nu) & \Delta_\nu \\ \Delta_\nu & -v_\nu(k - k_\nu) \end{pmatrix}, \quad (15.17)$$

and similarly for $k \approx -k_\nu$. Using $ika = \ln z$, the condition $\det[\omega - \mathcal{H}_{\text{eff},\nu}(z)] = 0$ can readily be solved. In effect, the roots are given by

$$z_\nu(\omega) \simeq \left(1 \pm \frac{a}{v_\nu} \sqrt{\Delta_\nu^2 - \omega^2}\right) e^{ik_\nu a}, \quad (15.18)$$

plus the complex conjugate values. Inspired by Eq. (15.18), we propose the following ansatz for the roots $z_n(\omega)$ located *inside* the unit circle:

$$z_\nu(\omega) = \left(1 - \tau_\nu \sqrt{\Delta_\nu^2 - \omega^2}\right) e^{i\delta_\nu}, \quad (15.19)$$

where $\tau_{1,2}$ and $\delta_{1,2}$ are phenomenological coefficients. In addition, the complex conjugate root $z_\nu^*(\omega)$ is a solution. This ansatz is expected to work well in the topologically trivial regime $V_x < V_c$. For small Δ and $|\omega|$, Eq. (15.18) implies the limiting behavior $\tau_\nu = a/v_\nu$ and $\delta_\nu = k_\nu a$. In addition, we also impose the condition

$$\tau_1 \Delta_1 = \tau_2 \Delta_2 = \eta \ll 1, \quad (15.20)$$

where η is a small parameter. In the small- Δ case with $\tau_\nu \approx a/v_\nu$, Eq. (15.20) implies that the effective pairing gap Δ_ν is inversely proportional to the corresponding density of states $\propto 1/v_\nu$. Figure 95(b) shows that this condition is accurately fulfilled as long as V_x stays well below V_c . However, Eq. (15.20) becomes less precise for $V_x \rightarrow V_c$. In Appendix 15.8, we provide more refined analytical expressions that determine the parameters η and δ_ν in our ansatz for the roots [see Eqs. (15.19) and (15.20)].

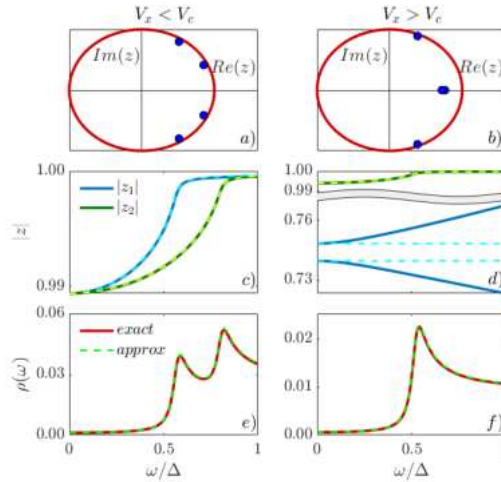


Figure 71: Behavior of the roots $z_n(\omega)$ for the spinful single-channel Majorana wire model [68, 69]. We use $V_x = 0.5V_c$ and $V_x = 1.2V_c$ as representatives for topologically trivial and nontrivial cases, respectively, with $\mu = 5$ meV and other parameters as specified in the main text. Upper panels: Roots $z_n(\omega = 0)$ (black dots) inside the unit circle (red) for (a) $V_x < V_c$ and (b) $V_x > V_c$. For illustrative purposes, we use $\Delta = 1$ meV in panels (a) and (b). For additional information, see Supplemental Material [73]. Middle panels: Modulus of the roots inside the unit circle vs ω/Δ for (c) $V_x < V_c$ and (d) $V_x > V_c$. Solid curves represent numerically exact results and dashed curves follow from Eqs. (15.19) and (15.21), respectively. Bottom panels: Energy dependence of the local bulk DoS, $\rho(\omega)$ (in meV^{-1}), for (e) $V_x < V_c$ and (f) $V_x > V_c$. The solid red curves depict numerically exact results using Eq. (15.3) and the dashed green curves show approximate results obtained from Eq. (15.25).

Next we turn to the topologically nontrivial regime $V_x > V_c$, where the momentum k_1 in Eq. (15.16) becomes purely imaginary. We should then replace $\delta_1 \rightarrow i\delta_1$ in the above ansatz for the roots. As a consequence, the $z_{\nu=1}(\omega)$ roots become real-valued, and the ansatz for $V_x > V_c$ takes the form

$$\begin{aligned} z_{1,\pm}(\omega) &= \left(1 \pm \tau_1 \sqrt{\Delta_1^2 - \omega^2}\right) e^{-\delta_1}, \\ z_{2,\pm}(\omega) &= \left(1 - \tau_2 \sqrt{\Delta_2^2 - \omega^2}\right) e^{\pm i\delta_2}, \end{aligned} \quad (15.21)$$

where both δ_1 and δ_2 are real positive. We thus have only a single pair of complex conjugate roots (z_2) near the unit circle for $V_x > V_c$. Accurate analytical results for the δ_ν and τ_ν parameters can be obtained by solving a cubic equation (see Appendix 15.8). As illustrated in Fig. 96(c) and 96(d), Eq. (15.21) captures the low-energy behavior of the roots rather well, especially in cases where electron-hole symmetry is approximately realized.

For this model of symmetry class D, the $\mathbb{Z}_2$ bulk topological invariant takes the form [1]

$$Q = \frac{\text{sgn Pf } \hat{\mathcal{H}}(k=0)}{\text{sgn Pf } \hat{\mathcal{H}}(k=\pi/a)} = \pm 1. \quad (15.22)$$

Interestingly, the number N_p of complex conjugate root pairs near (but inside) the unit circle is in correspondence with the topological invariant, $Q = (-1)^{N_p}$. These roots can be unambiguously identified as the ones approaching the unit circle from inside in the limit $\Delta \rightarrow 0$, corresponding to the Fermi points in the normal phase. For an odd (even) number of pairs, the phase is thus topologically nontrivial with $Q = -1$ (trivial with $Q = 1$). The upper panels in Fig. 96 illustrate the distribution of the roots inside the unit circle for the cases $V_x < V_c$ and $V_x > V_c$. We observe that upon entering the topologically nontrivial regime, the complex conjugate z_1 roots coalesce to form an almost degenerate root pair $z_{1,\pm}$ [see Eq. (15.21)] located on the real axis inside the unit circle. The roots on the real axis correspond to additional bands at high energies above Δ . At the same time, a single pair of complex conjugate roots (z_2) remains near (but inside) the unit circle, as one expects for a topologically nontrivial phase. As remarked above, this change in the structure of the roots across the transition is consistent with the corresponding change in the topological invariant. The transition between both regions happens when the Pfaffian, or equivalently the Hamiltonian determinant, at $k=0$ vanishes. Using the relation $\det \hat{\mathcal{H}}(k=0) = \prod_{n=1}^8 (1 - z_n(0))$, we thus reproduce property (iv) in Sec. 15.3 which signals the phase transition. It is also worth mentioning that the bulk invariant (15.22) can be directly expressed in terms of bGFs for the semi-infinite wire: Using $Q = \det \hat{r}_L = \det \hat{r}_R$ (see Ref. [9]) the reflection matrices $\hat{r}_{L/R}$ and therefore also Q can be obtained from the bGFs [see Eq. (15.9)].

The knowledge of the roots also gives access to other electronic properties of interest. For instance, we can obtain a compact expression for the energy-dependent local DoS at, say, lattice site $j=0$ of the translation-invariant chain:

$$\rho(\omega) = -\frac{1}{\pi} \text{Im Tr} [\hat{G}_{00}(\omega)]. \quad (15.23)$$

We focus on the low-energy limit, where one can expand the cofactor matrix $\hat{A}(z, \omega)$ in Eq. (15.6) to linear order in ω . The local GF then follows as

$$\hat{G}_{00}(\omega) \simeq \sum_{|z_n| < 1} \frac{\hat{A}(z_n, \omega) + \omega \hat{A}'(z_n, \omega)}{\prod_{m \neq n} (z_n - z_m)}, \quad (15.24)$$

where $\hat{A}'(z_n, \omega) = \frac{d}{d\omega} \hat{A}(z_n(\omega), \omega)$. From our ansatz in Eqs. (15.19) and (15.21), the sum in Eq. (15.24) can be reconstructed. A simple approximate expression follows for small Δ in

the low-energy limit, where one needs to keep just the first-order terms $\propto \tau_\nu \sqrt{\Delta_\nu^2 - \omega^2}$ in the denominator. We then obtain

$$\hat{G}_{00}(\omega) \approx \sum_{\nu=1,2} \frac{\hat{A}_\nu + \omega \hat{A}'_\nu}{\sqrt{\Delta_\nu^2 - \omega^2}}, \quad (15.25)$$

where $\hat{A}_\nu$ and $\hat{A}'_\nu$ are specified in Appendix 15.8. We note that for $V_x > V_c$, the main contribution to Eq. (15.25) stems from the residues associated to z_2 . The results for $\rho(\omega)$ depicted in Figs. 96(e) and 96(f) demonstrate that Eq. (15.25) accurately reproduces numerically exact calculations, both below and above the topological transition.

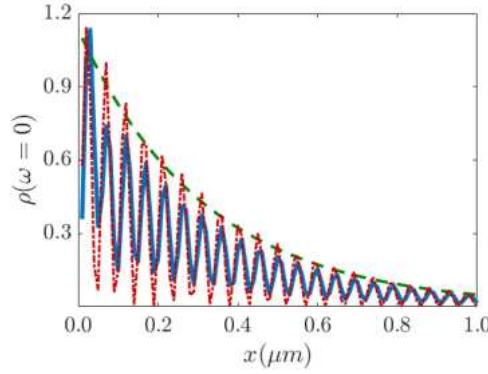


Figure 72: Spatial variation of the local DoS, $\rho_j(0)$ (in meV^{-1}), vs distance from the boundary, $x = ja$ (in μm), for the $\omega = 0$ Majorana state in a semi-infinite TS wire with $\mu = 1$ meV and $V_x = 2V_c$. The solid blue curve gives numerically exact results obtained from Eq. (15.7). Red-dotted and green-dashed curves show Eq. (15.26) with and without $2k_F$ oscillations, respectively.

Next we turn to the case of a semi-infinite chain in the topological phase, $V_x > V_c$. Using the Dyson equation in Eq. (15.7) and taking into account the behavior of the roots of the infinite chain discussed above, we can deduce the spatial decay profile of the zero-energy Majorana end state into the bulk. Noting that the GF components $\hat{G}_{j,0}$ and $\hat{G}_{0,j}$ in Eq. (15.7) are $\propto |z|^j$, we observe that for $V_x > V_c$, the decay is dominated by the z_2 roots since $|z_2| > |z_1|$. Moreover, the decay profile exhibits fast oscillations due to the complex phase δ_2 in Eq. (15.21), which for $\mu \gg \Delta$ can be approximated as $\delta_2 \simeq k_F a$ with $k_F \equiv k_2$. In this approximation, the local DoS of the $\omega = 0$ MBS thus has the spatial profile

$$\rho_j(\omega = 0) \propto |z_2(0)|^{2j} \cos^2(jk_F a + \chi_0), \quad (15.26)$$

where χ_0 describes a phase shift in the $2k_F$ oscillations. Equation (15.26) reproduces the numerically exact results obtained from Eq. (15.7) rather well, as illustrated in Fig. 97. The dashed curve shows that the envelope function is accurately described by $|z_2(0)|^{2j}$, corresponding to an exponential decay into the bulk of the chain.

15.5 Two-channel class-D nanowire

We next examine the case of spinful multichannel hybrid nanowires with broken time reversal symmetry. The bGF approach could in principle be applied to nanowire models with an arbitrary number of channels. In practice, however, the techniques in Sec. 15.3 are less efficient once the degree $2N$ of the secular polynomial (15.5) becomes very large. We here restrict ourselves to the two-channel case with $N = 8$, which can be realized for two single-channel nanowires coupled by tunneling terms. The resulting model already exhibits many of the features expected for generic multichannel nanowires [74, 75].

Our model Hamiltonian is given by

$$\hat{\mathcal{H}}_{2\text{ch}}(k) = \begin{pmatrix} \hat{\mathcal{H}}(k) & \hat{T} \\ \hat{T}^\dagger & \hat{\mathcal{H}}(k) \end{pmatrix}, \quad (15.27)$$

where the 2×2 structure refers to wire space. We consider two identical spinful single-channel Majorana wires described by the model of Refs. [68, 69] with $\hat{\mathcal{H}}(k)$ in Eq. (15.10). The interwire tunnel couplings are modeled by

$$\hat{T} = -t_y \sigma_0 \tau_z + i\alpha_y \sigma_x \tau_z + \Delta_y \sigma_0 \tau_x, \quad (15.28)$$

where t_y and α_y are spin-conserving and spin-flipping hopping amplitudes, respectively. The coupling α_y may arise due the presence of a Rashba spin-orbit coupling produced by an electric field along the z -direction. As in Sec. 15.4, we write $t_y = \hbar^2/(2m^*a_y^2)$ and $\alpha_y = \hbar u/a_y$, with the minimal distance a_y between the wires. In the concrete examples shown below, we assume $a_y = 3a$, which corresponds to a subband separation of ≈ 3 meV. The interwire coupling (15.28) also includes a non-local inter-wire pairing amplitude Δ_y . For the present class D case, however, we find that allowing for a small $\Delta_y \neq 0$ does not lead to significant changes in the phase diagram. We thus put $\Delta_y = 0$ in this section.

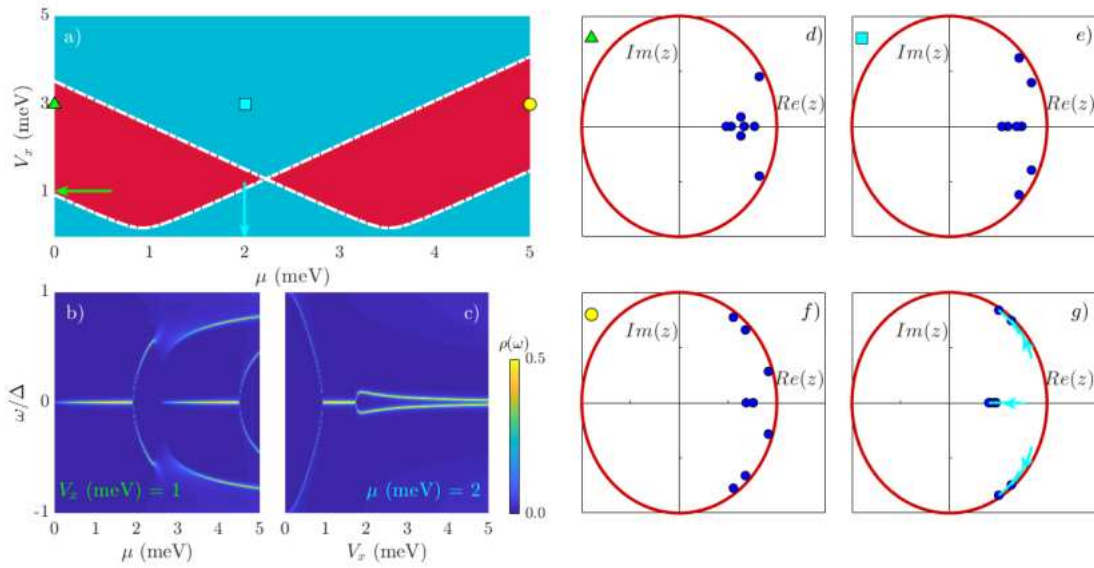


Figure 73: Two-channel spinful Majorana wire model of class D [see Eq. (15.27)] with parameters as explained in the main text. Panel (a) shows the bulk phase diagram in the μ - V_x plane. Topological nontrivial (trivial) phases are shown in red (blue). Panels (b) and (c) show the energy dependence of the local DoS, $\rho_{j=1}(\omega)$ (in meV^{-1}), at the boundary of a semi-infinite two-channel wire along the trajectories marked by arrows in panel (a). Panels (d)–(f) illustrate the roots $z_n(0)$ inside the unit circle at the three points indicated in panel (a) by a triangle (d), a square (e), and a circle (f), respectively. For additional insights, see Supplemental Material [73]. Panel (g) shows the evolution of the roots within the topologically trivial regime as V_x increases from 3 to 8 meV at constant chemical potential $\mu = 2$ meV. In panels (d)–(g), we use $\Delta = 1$ meV.

One can characterize the phase diagram of a translationally invariant two-channel wire by using the bulk topological invariant in Eq. (15.22) with the replacement $\hat{\mathcal{H}}(k) \rightarrow \hat{\mathcal{H}}_{2\text{ch}}(k)$. The

Pfaffian at $k = 0$ is here given by

$$\begin{aligned} \text{Pf} \hat{\mathcal{H}}_{2\text{ch}}(0) = & \alpha_y^4 + [(\mu - 3t_y)^2 - V_x^2 + \Delta^2] [(\mu - t_y)^2 - V_x^2 + \Delta^2] + \\ & + 2\alpha_y^2 [-(\mu - 3t_y)(\mu - t_y) - V_x^2 + \Delta^2]. \end{aligned} \quad (15.29)$$

The boundaries of the topological phase correspond to a vanishing Pfaffian at $k = 0$, where Eq. (15.29) implies the two critical Zeeman fields

$$V_{c,\pm} = \left(\alpha_y^2 + \mu^2 - 4\mu t_y + 5t_y^2 + \Delta^2 \pm 2|\mu - 2t_y|\sqrt{t_y^2 + \alpha_y^2} \right)^{1/2}.$$

The resulting phase diagram in the μ - V_x plane is illustrated in Fig. 98(a). We observe that the two-channel model (15.27) exhibits a richer phase diagram than in the single-channel case (see also Refs. [74, 75]).

We next construct the bGF of a semi-infinite wire by determining the roots of the secular polynomial in Eq. (15.5), which here is a 16th-order polynomial equation that we solve numerically. Figure 98(b) and 98(c) illustrates the evolution of the energy-dependent local DoS, $\rho_1(\omega)$, at the boundary, i.e., taken at site $j = 1$ of a semi-infinite two-channel wire. We consider two different trajectories in the μ - V_x plane as indicated by the arrows in Fig. 98(a). For constant V_x [panel b)], there are both topologically nontrivial and trivial regions as μ is varied. In the topologically nontrivial regions, we observe a zero-energy peak in the local DoS, signaling the presence of MBSs. This $\omega = 0$ peak is absent in the trivial regime. For fixed μ [panel (c)], the topologically nontrivial phase is reached for intermediate values of V_x . For larger V_x , even though the system is in a trivial phase, we find low-energy Andreev bound states that approach zero energy as V_x increases. This effect has also been described in Ref. [75].

Additional insights follow by analyzing the evolution of the roots $z_n(\omega = 0)$ inside the unit circle in the complex momentum plane. In Fig. 98(d)–(f), we illustrate their distribution for three different points in the phase diagram. For panels (d) and (f), the system is in a topological phase and, as expected, one finds an *odd* number of pairs of complex conjugate roots close to the unit circle. As in Sec. 15.4, the roots on the real axis correspond to additional bands at higher energies well above Δ . Panel (e) instead corresponds to a topologically trivial phase with an *even* number of conjugate root pairs near the unit circle. Finally, Fig. 98(g) illustrates the evolution of the roots in the topologically trivial regime as the Zeeman parameter V_x increases. We find that the both roots near the unit circle in the first quadrant become almost degenerate for large V_x . Such a behavior effectively amounts to having two replicas of a single-channel TS wire, which in turn helps to explain why Andreev bound states approach the zero energy limit for strong Zeeman field [see Fig. 98(c) and Ref. [75]].

15.6 TRITOPS nanowires

Next we turn to models for hybrid nanowires of symmetry class DIII. Such TRITOPS wires constitute another interesting system with topologically nontrivial phases. Below we first study single-channel wires and subsequently turn to the two-channel case.

15.6.1 Single-channel case

Many different proposals for physical realizations of single-channel TRITOPS wires have been put forward in the recent past [18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35]. For concreteness, we will here focus on the model introduced by Zhang *et al.* [21]. Using the spin-Nambu basis with $N = 4$ in Sec. 15.3, the Hamiltonian is given by

$$\hat{\mathcal{H}}_{\text{DIII}}(k) = \epsilon_k \sigma_0 \tau_z + \alpha_k \sigma_z \tau_z + \Delta_k \sigma_0 \tau_x, \quad (15.30)$$

where in this section, we use

$$\begin{aligned}\epsilon_k &= -2t \cos(ka) - \mu, \\ \alpha_k &= 2\alpha \sin(ka),\end{aligned}\tag{15.31}$$

$$\Delta_k = 2\Delta \cos(ka).\tag{15.32}$$

Again t corresponds to a nearest-neighbor hopping amplitude, μ is the chemical potential, a the lattice spacing, and α the spin-orbit coupling strength. The parameter Δ corresponds to a nearest-neighbor pairing interaction. In the examples below, we use $a = 10$ nm, $t = 10$ meV and $\alpha = 4$ meV as in Secs. 15.4 and 15.5.

By a simple rearrangement of the spin-Nambu spinor $\hat{\Psi}_k$, one can block-diagonalize the Hamiltonian in Eq. (15.30), $\hat{\mathcal{H}}_{\text{DIII}} = \text{diag}(\hat{\mathcal{H}}_-, \hat{\mathcal{H}}_+)$. To that end, upon replacing $\hat{\Psi}_k^T \rightarrow (c_{k\uparrow}, c_{-k\downarrow}^\dagger, c_{k\downarrow}, -c_{-k\uparrow}^\dagger)$, we arrive at the 2×2 block Hamiltonians

$$\hat{\mathcal{H}}_\pm(k) = (\epsilon_k \mp \alpha_k) \tilde{\sigma}_z + \Delta_k \tilde{\sigma}_x = \beta_\pm(k) \cdot \tilde{\sigma},\tag{15.33}$$

where $\tilde{\sigma}$ is the vector of $\tilde{\sigma}_{x,y,z}$ Pauli matrices in the respective 2×2 space obtained after block diagonalization. Each Hamiltonian $\hat{\mathcal{H}}_\pm(k)$ corresponds to a Dirac-type model where

$$\beta_\pm(k) = \begin{pmatrix} 2\Delta \cos(ka) \\ 0 \\ -\mu - 2t \cos(ka) \pm 2\alpha \sin(ka) \end{pmatrix}\tag{15.34}$$

is a vector field mapping the first Brillouin zone onto a closed curve.

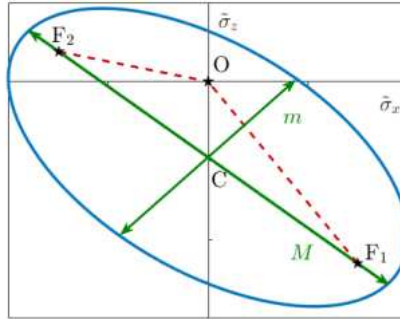


Figure 74: Curve traced out by $\beta_-(k)$ in the $\tilde{\sigma}_x$ - $\tilde{\sigma}_z$ plane for a single-channel TRITOPS wire in a topologically nontrivial phase [see Eqs. (15.33) and (15.34)] with $t = 0.5$, $\alpha = 0.8$, $\Delta = 1$, and $\mu = 1.04$ (all in meV). The evolution of the bulk Hamiltonian $\hat{\mathcal{H}}_-(k)$ upon traversal of the Brillouin zone is described by an ellipse containing the origin (O). For details, see main text and Appendix 15.9.

At this stage, we can apply the formalism of Ref. [76] for analyzing the roots of the secular polynomial of Dirac-like Hamiltonians. By projecting $\hat{\mathcal{H}}_\pm$ to the $\tilde{\sigma}_x$ - $\tilde{\sigma}_z$ plane, we obtain an elliptic curve as illustrated in Fig. 99. According to the arguments in Ref. [76], if the ellipse encloses the origin of the $\tilde{\sigma}_x$ - $\tilde{\sigma}_z$ plane, we know that for a semi-infinite wire, $\hat{\mathcal{H}}_\pm(k)$ will generate an edge state with energy equal to the modulus of the component of $\beta_\pm(k)$ perpendicular to this plane. In our case, $[\beta_\pm(k)]_y = 0$ implies that we have a pair of zero-energy boundary states in the topological phase. In addition, this argument also shows that there are no finite-energy Andreev bound states in the trivial phase (where the ellipse does not contain the origin). For the case in Fig. 99, where the origin is displaced along the $\tilde{\sigma}_z$ -axis, the topological transition

occurs at $ka = \pm\pi/2$ and $|\mu| = 2\alpha$ [see Eq. (15.34)]. This conclusion is consistent with the fact that at the topological transition, one finds roots at $z = e^{ika} = \pm i$ (see also Ref. [77]), in agreement with property (v) in Sec. 15.3.

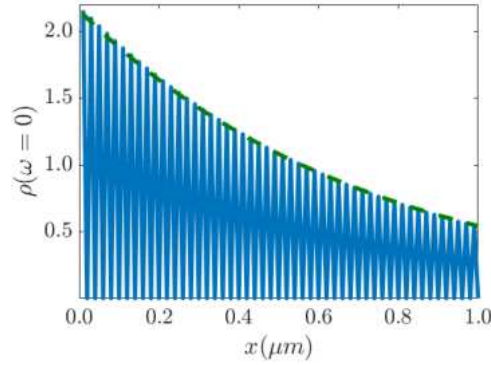


Figure 75: Spatial variation of the local DoS at zero energy (in meV^{-1}), corresponding to Majorana end states of a semi-infinite TRITOPS wire in its topological phase [see Eq. (15.30)] for $\mu = 0$ (blue solid curve). The green dashed curve shows an exponential decay on the length scale $\lambda_e = -\frac{a}{2} \ln |z_{\max}|$ [see Eq. (15.83)].

More generally, by determining the roots $z_n(\omega)$, we can again construct the bGF of a semi-infinite wire. In particular, we thereby obtain the class DIII bulk topological invariant via the reflection matrices in Eq. (15.9). In the present case, the invariant is given by $Q = \text{Pf}(i\hat{r}_{L,R})$ [9]. Furthermore, using the results of Refs. [76, 77], an analytical expression for the largest-modulus zero-frequency root, $z_{\max}$, inside the unit circle can be computed from purely geometrical considerations for the ellipse in Fig. 99 (see Appendix 15.9 for details). The length scale governing the spatial decay profile of the pair of Majorana states localized near the boundary of a semi-infinite TRITOPS wire then follows as $\lambda_e = -\frac{a}{2} \ln |z_{\max}|$ [see Eq. (15.83) in Appendix 15.9]. The validity of this expression is confirmed in Fig. 100, where we show numerically exact results for the spatial variation of the local DoS at $\omega = 0$ together with the prediction obtained from Eq. (15.83).

15.6.2 Two-channel case

As in Sec. 15.5, we can also extend the TRITOPS model to the two-channel case by coupling two single-channel wires. More general multichannel wire constructions are also possible but will not be pursued here. The corresponding Hamiltonian is with Eq. (15.30) given by

$$\hat{\mathcal{H}}_{\text{DIII},2\text{ch}}(k) = \begin{pmatrix} \hat{\mathcal{H}}_{\text{DIII}}(k) & \hat{T}_{\text{DIII}} \\ \hat{T}_{\text{DIII}}^\dagger & \hat{\mathcal{H}}_{\text{DIII}}(k) \end{pmatrix}, \quad (15.35)$$

where the interwire tunneling couplings are modeled in a similar manner as in Eq. (15.28):

$$\hat{T}_{\text{DIII}} = -t_y \sigma_0 \tau_z + i\alpha_y \sigma_y \tau_z + \Delta_y \sigma_0 \tau_z. \quad (15.36)$$

We here allow for spin-conserving (t_y) and spin-flipping (α_y) hopping processes, as well as for nonlocal pairing terms (Δ_y). Below, t_y and α_y are parametrized as specified in Sec. 15.5.

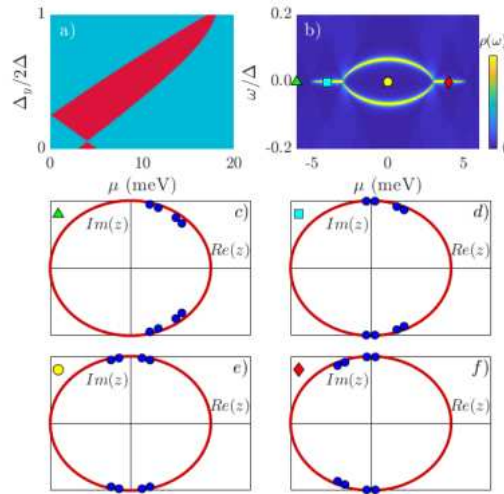


Figure 76: Two-channel TRITOPS nanowire [see Eq. (15.35)], with parameters as explained in the main text. Panel (a) shows the phase diagram in the μ - Δ_y plane, with the topologically nontrivial (trivial) phase in red (blue). (b) Local DoS, $\rho_{j=1}(\omega)$ (in meV^{-1}), at the boundary of a semi-infinite wire in the μ - ω plane for $\Delta_y = 0$. Panels (c) to (f) depict the roots $z_n(\omega = 0)$ inside the unit circle for different μ as indicated by the respective symbol in panel (b). We use $\Delta = 1$ meV in panels (c)–(f).

The resulting phase diagram is illustrated in Fig. 101(a). To make analytical progress, from now on we consider the case $\Delta_y = 0$ and determine the conditions for gap closings, and thus for phase transition curves in the two-channel TRITOPS case. The gap closes again for $ka = \pm\pi/2$ as in Sec. 15.6.1 but now for the chemical potential set to one of the critical values

$$|\mu_{\pm}| = \sqrt{\alpha_y^2 + (t_y \pm 2\alpha)^2}. \quad (15.37)$$

where the topological invariant is related to the product of the signs of the effective pairing amplitude at different Fermi points [66]. As the critical momenta are as in Sec. 15.6.1, the pairing function is directly determined by $\Delta \cos(ka)$ [see Eq. (15.30)]. For this reason, the topologically nontrivial (trivial) phase has an odd (even) number of Fermi points between $ka = 0$ and $ka = \pi/2$.

The bGF can again be computed from the roots of the secular polynomial. The latter also determine the behavior of the edge modes of a semi-infinite two-channel TRITOPS wire in different regions of the phase diagram. By continuity, the condition of having an odd number of Fermi points with $0 < k_F < \pi/2a$ corresponds to an odd number N_p of roots near the unit circle in the first quadrant. Our results for the roots are illustrated in Fig. 101(c)–101(f). As expected, N_p is odd for panels (d) and (f), where panel (b) shows that Majorana end states are present and thus a topological phase is realized. By contrast, panels (c) and (e) show topologically trivial cases with even N_p .

15.7 Phase-biased topological Josephson junctions

In this section, we consider different examples for the equilibrium supercurrent-phase relation in two- and three-terminal Josephson junctions made of nanowires in topologically nontrivial superconducting phases. These wires are coupled together by tunnel junctions. We start in Sec. 15.7.1 with the case of a trijunction of TS nanowires (see also Ref. [61]), and then turn to TRITOPS-TS Josephson junctions in Sec. 15.7.2.

15.7.1 Three-terminal TS junctions

We first consider a three-terminal junction formed by spinful single-channel nanowires in the TS phase. For a schematic layout, see Fig. 102. Such devices have been suggested, e.g., for Majorana braiding implementations [78, 79, 80], for the engineering of artificial topological Weyl semimetal phases [81, 82], and for the observation of giant shot noise features induced by the single zero-energy MBS localized at the trijunction [83]. While most previous studies have been based on minimal models or on spinless Kitaev chain models, a more realistic description using the spinful nanowire model of Refs. [68, 69] discussed in Sec. 15.4 is desirable. In particular, one can then assess the role of the spin degree of freedom and the effects of various microscopic parameters such as the angle θ in Fig. 102. We assume that each wire is sufficiently long such that the overlap between MBSs located at different ends of the same wire is negligibly small.

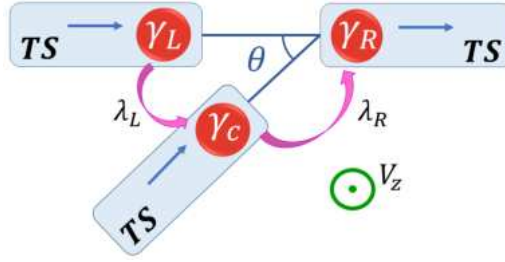


Figure 77: Three-terminal junction of spinful TS nanowires (see Sec. 15.4), with two parallel wires (L, R) and a central (C) wire at angle θ . The red dots indicate MBSs with Majorana operators $\gamma_{L,R,C}$ near the junction, with tunnel couplings $\lambda_{L,R}$ connecting the L, R wires to the C wire. We assume that no direct tunnel coupling between the L and R wires is present. A Zeeman field V_z is applied perpendicular to the plane containing the three wires. Blue arrows show the positive momentum direction in each wire.

We model each nanowire in the setup of Fig. 102 in terms of the spinful single-channel Hamiltonian of Eq. (15.10). All three wires lie in a plane, with two of them aligned (L and R in Fig. 102) and the third (the central wire, C , in Fig. 102) at an arbitrary angle θ to the other two. We here assume that the Zeeman field V_z is oriented perpendicular to the plane (see Ref. [1]). For simplicity, we consider identical material parameters for the three wires which are chosen such that the TS phase is realized.

Let us next discuss the unitary rotations necessary to adapt the bGFs of Sec. 15.4 to a common reference frame for all three wires in Fig. 102. We first perform a $\pi/2$ rotation of the spin axis around the y axis, which connects the intrinsic coordinate system of the L and R wires to the common reference frame. Defining

$$R(\vartheta) = [\sigma_0 \cos(\vartheta/2) - i\sigma_y \sin(\vartheta/2)] \tau_0, \quad (15.38)$$

the corresponding rotation matrix, $R_y = R(\vartheta = \pi/2)$, transforms a Zeeman field along the x -direction (see Sec. 15.4) into a Zeeman field along the negative z -direction (as in Fig. 102). The bGFs for the L and R wires in Fig. 102 are thus given by

$$\hat{\mathcal{G}}'_{L/R} = R_y \hat{\mathcal{G}}_{L,R} R_y^{-1}, \quad (15.39)$$

with $\hat{\mathcal{G}}_{L,R}$ as described in Sec. 15.4. For the C lead, we additionally have to rotate by the angle θ around the global z -axis. The corresponding rotation matrix, $R_z(\theta)$, follows from Eq. (15.38) with the replacements $\sigma_y \rightarrow \sigma_z$ and $\vartheta \rightarrow \theta$. We thereby obtain

$$\hat{\mathcal{G}}_C = R_z(\theta) R_y \hat{\mathcal{G}}_L R_y^{-1} R_z^{-1}(\theta). \quad (15.40)$$

In what follows, we rewrite $\hat{\mathcal{G}}'_{L/R} \rightarrow \hat{\mathcal{G}}_{L/R}$ to keep the notation simple.

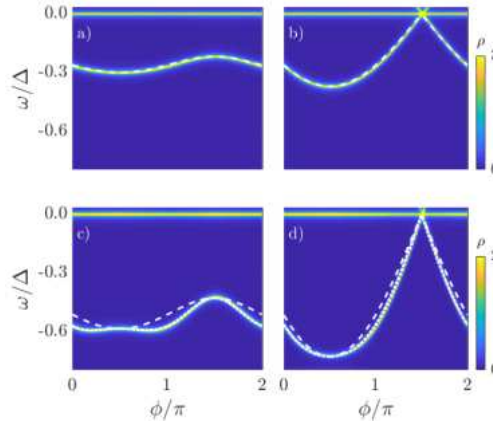


Figure 78: Phase dependence of the subgap spectrum of the trijunction of TS wires in Fig. 102, with the superconducting phases $\phi_L = -\phi_R = \phi$ and $\phi_C = 0$. The TS wires are modeled as spinful nanowires with $\mu = 2$ meV, $V_z = 3V_c$, and symmetric couplings, $\lambda_L = \lambda_R = \lambda$. For other parameters, see Sec. 15.4. Panel (a) [(b)] is for $\lambda = 2$ meV and $\theta = \pi/2$ [$\theta = \pi/10$]. Panel (c) [(d)] is for $\lambda = 5$ meV and $\theta = \pi/2$ [$\theta = \pi/10$]. From blue to yellow, $\rho_{3\text{TS}}(\omega)$ (in meV^{-1}) gradually increases, where Eq. (15.43) has been evaluated in a numerically exact manner. White dotted [dashed] curves show the approximate Andreev bound state dispersion relation in Eq. (15.54) [Eq. (15.51)].

The coupling between the L, R wires and the C wire is modeled by a spin-conserving tunneling term,

$$H_T = \frac{1}{2} \sum_{\nu=L,R} \hat{\Psi}_\nu^\dagger \hat{\lambda}_\nu \hat{\Psi}_C + \text{H.c.}, \quad \hat{\lambda}_\nu = \lambda_\nu \sigma_0 \tau_z e^{i\tau_z \phi_\nu/2}, \quad (15.41)$$

where $\hat{\Psi}_{L,R,C}$ are boundary spinor fields and ϕ_ν is the phase of the superconducting order parameter in the respective wire. We choose a gauge with $\phi_C = 0$ and real-valued tunnel couplings λ_ν . The physical properties of the trijunction are then determined by the full bGF,

$$\hat{G}_{3\text{TS}} = \begin{pmatrix} \hat{\mathcal{G}}_L^{-1} & \hat{\lambda}_L & 0 \\ \hat{\lambda}_L^\dagger & \hat{\mathcal{G}}_C^{-1} & \hat{\lambda}_R \\ 0 & \hat{\lambda}_R^\dagger & \hat{\mathcal{G}}_R^{-1} \end{pmatrix}^{-1}, \quad (15.42)$$

where the 3×3 structure refers to wire space. From Eq. (15.42), the energy dependence of the local DoS at the junction will be given by

$$\rho_{3\text{TS}}(\omega) = -\frac{1}{\pi} \text{Im Tr} [\hat{G}_{3\text{TS}}(\omega)]. \quad (15.43)$$

Figure 78 shows the phase dependence of $\rho_{3\text{TS}}(\omega)$ obtained by numerical evaluation of Eqs. (15.42) and (15.43) for a trijunction with $\phi_L = -\phi_R = \phi$ and $\phi_C = 0$. (This is the series configuration in the parlance of Ref. [61].)

Deep in the topological regime, the low-energy properties of the trijunction are well described by a minimal model keeping only the MBSs at the junction. To show this from the above bGFs, we first derive an effective Hamiltonian for each wire that only keeps track of the respective MBS:

$$H_{\text{eff},\nu} = \lim_{\omega \rightarrow 0} \hat{\mathcal{G}}_\nu^{-1}(\omega). \quad (15.44)$$

Using Eq. (15.44) and recalling that the z_2 roots dominate for $V_x > V_c$, we can read off the boundary spinors for each of the wires ($\nu = L, R, C$; see Ref. [61]):

$$\hat{\Psi}_L \simeq \sqrt{\frac{\Delta_2}{t}} \begin{pmatrix} 0 \\ 1 \\ -i \\ 0 \end{pmatrix} \gamma_L, \quad \hat{\Psi}_R \simeq \sqrt{\frac{\Delta_2}{t}} \begin{pmatrix} 0 \\ -i \\ 1 \\ 0 \end{pmatrix} \gamma_R, \quad \hat{\Psi}_C \simeq \sqrt{\frac{\Delta_2}{t}} R_z(\theta) \begin{pmatrix} 0 \\ 1 \\ -i \\ 0 \end{pmatrix} \quad (15.45)$$

where the Majorana operators γ_ν satisfy the anticommutation relations $\{\gamma_\nu, \gamma_{\nu'}\} = \delta_{\nu\nu'}$. The pairing gap Δ_2 has been defined in Sec. 15.4 [see also Fig. 95 and Eq. (15.20)].

Next, we project the tunneling Hamiltonian (15.41) to the Majorana sector by means of Eq. (15.45). We thereby arrive at a minimal model Hamiltonian,

$$H_{\text{mm}} = -i\Omega_L(\phi)\gamma_L\gamma_C - i\Omega_R(\phi)\gamma_R\gamma_C, \quad (15.46)$$

with the energies

$$\begin{aligned} \Omega_L(\phi) &= \frac{2\Delta_2\lambda_L}{t} \sin\left(\frac{\phi + \theta}{2}\right), \\ \Omega_R(\phi) &= -\frac{2\Delta_2\lambda_R}{t} \cos\left(\frac{\phi - \theta}{2}\right). \end{aligned} \quad (15.47)$$

Equation (15.46) is easily diagonalized by rotating the $\gamma_{L,R}$ operators to new Majorana operators $\tilde{\gamma}_{L,R}$,

$$\begin{pmatrix} \gamma_L \\ \gamma_R \end{pmatrix} = \begin{pmatrix} \sin \kappa & -\cos \kappa \\ \cos \kappa & \sin \kappa \end{pmatrix} \begin{pmatrix} \tilde{\gamma}_L \\ \tilde{\gamma}_R \end{pmatrix}, \quad (15.48)$$

with $\sin \kappa = \Omega_L/\Omega$ and

$$\Omega(\phi) = \sqrt{\Omega_L^2(\phi) + \Omega_R^2(\phi)}. \quad (15.49)$$

We thereby arrive at

$$H_{\text{mm}} = -i\Omega(\phi)\tilde{\gamma}_L\gamma_C, \quad (15.50)$$

where the decoupled Majorana operator $\tilde{\gamma}_R$ describes the remaining zero-energy state [83]. The eigenstates of Eq. (15.50) correspond to Andreev bound states with the phase-dependent subgap energy [see Eq. (15.47)],

$$E_{\pm}(\phi) = \pm \frac{1}{2} \sqrt{\Omega_L^2(\phi) + \Omega_R^2(\phi)}. \quad (15.51)$$

The phase derivative $\partial_\phi E_-(\phi)$ then yields the Josephson current-phase relation. As illustrated in Fig. 78, Eq. (15.51) reproduces our numerically exact bGF calculations for small tunnel couplings $\lambda_{L,R}$.

However, for intermediate-to-large values of the tunnel couplings, the Andreev bound state dispersion may deviate from Eq. (15.51) [see, e.g., the ‘bump’-like features in Fig. 78(c)]. Such deviations are due to the fact that the Majorana operators γ_L and γ_R will become connected through the virtual excitation of continuum quasiparticle states with above-gap energy $E > \Delta$. Within our minimal model, this physics can be taken into account by adding an effective coupling λ_{LR} between the L and R wires. For $\lambda_\nu \ll \Delta$, we estimate $\lambda_{LR} \simeq \lambda_L\lambda_R/\Delta$. The corresponding tunneling term is given by

$$H_{T,LR} = \frac{1}{2} \lambda_{LR} \hat{\Psi}_L^\dagger \sigma_0 e^{i\tau_z \phi} \tau_z \hat{\Psi}_R + \text{H.c.} \quad (15.52)$$

Using the Majorana spinors in Eq. (15.45) together with Eq. (15.46), we arrive at an improved version of the minimal model Hamiltonian:

$$\begin{aligned} H_{\text{mm}} &= -i\Omega_L(\phi)\gamma_L\gamma_C - i\Omega_R(\phi)\gamma_R\gamma_C \\ &\quad - i\Omega_{LR}(\phi)\gamma_L\gamma_R, \\ \Omega_{LR}(\phi) &= \frac{2\Delta_2\lambda_{LR}}{t}\cos\phi. \end{aligned} \quad (15.53)$$

One can easily show that Eq. (15.53) still predicts a decoupled zero-energy MBS at the tri-junction. The hybridization between the remaining two Majorana states yields Andreev bound states with the dispersion relation

$$E_{\pm}(\phi) = \pm \frac{1}{2} \sqrt{\Omega_L^2(\phi) + \Omega_R^2(\phi) + \Omega_{LR}^2(\phi)}. \quad (15.54)$$

Of course, for $\lambda_{LR} \rightarrow 0$, we recover Eq. (15.51). Only by including the Ω_{LR} term in Eq. (15.54), however, the bumps found in the numerically exact dispersion in Fig. 78(c) can be accurately reproduced.

We conclude that the minimal model in Eq. (15.53), which has been derived from the bGF approach, captures the basic physics of the Josephson effect in the three-terminal TS junction shown in Fig. 102. In particular, the dependence of the current-phase relation on the angle θ between the wires resulting from the subgap spectrum in Fig. 102 will be correctly reproduced.

15.7.2 TRITOPS-TS junction

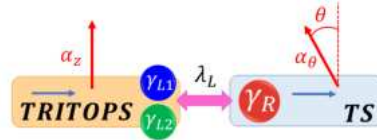


Figure 79: Sketch of a TRITOPS-TS Josephson junction. Colored dots indicate MBSs corresponding to the Majorana operators $\gamma_{L1,L2,R}$. The tunnel coupling λ_L connects both wires, where blue arrows shows the positive momentum direction in each wire. The spin-orbit axes on both sides are tilted by the relative angle θ .

We next consider the two-terminal Josephson junction in Fig. 79 between a TRITOPS wire [see Eq. (15.33) in Sec. 15.6.1], and a TS nanowire [see Eq. (15.10) in Sec. 15.4]. Denoting the respective boundary spin-Nambu spinors by $\hat{\Psi}_L$ and $\hat{\Psi}_R$, respectively, the tunneling Hamiltonian is given by

$$H_T = \frac{1}{2} \lambda_L \hat{\Psi}_L^\dagger \sigma_0 e^{i\tau_z\phi/2} \tau_z \hat{\Psi}_R + \text{H.c.}, \quad (15.55)$$

where ϕ is the superconducting phase difference across the junction and we assume a real-valued tunnel coupling λ_L . Below we assume for simplicity that the pairing gap Δ is identical for both nanowires. We will allow for a relative angle θ between the directions of the spin-orbit field in each wire, see the schematic device layout in Fig. 79. One could vary θ by changing the orientation of a local electric field applied to the TS wire only, which in turn will affect the corresponding Rashba spin-orbit field. In addition, we need a Zeeman field to induce the topological phase in the TS nanowire (see Sec. 15.4), while no Zeeman field should be present on the time-reversal invariant TRITOPS side. To achieve this goal, one may use mesoscopic ferromagnets for inducing a Zeeman field only locally [84].

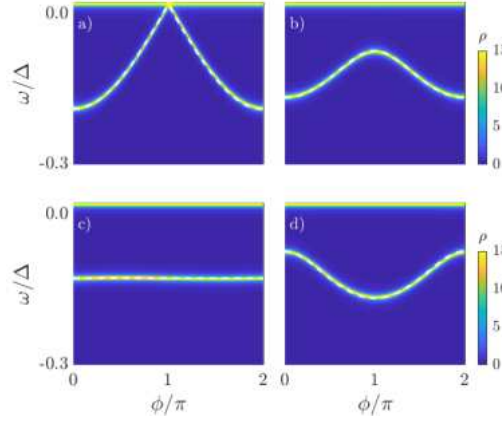


Figure 80: Phase-dependent subgap spectrum of a TRITOPS-TS Josephson junction for different values of the tilt angle θ in Fig. 79. The spinful single-channel model parameters are as described in Secs. 15.4 and 15.6, with $\mu = 1$ meV, $\lambda_L = 2$ meV, and $V_x = 1.5V_c$ on the TS side. The tilt angle is $\theta = 0$ in panel (a), $\theta = 0.3\pi$ in panel (b), $\theta = \pi/2$ in panel (c), and $\theta = 0.7\pi$ in panel (d). From blue to yellow, the color code indicates increasing DoS values at the junction, $\rho(\omega)$ (in meV^{-1}). White dashed curves show the Andreev bound states (15.59).

To account for the angle θ , we then apply the unitary transformation $R_y(\theta)$ to the bGF describing the TS nanowire. This rotation simultaneously affects the spin-orbit and the Zeeman field directions in the TS wire such that both directions can never be parallel to each other. The junction spectral properties then follow again from a Dyson equation as in Eq. (15.41). Assuming that both wires have model parameters putting them deeply into the respective topological regime, we can compare our numerically exact results for the subgap spectral properties to the corresponding predictions of a minimal model Hamiltonian. The latter is obtained by retaining only the MBS degrees of freedom indicated in Fig. 79. To that end, the approximate expression for the boundary spinors can again be derived from the respective bGFs as in Sec. 15.7.1. Those spinors involve the Majorana operators $\gamma_{L1,L2,R}$ in Fig. 79 and are given by

$$\begin{aligned}\hat{\Psi}_L &\simeq \sqrt{\frac{\Delta}{t}} \begin{pmatrix} 1 \\ 0 \\ i \\ 0 \end{pmatrix} \gamma_{L1} + \sqrt{\frac{\Delta}{t}} \begin{pmatrix} 0 \\ i \\ 0 \\ 1 \end{pmatrix} \gamma_{L2}, \\ \hat{\Psi}_R &\simeq \sqrt{\frac{\Delta_2}{t}} R_y(\theta) \begin{pmatrix} i \\ -i \\ 1 \\ 1 \end{pmatrix} \gamma_R.\end{aligned}\quad (15.56)$$

The resulting minimal model Hamiltonian is

$$H_{\text{min}} = -i [w_1(\phi) \gamma_{L1} + w_2(\phi) \gamma_{L2}] \gamma_R \quad (15.57)$$

with the energies

$$\begin{aligned}w_1(\phi) &= \frac{2\lambda_L \sqrt{\Delta \Delta_2}}{t} \cos \frac{\phi}{2} \cos \frac{\theta}{2}, \\ w_2(\phi) &= -\frac{2\lambda_L \sqrt{\Delta \Delta_2}}{t} \sin \frac{\phi}{2} \sin \frac{\theta}{2}.\end{aligned}\quad (15.58)$$

The structure of H_{min} in Eq. (15.57) is similar to the minimal model (15.46) for the TS trijunction in Sec. 15.7.1 without any coupling between the $\gamma_{L1,L2}$ operators. The subgap spectrum

is therefore characterized by a decoupled zero-energy Majorana state, and the hybridization of the two other Majorana operators yields the Andreev bound state dispersion:

$$E_{\pm}(\phi) = \pm \frac{1}{2} \sqrt{w_1^2(\phi) + w_2^2(\phi)}. \quad (15.59)$$

We compare Eq. (15.59) to numerically exact results for the subgap spectral properties of the TRITOPS-TS junction in Fig. 80. Clearly, the general subgap spectrum is rather well described by the minimal model (15.57). In contrast to the case of a tri-terminal TS junction, for TRITOPS-TS junctions it is not necessary to take into account higher-order tunneling processes for obtaining accurate agreement with numerically exact bGF calculations (but see Ref. [85]).

15.8 On the spinful single-channel model

In this appendix, we provide technical details pertaining to our discussion of the spinful single-channel Majorana wire model [68, 69] in Sec. 15.3. First, the explicit form of the coefficients $C_n(\omega)$ in Eq. (15.14) is given by

$$\begin{aligned} C_0 = & \frac{3\alpha^4}{8} + \Delta^4 + \mu^4 - 8\mu^3t + 36\mu^2t^2 - 80\mu t^3 + 70t^4 - 2\mu^2V_x^2 + 8\mu tV_x^2 - 12t^2V_x^2 \\ & + V_x^4 - 2(\mu^2 - 4\mu t + 6t^2 + V_x^2)\omega^2 + \omega^4 + 2\Delta^2(\mu^2 - 4\mu t + 6t^2 - V_x^2 - \omega^2) + \\ & + \alpha^2 [\Delta^2 - \mu^2 + 4\mu t - 5t^2 + V_x^2 - \omega^2], \end{aligned} \quad (15.60)$$

$$\begin{aligned} C_1 = & -(\mu - 2t)t(\alpha^2 - 4(\Delta^2 + \mu^2 - 4\mu t + 7t^2 - V_x^2 - \omega^2)), \\ C_2 = & \left\{ -\alpha^4 + 8t^2(\Delta^2 + 3\mu^2 - 12\mu t + 14t^2 - V_x^2 - \omega^2) + 2\alpha^2 [-\Delta^2 + (\mu - 2t)^2 - V_x^2 + \omega^2] \right\} / 4, \\ C_3 = & (t\mu - 2t^2)(\alpha^2 + 4t^2), \end{aligned} \quad (15.61)$$

$$C_4 = [t^2 + (\alpha/2)^2]^2 \quad (15.62)$$

It is convenient to renormalize these coefficients such that C_4 appears as common factor of the polynomial.

The C_n coefficients in turn determine the coefficients $a_m(\omega)$ appearing in the eighth-order polynomial equation (15.15). The roots $z_n(\omega)$ therefore have satisfy the Vieta relations

$$\begin{aligned} S_k(z_1, \dots, z_8) &= \sum_{i_1 < i_2 < \dots < i_k} z_{i_1} z_{i_2} \dots z_{i_k} \\ &= (-1)^{8-k} \frac{a_k}{a_8}. \end{aligned} \quad (15.63)$$

Using the condition (15.20) and the ansatz (15.19), the first three invariants are given by $S_1 = 2AB$, $S_2 = 2(A^2 - 2)(1 + C) + 4B^2$, and $S_3 = 2AB(A^2 - 1) + 4ABC$ with

$$\begin{aligned} A &= 1 - \eta + \frac{1}{1 - \eta}, \quad B = \cos(\delta_1) + \cos(\delta_2), \\ C &= 2 \cos(\delta_1) \cos(\delta_2). \end{aligned} \quad (15.64)$$

As a consequence, the parameter C obeys a cubic equation that can be solved analytically,

$$w_3 + w_2 C + w_1 C^2 + C^3 = 0, \quad (15.65)$$

with the coefficients

$$\begin{aligned} w_1 &= 1 - \frac{S_3}{S_1}, \\ w_2 &= \frac{S_2}{4} - \frac{S_3}{S_1} - \frac{1}{4} + \left(\frac{S_3}{2S_1} \right)^2, \\ w_3 &= -\frac{S_2 S_3}{8S_1} - \frac{S_2}{8} + S_1^2 - \frac{1}{4} + \left(\frac{S_3}{2S_1} \right)^2. \end{aligned} \quad (15.66)$$

For $V_x < V_c$, the physical solution of Eq. (15.65) is given by

$$C = -2\sqrt{-Q} \cos(\theta_0/3) - w_1/3, \quad (15.67)$$

with

$$\begin{aligned} \theta_0 &= \cos^{-1} \left(-\frac{R}{\sqrt{-Q^3}} \right), \\ Q &= \frac{3a_2 - w_1^2}{9}, \\ R &= \frac{9w_1 w_2 - 27w_3 - 2w_1^3}{54}. \end{aligned} \quad (15.68)$$

For $V_x > V_c$, the solution is given by $C = P_1 - Q/P_1 - w_1/3$ (assuming $P_1 \neq 0$), with

$$P_1 = \text{sgn}(R) \left(|R| + \sqrt{R^2 + Q^3} \right)^{1/3}. \quad (15.69)$$

The coefficients A and B then follow from

$$A^2 = \frac{S_3}{S_1} + 1 - 2C, \quad B = \frac{S_1}{2A}. \quad (15.70)$$

Finally, the parameters in our ansatz [see Eqs. (15.19) and (15.21)], can be determined from the relations

$$\begin{aligned} \cos \delta_1 &= \frac{B + \sqrt{B^2 - 2C}}{2}, \\ \cos \delta_2 &= \frac{B - \sqrt{B^2 - 2C}}{2}, \\ \eta &= 1 - \frac{A}{2} + \sqrt{\frac{A^2}{4} - 1}. \end{aligned} \quad (15.71)$$

We proceed by providing the detailed form of the matrices $\hat{A}_\nu$ and $\hat{A}'_\nu$ in Eq. (15.25). Using the definition in the main text, for $V_x < V_c$, they are with $z_\nu(\omega)$ in Eq. (15.19) given by

$$\hat{A}_\nu = \frac{\hat{A}(z_\nu)}{b_\nu} + \frac{\hat{A}(z_\nu^*)}{b_\nu^*}, \quad \hat{A}'_\nu = \frac{\hat{A}'(z_\nu)}{b_\nu} + \frac{\hat{A}'(z_\nu^*)}{b_\nu^*}, \quad (15.72)$$

where an expansion of $\prod_{z_\nu \neq z_m} (z_\nu - z_m)$ to first order in $\tau_\nu \sqrt{\Delta_\nu^2 - \omega^2}$ yields

$$b_\nu = 32e^{3i\delta_\nu} \tau_\nu \sin^2(\delta_\nu) [\cos(\delta_2) - \cos(\delta_1)]^2. \quad (15.73)$$

Explicitly, the components of the symmetric 4×4 matrix $\hat{A}$ in Eq. (15.72), $\hat{A}_{ij} = \hat{A}_{ji}$, follow from

$$\begin{aligned}
\hat{A}_{11}(z) &= -\hat{A}_{33}(z) = z^3 V_x^2 [\epsilon(z) - \tilde{\alpha}(z)] + z^3 [\epsilon(z) + \tilde{\alpha}(z)] \{-\Delta^2 - (\epsilon(z) - \tilde{\alpha}(z))^2\}, \\
\hat{A}_{22}(z) &= -\hat{A}_{44}(z) = z^3 V_x^2 [\epsilon(z) + \tilde{\alpha}(z)] + z^3 [\epsilon(z) - \tilde{\alpha}(z)] \{-\Delta^2 - (\epsilon(z) + \tilde{\alpha}(z))^2\}, \\
\hat{A}_{12}(z) &= \hat{A}_{34}(z) = z^3 V_x [\Delta^2 + \epsilon^2(z) - \tilde{\alpha}^2(z) - V_x^2], \\
\hat{A}_{13}(z) &= z^3 \Delta \{V_x^2 - \Delta^2 - [\epsilon(z) - \tilde{\alpha}(z)]^2\}, \\
\hat{A}_{14}(z) &= -\hat{A}_{23}(z) = 2z^3 V_x \Delta \tilde{\alpha}(z), \\
\hat{A}_{24}(z) &= z^3 \Delta \{V_x^2 - \Delta^2 - [\epsilon(z) + \tilde{\alpha}(z)]^2\}.
\end{aligned} \tag{15.74}$$

Similarly, by taking a derivative with respect to ω , the nonvanishing matrix elements of the symmetric matrix $\hat{A}'_{ij} = \hat{A}'_{ji}$ follow as

$$\begin{aligned}
\hat{A}'_{11}(z) &= \hat{A}'_{33}(z) = -z^3 \{\Delta^2 + [\epsilon(z) - \alpha(z)]^2 + V_x^2\}, \\
\hat{A}'_{22}(z) &= \hat{A}'_{44}(z) = -z^3 \{\Delta^2 + [\epsilon(z) + \alpha(z)]^2 + V_x^2\}, \\
\hat{A}'_{12}(z) &= -\hat{A}'_{34}(z) = 2z^3 V_x \epsilon(z), \\
\hat{A}'_{14}(z) &= \hat{A}'_{23}(z) = 2z^3 V_x \Delta.
\end{aligned} \tag{15.75}$$

In the topologically nontrivial phase, $V_x > V_c$, trigonometric functions associated with the roots $z_{1,\pm}$ in Eq. (15.21) turn into hyperbolic functions. The matrices with $\nu = 1$ in Eq. (15.72) are then replaced by

$$\begin{aligned}
\hat{A}_1 &= \frac{\hat{A}(z_{1,+})}{\tilde{b}_1} - \frac{\hat{A}(z_{1,-})}{\tilde{b}_1}, \\
\hat{A}'_1 &= \frac{\hat{A}'(z_{1,+})}{\tilde{b}_1} - \frac{\hat{A}'(z_{1,-})}{\tilde{b}_1},
\end{aligned} \tag{15.76}$$

with the quantities

$$\begin{aligned}
\tilde{b}_1 &= 32e^{-3\delta_1} \tau_1 \sinh^2(\delta_1) [\cos(\delta_2) - \cosh(\delta_1)]^2, \\
\tilde{b}_2 &= 32e^{3i\delta_2} \tau_2 \sin^2(\delta_2) [\cos(\delta_2) - \cosh(\delta_1)]^2.
\end{aligned} \tag{15.77}$$

The $\nu = 2$ matrices follow from Eq. (15.72) with the replacement $b_2 \rightarrow \tilde{b}_2$. Finally, we note that for very large V_x , one approaches the Kitaev limit of the nanowire, and the relevant residues come from the z_2 roots only.

15.9 On TRITOPS wires

According to Theorem 1 of Ref. [76], the largest-modulus root $z_{\max}$ inside the unit circle can be determined from the relative position of the origin inside the ellipse discussed in Sec. 15.6.1. For that purpose, we first determine the major (M) and minor (m) axes of the ellipse in Fig. 99. Using Eq. (15.34) and focusing on the case of $\hat{\mathcal{H}}_-(k)$, the defining equation of the ellipse is given by

$$B^T \begin{pmatrix} (t^2 + \alpha^2)/\Delta^2 & t/\Delta \\ t/\Delta & 1 \end{pmatrix} B = 4\alpha^2 \tag{15.78}$$

with $B^T = (\beta_{-,x}, \beta_{-,z} + \mu)$. From the eigenvalues of the 2×2 matrix in Eq. (15.78),

$$\lambda_{\pm} = \frac{t^2 + \Delta^2 + \alpha^2}{2\Delta^2} \pm \frac{\sqrt{(t^2 + \Delta^2)^2 + 2(t^2 - \Delta^2)\alpha^2 + \alpha^4}}{2\Delta^2}, \tag{15.79}$$

we obtain

$$m = 4\alpha/\sqrt{\lambda_+}, \quad M = 4\alpha/\sqrt{\lambda_-}. \quad (15.80)$$

The distance between the foci of the ellipse then follows as $f = \sqrt{M^2 - m^2}$.

To obtain the distance $l = |OF_1| + |OF_2|$ between the foci and the origin (corresponding to the red dashed line in Fig. 99), we first compute the rotation angle θ of the ellipse using the eigenvectors of the conic section matrix,

$$\begin{aligned} \cos \theta &= \frac{1}{\sqrt{1 + X^2/(2t\Delta)^2}}, \\ X &= \Delta^2 - t^2 - \alpha^2 - \sqrt{(t^2 + \Delta^2)^2 + 2(t^2 - \Delta^2)\alpha^2 + \alpha^4}. \end{aligned} \quad (15.81)$$

As a consequence, l follows from the relation

$$|OF_{1,2}| = \sqrt{(f/2)^2 + \mu^2 \pm \mu f \sin \theta}. \quad (15.82)$$

The largest-modulus root inside the unit circle is then given by (see Refs. [76, 77])

$$|z_{\max}| = \frac{l + \sqrt{l^2 - f^2}}{M + m}. \quad (15.83)$$

The same result follows for the other block, $\hat{\mathcal{H}}_+(k)$. As discussed in Sec. 15.6.1, Eq. (15.83) determines the decay length of Majorana end states into the bulk of a TRITOPS wire.

15.10 Easy illustrative examples

15.10.1 Example 1

15.10.2 Example 2

16 Conductance calculations for quantum wires and interfaces: mode matching and Green functions by Khomyakov, Brocks, Karpan, Zwierzycki, Kelly

Landauer's formula relates the conductance of a quantum wire or interface to transmission probabilities. Total transmission probabilities are frequently calculated using Green function techniques and an expression first derived by Caroli. Alternatively, partial transmission probabilities can be calculated from the scattering wave functions that are obtained by matching the wave functions in the scattering region to the Bloch modes of ideal bulk leads. An elegant technique for doing this, formulated originally by Ando, is here generalized to any Hamiltonian that can be represented in tight-binding form. A more compact expression for the transmission matrix elements is derived and it is shown how all the Green function results can be derived from the mode matching technique. We illustrate this for a simple model which can be studied analytically, and for an Fe|vacuum|Fe tunnel junction which we study using first-principles calculations.

16.1 Summary and Conclusions

We have demonstrated the equivalence of the mode matching and the Green function approaches to calculating the conductance of quantum wires and interfaces. In the mode matching technique, the scattering problem is solved by matching the wave function in the scattering

region to the Bloch modes in the leads. The technique is formulated for tight-binding Hamiltonians and covers all representations that can be expressed in tight-binding form, including first-principles implementations using localized orbital basis sets or a real space grid.

Alternatively, the scattering information can be extracted from the Green function, which is calculated by partitioning the system into a scattering region and leads. We demonstrate that the Green function technique can be reformulated in terms of mode matching. In addition we prove that the mode matching expression for the transmission matrix does not depend on where in the leads the scattering region is matched to the ideal leads, which was called into question in Ref. [22].

Calculating the full transmission matrix allows us to study individual scattering amplitudes from and to every possible mode. We have derived a compact expression for the transmission matrix elements, Eq. (16.67). Only propagating modes enter this expression, since evanescent modes do not contribute to the transmission directly. The evanescent modes are important, however, for setting up the correct matching conditions at the boundaries between the scattering region and the leads. Alternatively, the total transmission probability can be calculated from the Caroli expression, Eq. (16.53). Formally this expression sums over all possible modes, i.e. propagating and evanescent. However, by deriving the Caroli expression from the mode matching approach we show explicitly that only the propagating modes give a non-zero contribution.

The mode matching approach and its equivalence to the Green function approach were illustrated for a simple analytical model, as well as by numerical calculations on an Fe|vacuum|Fe tunnel junction. In the latter we treat the electronic structure within DFT, using a TB-MTO-ASA basis set. We demonstrate that the layer densities of states that follow from the mode matching and Green function approaches are numerically indistinguishable. We identify the Fe(001) surface state in the density of states and establish its contribution to the transmission through the tunnel junction. Finally, the invariance property discussed above is demonstrated numerically on the Fe|vacuum|Fe tunnel junction.

16.2 Introduction

Since the discovery of giant magnetoresistance in metallic multilayers there has been considerable interest in studying electronic transport in layered materials.[1, 2] At the same time, experimental control of the lateral scale has enabled studies of electronic transport in quantum wires of atomic dimensions.[3] Because of the small dimensions involved, the transport properties of both of these systems should be understood on the basis of their atomic structure. This perception has generated a large effort in recent years to calculate the conductance of multilayers and quantum wires from first principles. Several different approaches have been formulated which have a common basis in the Landauer-Büttiker approach or are equivalent to it. In the linear response regime, the conductance $\mathcal{G}$ is expressed as a quantum mechanical scattering problem[4] and can be simply related to the total transmission probability at the Fermi energy, $T(E_F)$, as

$$\mathcal{G} = \frac{e^2}{h} T(E_F). \quad (16.1)$$

The multilayer or quantum wire is generally considered as a scattering region of finite size, sandwiched between two semi-infinite ballistic wires. Aiming at a materials specific description, most current approaches treat the electronic structure within the framework of density functional theory (DFT)[5, 6, 7, 8, 9, 10, 11, 12, 13, 14].

Frequently the conductance is calculated using a Green function expression first derived by Caroli *et al.*[15, 16] An alternative technique, suitable for Hamiltonians that can be represented in tight-binding form, has been formulated by Ando.[17] It is based upon directly matching the

wave function in the scattering region to the Bloch modes of the leads. The latter technique has been applied to conductance calculations at the empirical tight-binding level, [18] as well as on the first-principles DFT level. [8, 19, 20, 21] The relationship between the mode matching[17] and Green function [15, 16, 22] approaches is not immediately obvious. Indeed, it was recently stated that Ando's approach is incomplete and does not yield the correct expression for the conductance. [22]

In this paper we demonstrate that the two approaches are completely equivalent. In the Green function approach, a small imaginary part must be added to or subtracted from the energy in order to distinguish between the retarded and advanced forms.[5, 6, 7, 10, 11, 12, 13, 14, 22] In mode matching, scattering wave functions are calculated that incorporate the retarded or advanced boundary conditions directly. This makes it possible to solve the scattering problem also at real, instead of complex energies. In addition to yielding the total conductance, by focussing on wave functions the contribution of each individual scattering channel can be identified. In particular, we derive a simple, compact expression for the transmission matrix elements, see Eq. (16.67).

The paper is organized as follows. In the next section the Hamiltonian we will use is introduced. This model allows us to study both quantum wires that are finite in the directions perpendicular to the wire, and systems that are periodic in these directions such as single interfaces, sandwiches and multilayers. We will use the single term "quantum wire" to describe both systems. In Secs. 16.4 and 16.5 the mode matching and Green function techniques are summarized. The equivalence of the transmission matrices obtained using these two techniques is demonstrated in Sec. 16.6 and the Caroli expression for the conductance is derived from the mode matching expressions. In Sec. 16.7 the two techniques are applied first to a simple analytical model,[23] and then to an Fe|vacuum|Fe tunnel junction using numerical first-principles calculations. The main conclusions are summarized in Sec. 16.1.

16.3 Hamiltonian

We set up a tight-binding representation of the Hamiltonian. This is not a severe restriction since a first-principles DFT implementation that uses a representation on an atomic orbital basis set has the same mathematical structure as a tight-binding model.[24, 25] Alternatively, an implementation that uses a representation of the Hamiltonian on a grid in real space, can also be mapped onto a tight-binding model.[26] We begin by dividing the system into slices ("principal layers") perpendicular to the wire direction. [27] The thickness of these slices is chosen such that there is only an interaction between neighboring slices. Labelling each slice with an index i , the Schrödinger equation of the quantum wire becomes

$$-\mathbf{H}_{i,i-1}\mathbf{c}_{i-1} + (E\mathbf{I} - \mathbf{H}_{i,i})\mathbf{c}_i - \mathbf{H}_{i,i+1}\mathbf{c}_{i+1} = 0, \quad (16.2)$$

for $i = -\infty, \dots, \infty$. Assuming that each slice contains N sites and/or orbitals, $\mathbf{c}_i$ is a vector of dimension N containing the wave function coefficients on all sites and/or orbitals of slice i . The $N \times N$ matrices $\mathbf{H}_{i,i}$ and $\mathbf{H}_{i,i\pm 1}$ consist of on-slice and hopping matrix elements of the Hamiltonian, respectively. $\mathbf{I}$ is the $N \times N$ identity matrix. A schematic representation of the structure of the Hamiltonian is given in Fig. 81.

Eq. (16.2) is valid both for quantum wires that are finite in the directions parallel to the slices, and for layered systems that are periodic in these directions. In the latter case, translations in the transverse direction can be described in terms of a Bloch wave vector in the two-dimensional Brillouin zone, $\mathbf{k}_{\parallel}$, which is a good quantum number and the system becomes effectively one dimensional. Explicit expressions for the Hamilton matrix elements depend upon the particular localized orbital basis or real space grid representation used.[28] Since details of the tight-binding muffin tin orbital scheme used in Refs. [8, 19, 20, 21] are given in Ref. [25]

and of the real-space high-order finite difference method can be found in Ref. [26], they will not be discussed further here.

The system is divided into three parts, with $i = -\infty, \dots, 0$ corresponding to the left lead (L), $i = 1, \dots, S$ to the scattering region (S) and $i = S + 1, \dots, \infty$ to the right lead (R). The leads are assumed to be ideal wires characterized by a periodic potential. It is then sufficient to identify a slice with a translational period along the wire. By construction, the Hamiltonian matrix is the same for each slice of the leads, i.e. $\mathbf{H}_{i,i} \equiv \mathbf{H}_{L/R}$, $\mathbf{H}_{i,i-1} \equiv \mathbf{B}_{L/R}$ and $\mathbf{H}_{i,i+1} \equiv \mathbf{B}_{L/R}^\dagger$ for the left/right leads. Fig. 81 summarizes our model of a quantum wire.

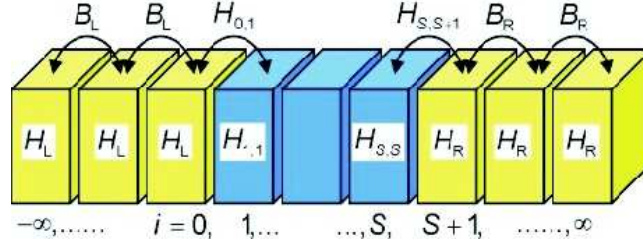


Figure 81: Hamiltonian matrix of a quantum wire divided into slices. The left (L) and right (R) leads are ideal periodic wires that span the cells $i = -\infty, \dots, 0$ and $i = S + 1, \dots, \infty$, respectively. The scattering region spans cells $i = 1, \dots, S$.

16.4 Mode matching approach

Eq. (16.2) can be solved in a particularly convenient way by a technique which we call mode matching. In this, we follow Ando and generalize his approach to a slice geometry.[17]

16.4.1 Bloch matrices

The first step consists of finding solutions for the leads, for which Eq. (16.2) can be simplified to

$$-\mathbf{B}\mathbf{c}_{i-1} + (E\mathbf{I} - \mathbf{H})\mathbf{c}_i - \mathbf{B}^\dagger\mathbf{c}_{i+1} = 0. \quad (16.3)$$

These equations hold for $i = -\infty, \dots, -1$ and $i = S + 2, \dots, \infty$, i.e. the left and right leads. To keep the notation as simple as possible, we have omitted the subscripts L and R, see Fig. 81. Since the leads are periodic wires, one can make the ansatz that the solutions have Bloch symmetry, i.e. $\mathbf{c}_i = \lambda\mathbf{c}_{i-1}$, $\mathbf{c}_{i+1} = \lambda^2\mathbf{c}_{i-1}$, where λ is the Bloch factor. Substituting this into Eq. (16.3) results in a quadratic eigenvalue equation of dimension N . The latter can be solved most easily by transforming it to an equivalent linear (generalized) eigenvalue problem of dimension $2N$ and solving this by a standard algorithm. [29, 30]

It can be shown that the equation generally has $2N$ solutions, which can be divided into N right-going modes and N left-going modes, [31] labelled as (+) and (−) in the following. Right-going modes are either evanescent waves that are decaying to the right, or waves of constant amplitude that are propagating to the right, whereas left-going modes are decaying or propagating to the left. We denote the eigenvalues by $\lambda_n(\pm)$ where $n = 1, \dots, N$, the corresponding eigenvectors by $\mathbf{u}_n(\pm)$ and write the eigenvalue equation as

$$-\mathbf{B}\mathbf{u}_n(\pm) + (E\mathbf{I} - \mathbf{H})\lambda_n(\pm)\mathbf{u}_n(\pm) - \mathbf{B}^\dagger\lambda_n(\pm)^2\mathbf{u}_n(\pm) = 0. \quad (16.4)$$

In the following we assume that the vectors $\mathbf{u}_n(\pm)$ are normalized; note that in general they are *not* orthogonal.

One can distinguish right- from left-going evanescent modes on the basis of their eigenvalues; right going evanescent modes have $|\lambda(+)| < 1$ and left going evanescent modes have $|\lambda(-)| > 1$.

Propagating modes have $|\lambda(\pm)| = 1$, so here one has to determine the Bloch velocity and use its sign to distinguish right from left propagation. The Bloch velocities are given by the expression

$$v_n(\pm) = -\frac{2a}{\hbar} \text{Im} [\lambda_n(\pm) \mathbf{u}_n(\pm)^\dagger \mathbf{B}^\dagger \mathbf{u}_n(\pm)], \quad (16.5)$$

where a is the slice thickness. A derivation of this equation is given in Appendix A.

Since the eigenvectors are non-orthogonal, it is convenient to define dual vectors $\tilde{\mathbf{u}}_n(\pm)$

$$\tilde{\mathbf{u}}_n^\dagger(\pm) \mathbf{u}_m(\pm) = \delta_{n,m}; \quad \mathbf{u}_n^\dagger(\pm) \tilde{\mathbf{u}}_m(\pm) = \delta_{n,m}. \quad (16.6)$$

Any wave function in the leads can be expressed as a linear combination of the lead modes. This can be done in a compact form using the two $N \times N$ Bloch matrices for right- and left-going solutions

$$\mathbf{F}(\pm) = \sum_{n=1}^N \lambda_n(\pm) \mathbf{u}_n(\pm) \tilde{\mathbf{u}}_n^\dagger(\pm). \quad (16.7)$$

For any integer i , $\mathbf{F}^i$ is given by a similar expression but with λ_n replaced by λ_n^i . From the foregoing it is easy to show that the Bloch matrices obey the equation

$$-\mathbf{B}\mathbf{F}^{-1}(\pm) + (E\mathbf{I} - \mathbf{H}) - \mathbf{B}^\dagger \mathbf{F}(\pm) = 0. \quad (16.8)$$

A general solution in the leads can be expressed in terms of a recursion relation

$$\begin{aligned} \mathbf{c}_i &= \mathbf{c}_i(+) + \mathbf{c}_i(-) \\ &= \mathbf{F}^{i-j}(+) \mathbf{c}_j(+) + \mathbf{F}^{i-j}(-) \mathbf{c}_j(-). \end{aligned} \quad (16.9)$$

Fixing the coefficients in one slice then sets the boundary conditions and determines the solution for the whole lead.

16.4.2 Transmission matrix

The scattering region is defined by $i = 1, \dots, S$, see Fig. 81. Right and left of the scattering regions one has the recursion relations for the states in the leads from Eq. (16.9)

$$\begin{aligned} \mathbf{c}_{-1} &= \mathbf{F}_L^{-1}(+) \mathbf{c}_0(+) + \mathbf{F}_L^{-1}(-) \mathbf{c}_0(-) \\ &= [\mathbf{F}_L^{-1}(+) - \mathbf{F}_L^{-1}(-)] \mathbf{c}_0(+) + \mathbf{F}_L^{-1}(-) \mathbf{c}_0, \end{aligned} \quad (16.10)$$

with $\mathbf{c}_0 = \mathbf{c}_0(+) + \mathbf{c}_0(-)$ and

$$\mathbf{c}_{S+2} = \mathbf{F}_R(+) \mathbf{c}_{S+1}(+) + \mathbf{F}_R(-) \mathbf{c}_{S+1}(-), \quad (16.11)$$

where the subscripts L and R distinguish between the Bloch matrices of the left and right leads.

The boundary conditions for the scattering problem are set up in the usual way. The vector $\mathbf{c}_0(+)$ is treated as the source, i.e. as a fixed incoming wave from the left lead. There is no incoming wave from the right lead, so we set $\mathbf{c}_{S+1}(-) = 0$.

Having set the boundary conditions, Eqs. (16.10) and (16.11) can be used to rewrite Eq. (16.2) in the region not covered by Eq. (16.3), i.e. for $i = 0, \dots, S+1$. This describes the wave function in the scattering region and the matching to the solutions in the leads. Eq. (16.2) in this region is rewritten as

$$-\mathbf{H}'_{i,i-1} \mathbf{c}_{i-1} + (E\mathbf{I} - \mathbf{H}'_{i,i}) \mathbf{c}_i - \mathbf{H}'_{i,i+1} \mathbf{c}_{i+1} = \mathbf{Q}_i \mathbf{c}_0(+), \quad (16.12)$$

with a modified Hamilton matrix defined so $\mathbf{H}'_{i,j} = \mathbf{H}_{i,j}$ for the elements $\{i, j = 0, 1\}$, $\{i = 1, \dots, S; j = i, i \pm 1\}$ and $\{i, j = S + 1, S\}$, but with

$$\begin{aligned}\mathbf{H}'_{0,0} &= \mathbf{H}_L + \mathbf{B}_L \mathbf{F}_L^{-1}(-), \\ \mathbf{H}'_{S+1,S+1} &= \mathbf{H}_R + \mathbf{B}_R^\dagger \mathbf{F}_R(+).\end{aligned}\quad (16.13)$$

$\mathbf{H}'_{0,-1} = 0$ and $\mathbf{H}'_{S+1,S+2} = 0$, so the modified scattering region is decoupled from the leads. On the left-hand side of Eq. (16.12), we have a source term with

$$\mathbf{Q}_0 = \mathbf{B}_L [\mathbf{F}_L^{-1}(+) - \mathbf{F}_L^{-1}(-)], \quad (16.14)$$

and $\mathbf{Q}_i = 0$ for $i = 1, \dots, S + 1$. Eq. (16.12) defines a set of $(S + 2) \times N$ linear equations. Because the Hamilton matrix is block tridiagonal, each block being of dimension N , this set of equations can be solved efficiently using a block Gaussian elimination scheme.[30] The total wave function $\mathbf{c}_i$ can then be obtained by back substitution.

The transmission is obtained from the wave function in the right lead $\mathbf{c}_{S+1}(+)$. In particular, choosing the incoming wave as one of the modes of the left lead, i.e. $\mathbf{c}_0(+) = \mathbf{u}_{L,m}(+)$, generalized transmission matrix elements $\tau_{n,m}$ are defined by expanding $\mathbf{c}_{S+1}(+)$ in modes of the right lead

$$\mathbf{c}_{S+1}(+) = \sum_{n=1}^N \mathbf{u}_{R,n}(+) \tau_{n,m}. \quad (16.15)$$

By letting $\mathbf{c}_0(+) run over all possible incoming modes of the left lead $\mathbf{u}_{L,m}(+)$; $m = 1, \dots, N$, a full transmission matrix is obtained.$

Matrix elements can be defined for all modes, propagating and evanescent, but of course only matrix elements where n, m denote propagating modes contribute to the real physical transmission. These modes can be selected by making use of their eigenvalues; see the discussion following Eq. (16.4). The physical transmission matrix elements are then found by normalizing with respect to the current[16]

$$t_{n,m} = \sqrt{\frac{v_{R,n}(+)a_L}{v_{L,m}(+)a_R}} \tau_{n,m}, \quad (16.16)$$

where $v_{L,m}(+)$ and $v_{R,n}(+)$ are the Bloch velocities in the direction of the wire for the right-propagating modes m and n in the left and right leads, respectively, see Eq. (16.5); a_L and a_R are the slice thicknesses of left and right leads.[32] The total transmission probability is given by

$$T(E) = \sum_{n,m}^{(+)} |t_{n,m}|^2, \quad (16.17)$$

and the conductance is given by Eq. (16.1) evaluated at $E = E_F$.

16.4.3 Green function matrix

Solving the set of linear equations Eq. (16.12) directly leads to the conductance. However, to facilitate a connection to the Green function approach discussed in the next section, we can formulate the solution in a slightly different way. A finite Green function matrix $\mathbf{G}'_{i,j}(z)$, $i, j = 0, \dots, S + 1$ can be defined by

$$-\mathbf{H}'_{i,i-1} \mathbf{G}'_{i-1,j} + (z\mathbf{I} - \mathbf{H}'_{i,i}) \mathbf{G}'_{i,j} - \mathbf{H}'_{i,i+1} \mathbf{G}'_{i+1,j} = \mathbf{I} \delta_{i,j}, \quad (16.18)$$

with z complex. Note, however, that the matrices $\mathbf{H}'_{0,0}$ and $\mathbf{H}'_{S+1,S+1}$ are non-Hermitian and $\mathbf{G}'_{i,j}(E)$ is also uniquely defined for real energies. The Green function matrix allows the solution of Eq. (16.12) to be written as

$$\mathbf{c}_i = \mathbf{G}'_{i,0}(E) \mathbf{Q}_0 \mathbf{c}_0(+). \quad (16.19)$$

As before, the transmission can be extracted at $i = S + 1$ and comparison with Eq. (16.15) gives

$$\tau_{n,m} = \tilde{\mathbf{u}}_{R,n}^\dagger(+) \mathbf{G}'_{S+1,0}(E) \mathbf{Q}_0 \mathbf{u}_{L,m}(+). \quad (16.20)$$

which can be used in Eq. (16.16). The Green function matrix block $\mathbf{G}'_{S+1,0}(E)$ can be found by solving Eq. (16.18) using a recursive algorithm that resembles a Gaussian elimination scheme.[25, 33]

16.5 Green function approach

An apparently quite different route to the transmission matrix starts by defining an infinite Green function matrix $\mathbf{G}_{i,j}(z)$ for $i, j = -\infty, \dots, \infty$ with respect to the original Hamiltonian of Eq. (16.2).

$$-\mathbf{H}_{i,i-1} \mathbf{G}_{i-1,j} + (z\mathbf{I} - \mathbf{H}_{i,i}) \mathbf{G}_{i,j} - \mathbf{H}_{i,i+1} \mathbf{G}_{i+1,j} = \mathbf{I} \delta_{i,j}. \quad (16.21)$$

Choosing $z = \lim_{\eta \rightarrow 0} (E \pm i\eta)$ defines as usual the retarded/advanced Green function matrix. We shall use $\mathbf{G}_{i,j}(E)$ to denote the *retarded* Green function matrix and $\mathbf{G}_{i,j}^a(E)$ to denote the *advanced* Green function matrix.

16.5.1 Partitioning

Eq. (16.21) is most conveniently solved by applying a partitioning technique.[16, 34] It is straightforward to show that the finite part $\mathbf{G}_{i,j}(z)$ defined for $i, j = 0, \dots, S+1$ can be derived from a closed set of equations, similar to Eq. (16.21), but with $\mathbf{H}_{i,j}$ replaced by $\mathbf{H}_{i,j}''$ where $\mathbf{H}_{i,j}'' = \mathbf{H}_{i,j}$ for the elements $\{i, j = 0, 1\}$, $\{i = 1, \dots, S; j = i, i \pm 1\}$ and $\{i, j = S+1, S\}$, but with

$$\begin{aligned} \mathbf{H}_{0,0}''(z) &= \mathbf{H}_L + \mathbf{B}_L \mathbf{g}_L(z) \mathbf{B}_L^\dagger, \\ \mathbf{H}_{S+1,S+1}''(z) &= \mathbf{H}_R + \mathbf{B}_R^\dagger \mathbf{g}_R(z) \mathbf{B}_R. \end{aligned} \quad (16.22)$$

Here $\mathbf{g}_L(z)$ and $\mathbf{g}_R(z)$ are the surface Green functions of the semi-infinite left and right leads, respectively, which can be calculated using an iterative technique: denoting $\mathbf{G}_{i,j}^{[n]}(z)$ as the solution of an equation similar to Eq. (16.21), but with $\mathbf{H}_{i,j} = 0$ for $\{i > n \vee j > n\}$, one can easily derive the right-going recursion relation

$$\left[z\mathbf{I} - \mathbf{H}_{n+1,n+1} - \mathbf{H}_{n+1,n} \mathbf{G}_{n,n}^{[n]}(z) \mathbf{H}_{n,n+1} \right] \mathbf{G}_{n+1,n+1}^{[n+1]}(z) = \mathbf{I}. \quad (16.23)$$

For an ideal wire with $i, j = -\infty, \dots, n$, $\mathbf{G}_{n,n}^{[n]}(z) = \mathbf{g}_L(z)$ should be independent of n resulting in the following equation for the surface Green function,

$$\left[z\mathbf{I} - \mathbf{H}_L - \mathbf{B}_L \mathbf{g}_L(z) \mathbf{B}_L^\dagger \right] \mathbf{g}_L(z) = \mathbf{I}. \quad (16.24)$$

Several iterative algorithms have been formulated for solving this non-linear matrix equation.[27, 35, 36] A similar reasoning based upon a left-going recursion for the right lead results in an equation for the surface Green function $\mathbf{g}_R(z)$ of the right lead

$$\left[z\mathbf{I} - \mathbf{H}_R - \mathbf{B}_R^\dagger \mathbf{g}_R(z) \mathbf{B}_R \right] \mathbf{g}_R(z) = \mathbf{I}. \quad (16.25)$$

Again, setting $z = E + i\eta$ in Eqs. (16.24) and (16.25) defines the usual retarded surface Green functions $\mathbf{g}_{L/R}(E)$. Although we are mainly interested in the physical limit $\lim_{\eta \rightarrow 0}$, in practice a finite value of η is often used in order to make the iterative algorithms stable.

The quantities

$$\Sigma_L(E) = \mathbf{B}_L \mathbf{g}_L(E) \mathbf{B}_L^\dagger; \quad \Sigma_R(E) = \mathbf{B}_R^\dagger \mathbf{g}_R(E) \mathbf{B}_R, \quad (16.26)$$

which appear in Eqs. (16.22)-(16.25), are called the self-energies of the left and right leads, respectively.[16] Once these are obtained, the finite Hamilton matrix of Eq. (16.22) is constructed and the retarded Green function matrix $\mathbf{G}_{i,j}(E)$ can be found using a recursive algorithm.[33] Using the lead modes the transmission matrix elements can then be calculated, as will be shown in the next section, Sec. 16.5.2. Alternatively, the total transmission probability can be expressed in a form that does not require the lead modes explicitly, which is discussed in Sec. 16.6.

16.5.2 Transmission matrix

The transmission matrix can be obtained from the Green function matrix of Eq. (16.21). To do this, we adapt a Fisher-Lee type of approach to our tight-binding formulation.[37] Assuming that the unperturbed reference wave function is the Bloch mode $\mathbf{u}_{L,m}(+)$ that comes in from the left lead, the Lippmann-Schwinger equation[38] in tight-binding form is

$$\begin{aligned} \mathbf{c}_i &= \mathbf{u}_{L,m,i}(+) + \sum_{j,k} \mathbf{G}_{i,j} \mathbf{V}_{j,k} \mathbf{u}_{L,m,k}(+) \\ &= \left[\mathbf{F}_L^i(+) + \sum_{j,k} \mathbf{G}_{i,j} \mathbf{V}_{j,k} \mathbf{F}_L^k(+) \right] \mathbf{u}_{L,m}(+). \end{aligned} \quad (16.27)$$

Here $\mathbf{u}_{L,m,i}(+)$ is the reference wave function in slice i . It obeys Bloch symmetry and $\mathbf{u}_{L,m}(+) \equiv \mathbf{u}_{L,m,0}(+)$ is the Bloch mode at the origin, see Sec. 16.4.1. The matrix $\mathbf{V}_{j,k}$ represents the perturbation with respect to the ideal left lead.

Eq. (16.27) can be simplified using the Dyson equation, which in tight-binding form reads

$$\begin{aligned} \mathbf{G}_{i,0} &= \mathbf{G}_{i,0}^{(0)} + \sum_{j,k} \mathbf{G}_{i,j} \mathbf{V}_{j,k} \mathbf{G}_{k,0}^{(0)} \\ &= \left[\mathbf{F}_L^i(+) + \sum_{j,k} \mathbf{G}_{i,j} \mathbf{V}_{j,k} \mathbf{F}_L^k(+) \right] \mathbf{G}_{0,0}^{(0)}, \end{aligned} \quad (16.28)$$

using Eq. (16.32). Comparing Eqs. (16.27) and (16.28) one finds the simple expression

$$\mathbf{c}_i = \mathbf{G}_{i,0}(E) \left[\mathbf{G}_{0,0}^{(0)}(E) \right]^{-1} \mathbf{u}_{L,m}(+). \quad (16.29)$$

From the definition of the generalized transmission matrix elements, cf. Eq. (16.15), one then obtains the expression

$$\tau_{n,m} = \tilde{\mathbf{u}}_{R,n}^\dagger(+) \mathbf{G}_{S+1,0}(E) \left[\mathbf{G}_{0,0}^{(0)}(E) \right]^{-1} \mathbf{u}_{L,m}(+). \quad (16.30)$$

To find $\tau_{n,m}$ one needs to calculate only the Green function matrix blocks $\mathbf{G}_{S+1,0}(E)$ of the full system and $\mathbf{G}_{0,0}^{(0)}(E)$ of the ideal left lead. The physical transmission matrix elements and the total transmission probability can then be obtained from Eqs. (16.16) and (16.17).

16.6 Mode matching versus Green functions

The two seemingly different formalisms introduced in Secs. 16.4 and 16.5 are in fact closely related. In this section we will show how all Green function results can be obtained from the mode matching approach. We begin by expressing the Green function matrices of ideal wires in terms of the Bloch matrices, $\mathbf{F}(\pm)$. These expressions are then used to prove that the transmission matrix elements obtained from the mode matching and Green function approaches, cf. Eqs. (16.20) and (16.30), are in fact identical. After that, we show that the transmission matrix elements are independent of the exact positions within the leads that are used to match the leads to the scattering region, apart from a trivial phase factor. Then we derive from the mode matching expression for the total transmission probability the Green function expression known as the Caroli expression.[15] Finally, a more compact expression for the transmission matrix elements is derived.

16.6.1 Green functions of ideal wires in terms of Bloch matrices

We begin by deriving an expression for the retarded Green function matrix $\mathbf{G}_{i,j}^{(0)}$ of an *ideal infinite wire* in terms of its eigenmodes. The columns of such a Green function obey the equation

$$-\mathbf{B}\mathbf{G}_{i-1,j}^{(0)} + (E^+\mathbf{I} - \mathbf{H})\mathbf{G}_{i,j}^{(0)} - \mathbf{B}^\dagger\mathbf{G}_{i+1,j}^{(0)} = \mathbf{I}\delta_{i,j}, \quad (16.31)$$

where $E^+ = E + i\eta$. For $i \neq j$ the solution is similar to that of the wave functions, see Eq. (16.3). In addition, the retarded Green function should consist only of propagating waves that move outwards from the δ -source and/or evanescent states that decay away from the source.[38] From Eq. (16.9), we have the *column* recursion relations

$$\begin{aligned} \mathbf{G}_{i,j}^{(0)}(E) &= \mathbf{F}^{i-j}(-)\mathbf{G}_{j,j}^{(0)}(E), \quad i < j, \\ \mathbf{G}_{i,j}^{(0)}(E) &= \mathbf{F}^{i-j}(+)\mathbf{G}_{j,j}^{(0)}(E), \quad i > j. \end{aligned} \quad (16.32)$$

The diagonal block $\mathbf{G}_{j,j}^{(0)}(E)$ can now be obtained by combining Eqs. (16.31) and (16.32), which gives for $i = j$

$$[-\mathbf{B}\mathbf{F}^{-1}(-) + E^+\mathbf{I} - \mathbf{H} - \mathbf{B}^\dagger\mathbf{F}(+)]\mathbf{G}_{j,j}^{(0)} = \mathbf{I}. \quad (16.33)$$

Eliminating $E^+\mathbf{I} - \mathbf{H}$ using Eq. (16.8) then yields

$$[\mathbf{G}_{j,j}^{(0)}(E)]^{-1} = \mathbf{B} [\mathbf{F}^{-1}(+) - \mathbf{F}^{-1}(-)], \quad (16.34)$$

or the equivalent

$$[\mathbf{G}_{j,j}^{(0)}(E)]^{-1} = \mathbf{B}^\dagger [\mathbf{F}(-) - \mathbf{F}(+)], \quad (16.35)$$

Eqs. (16.32) and (16.34) represent the full expression for the Green function $\mathbf{G}_{i,j}^{(0)}$ of an infinite ideal wire in terms of the Bloch matrices $\mathbf{F}(\pm)$ and thus in terms of the eigenmodes. Note that we can set $E^+ = E$ since, in terms of the modes, the retarded Green function matrix is uniquely defined for real energies.

The *advanced* Green function matrix $\mathbf{G}_{i,0}^{(0)a}(E)$ can be found from a similar procedure. It should consist of propagating waves that move towards the source and/or evanescent states that grow towards the source. One can construct two new Bloch matrices $\mathbf{F}^a(\pm)$, which are similar to those defined in Eq. (16.7). In $\mathbf{F}^a(+)$ one collects the modes that are decaying to the right (growing to the left) and modes that are propagating to the left. $\mathbf{F}^a(-)$ then contains

modes that grow or propagate to the right.

$$\begin{aligned}\mathbf{F}^a(\pm) &= \sum_{n=1}^N \lambda_n^a(\pm) \mathbf{u}_n^a(\pm) \tilde{\mathbf{u}}_n^{a\dagger}(\pm), \text{ with} \\ \lambda_n^a(\pm) &= \lambda_n(\mp), \mathbf{u}_n^a(\pm) = \mathbf{u}_n(\mp) \text{ propagating} \\ \lambda_n^a(\pm) &= \lambda_n(\pm), \mathbf{u}_n^a(\pm) = \mathbf{u}_n(\pm) \text{ evanescent.}\end{aligned}\tag{16.36}$$

Using these definitions, expressions for the advanced Green function matrix are obtained from Eqs. (16.32) and (16.34) by replacing $\mathbf{F}(\pm)$ with $\mathbf{F}^a(\pm)$.

From the general relation between retarded and advanced Green functions, $\mathbf{G}_{i,j} = (\mathbf{G}_{j,i}^a)^\dagger$, the following *row* recursion relations can be deduced for the retarded Green function

$$\begin{aligned}\mathbf{G}_{i,j}^{(0)}(E) &= \mathbf{G}_{i,i}^{(0)}(E) [\mathbf{F}^{a\dagger}(+)]^{j-i}, \quad i < j, \\ \mathbf{G}_{i,j}^{(0)}(E) &= \mathbf{G}_{i,i}^{(0)}(E) [\mathbf{F}^{a\dagger}(-)]^{j-i}, \quad i > j.\end{aligned}\tag{16.37}$$

The retarded Green function $\mathbf{G}_{i,0}^{(s)}(E)$ of a *semi-infinite wire* extending from $i = -\infty, \dots, 0$ can be obtained using a similar technique. Instead of Eq. (16.32), we get

$$\mathbf{G}_{i,0}^{(s)}(E) = \mathbf{F}^i(-) \mathbf{g}(E), \quad i < 0.\tag{16.38}$$

where $\mathbf{g}(E) = \mathbf{G}_{0,0}^{(s)}(E)$ is the surface Green function. Using this in Eq. (16.31) gives for $i = 0$ and $j = 0$ and for $i = -1$ and $j = 0$, respectively,

$$\begin{aligned}[-\mathbf{B}\mathbf{F}^{-1}(-) + E^+\mathbf{I} - \mathbf{H}] \mathbf{g} &= \mathbf{I}, \\ [-\mathbf{B}\mathbf{F}^{-1}(-) + E^+\mathbf{I} - \mathbf{H}] \mathbf{F}^{-1}(-) \mathbf{g} &= \mathbf{B}^\dagger \mathbf{g}.\end{aligned}\tag{16.39}$$

Note that the $\mathbf{B}^\dagger$ term is absent in the first equation since we are dealing with a semi-infinite wire. These two equations can be easily solved to find an expression for the surface Green function

$$\mathbf{g}(E) = \mathbf{F}^{-1}(-) (\mathbf{B}^\dagger)^{-1}.\tag{16.40}$$

Eqs. (16.38) and (16.40) represent the Green function of a semi-infinite ideal wire extending from $i = -\infty, \dots, 0$. In a similar fashion, one gets for the Green function of a semi-infinite ideal wire extending from $i = 0, \dots, \infty$

$$\mathbf{G}_{i,0}^{(s)}(E) = \mathbf{F}^{i+1}(+) \mathbf{B}^{-1}, \quad i \geq 0.\tag{16.41}$$

Analogously to Eqs. (16.38)-(16.41), one can define the advanced Green function matrix $\mathbf{G}_{i,0}^{(s)a}(E)$ in terms of the Bloch matrices $\mathbf{F}^a(\pm)$. Moreover, since $[\mathbf{g}^a]^\dagger = \mathbf{g}$, we have the following relation between the Bloch matrices

$$\mathbf{B}^\dagger \mathbf{F}(\pm) = \mathbf{F}^{a\dagger}(\pm) \mathbf{B}.\tag{16.42}$$

16.6.2 Equivalence of mode matching and Green function approaches

The retarded surface Green functions of the left and right leads can be derived from Eqs. (16.40) and (16.41)

$$\mathbf{g}_L(E) = \mathbf{F}_L^{-1}(-) [\mathbf{B}_L^\dagger]^{-1}; \quad \mathbf{g}_R(E) = \mathbf{F}_R(+)\mathbf{B}_R^{-1}.\tag{16.43}$$

The retarded self-energies of Eq. (16.26) are then given by

$$\Sigma_L(E) = \mathbf{B}_L \mathbf{F}_L^{-1}(-); \Sigma_R(E) = \mathbf{B}_R^\dagger \mathbf{F}_R(+). \quad (16.44)$$

Comparing Eqs. (16.13) and (16.22) then establishes

$$\mathbf{G}'_{i,j}(E) = \mathbf{G}_{i,j}(E). \quad (16.45)$$

In other words, the two Green functions discussed in Secs. 16.4 and 16.5 are identical.

By comparing Eqs. (16.14) and (16.34) one has

$$\mathbf{Q}_0 = [\mathbf{G}_{0,0}^{(0)}(E)]^{-1}. \quad (16.46)$$

In conclusion, the two expressions for the generalized transmission matrix elements, Eqs. (16.20) and (16.30), are identical.

16.6.3 Invariance of transmission probability

Apart from trivial phase factors, the transmission matrix elements should not depend on where in the ideal lead the wave function matching is carried out. In a recent paper it was stated that Ando's expression for $t_{n,m}$, Eqs. (16.20) and (16.16), lacks this invariance property and is therefore incomplete.[22] One can however prove directly from Eq. (16.20) that the transmission matrix elements do have the required invariance property.[39] The proof becomes easier if the equivalence of Eqs. (16.20) and (16.30), established above, is used.

The scattering region runs from slices 0 to $S + 1$ if we include the boundaries with the left and right leads. This means that the Green function matrix $\mathbf{G}_{i,j}$ with indices i, j outside this region obeys the equations for the ideal leads. From the column and row recursion relations, Eqs. (16.32) and (16.37), one derives

$$\mathbf{G}_{S+1+i,j}(E) = \mathbf{F}_R^i(+)\mathbf{G}_{S+1,0}(E) [\mathbf{F}_L^{a\dagger}(-)]^j, \quad (16.47)$$

for $j < 0, i > 0$. In a similar way, one derives for the Green function matrix of the left lead

$$\mathbf{G}_{j,j}^{(0)}(E) = \mathbf{F}_L^j(+)\mathbf{G}_{0,0}^{(0)}(E) [\mathbf{F}_L^{a\dagger}(-)]^j, \quad (16.48)$$

for $j < 0$.

We now artificially extend the scattering region by including slices from the left and right leads and let it run from $j < 0$ to $S + 1 + i$ with $i > 0$. Eq. (16.30) gives for the transmission matrix element

$$\tau'_{n,m} = \tilde{\mathbf{u}}_{R,n}^\dagger(+) \mathbf{G}_{S+1+i,j}(E) [\mathbf{G}_{j,j}^{(0)}(E)]^{-1} \mathbf{u}_{L,m}(+) \quad (16.49)$$

Using (16.47) and (16.48) then gives

$$\begin{aligned} \tau'_{n,m} &= \tilde{\mathbf{u}}_{R,n}^\dagger(+) \mathbf{F}_R^i(+) \mathbf{G}_{S+1,0}(E) \\ &\quad [\mathbf{G}_{0,0}^{(0)}(E)]^{-1} \mathbf{F}_L^{-j}(+) \mathbf{u}_{L,m}(+) \\ &= \lambda_{R,n}^i(+) \lambda_{L,m}^{-j}(+) \tau_{n,m} \\ &= e^{i\alpha} \tau_{n,m}, \end{aligned} \quad (16.50)$$

with α real. The third line follows from applying Eq. (16.7). The last line follows from the fact that we are only interested in propagating states and for propagating states $|\lambda| = 1$. Using this result in (16.16) and (16.17) proves the invariance of the total transmission probability with respect to moving the boundaries between leads and scattering region into the leads.

16.6.4 The Caroli expression

The total transmission probability is given by Eq. (16.17), where the sum has to be over propagating states only. We can extend the summation to include all N states (propagating and evanescent) by defining an $N \times N$ transmission matrix

$$\mathbf{t} = \mathbf{V}_R^{\frac{1}{2}}(+) \tau \tilde{\mathbf{V}}_L^{\frac{1}{2}}(+), \quad (16.51)$$

where τ is the matrix whose elements are given by Eq. (16.20). $\mathbf{V}_R(+)$ is defined as the singular, diagonal matrix that has the velocities $v_{R,n}$ times the constant $\hbar/a_R$ on the diagonal for the right-propagating states and zeros for evanescent states. We call it the velocity matrix. Likewise a pseudo-inverse velocity matrix $\tilde{\mathbf{V}}_L(+)$ can be defined, which has $1/v_{L,n} \times a_L/\hbar$ on the diagonal for left-propagating states and all other matrix elements are zero. These velocity matrices project onto the space of the propagating states so that the transmission matrix has only non-zero values between propagating states. (16.17) can then be expressed in the familiar form

$$\begin{aligned} T &= \text{Tr} [\mathbf{t}^\dagger \mathbf{t}] \\ &= \text{Tr} [\tau^\dagger \mathbf{V}_R(+)\tau \tilde{\mathbf{V}}_L(+)]. \end{aligned} \quad (16.52)$$

Using the above definition of the velocity matrix and Ando's expressions for the transmission matrix elements, we will show how (16.52) can be rewritten as

$$T = \text{Tr} [\mathbf{\Gamma}_R \mathbf{G}^r \mathbf{\Gamma}_L \mathbf{G}^a], \quad (16.53)$$

where $\mathbf{G}^r, \mathbf{G}^a$ are short-hand notations for $\mathbf{G}_{S+1,0}(E)$ and $\mathbf{G}_{0,S+1}^a(E)$, respectively. The matrices $\mathbf{\Gamma}_{L/R}$ are defined as

$$\mathbf{\Gamma}_{L/R} = i [\Sigma_{L/R} - \Sigma_{L/R}^\dagger]. \quad (16.54)$$

Eq. (16.53) is known as the Caroli expression,[15] and it is often used to calculate transmission probabilities. [16, 22] It is equivalent to the Kubo-Greenwood expression for the linear response regime.[14, 40]

The first step is to construct $N \times N$ matrices $\mathbf{U}(\pm)$, the columns of which are the eigenmodes $\mathbf{u}_n(\pm)$, and diagonal matrices $\mathbf{\Lambda}(\pm)$, the elements of which are the eigenvalues $\lambda_n(\pm)$

$$\mathbf{U}(\pm) = (\mathbf{u}_1(\pm) \quad \mathbf{u}_2(\pm) \quad \dots \quad \mathbf{u}_N(\pm)), \quad (16.55)$$

$$\mathbf{\Lambda}(\pm)_{m,n} = \lambda_n(\pm) \delta_{m,n}. \quad (16.56)$$

From Eqs. (16.6) and (16.9) it is then easy to show that the dual vectors $\tilde{\mathbf{u}}_n(\pm)$ form the columns of the matrix $[\mathbf{U}(\pm)^{-1}]^\dagger$ and that the $\mathbf{F}(\pm)$ matrices obey the equation

$$\mathbf{F}(\pm) \mathbf{U}(\pm) = \mathbf{U}(\pm) \mathbf{\Lambda}(\pm). \quad (16.57)$$

In a similar way matrices $\mathbf{U}^a(\pm)$ and $\mathbf{\Lambda}^a(\pm)$ can be constructed, see Eq. (16.36).

Using these definitions, one can generalize the τ -matrix of Eq. (16.20) to

$$\tau = \mathbf{U}_R^{-1}(+) \mathbf{G}^r \mathbf{Q}_0 \mathbf{U}_L(+). \quad (16.58)$$

Note that τ -matrix elements are defined not only between propagating states, but also between evanescent states. However, as we remarked above already, only propagating states contribute to the physical transmission.

The second step is to express the velocity matrices in terms of the $\mathbf{\Gamma}$ -matrices. To do this we use an expression for the velocity matrix,

$$\mathbf{V}(\pm) = i [\mathbf{U}^\dagger(\pm) \mathbf{B}^\dagger \mathbf{U}(\pm) \mathbf{\Lambda}(\pm) - \mathbf{\Lambda}^\dagger(\pm) \mathbf{U}^\dagger(\pm) \mathbf{B} \mathbf{U}(\pm)], \quad (16.59)$$

which can be shown (see Appendix A for a proof) to be equivalent to the definition introduced in the first paragraph of this section. Using (16.57), this can be rewritten for the right lead as

$$\begin{aligned} \mathbf{V}_R(+) &= i \mathbf{U}_R^\dagger(+) [\mathbf{B}_R^\dagger \mathbf{F}_R(+) - \mathbf{F}_R^\dagger(+) \mathbf{B}_R] \mathbf{U}_R(+) \\ &= i \mathbf{U}_R^\dagger(+) [\mathbf{\Sigma}_R - \mathbf{\Sigma}_R^\dagger] \mathbf{U}_R(+) \\ &= \mathbf{U}_R^\dagger(+) \mathbf{\Gamma}_R \mathbf{U}_R(+). \end{aligned} \quad (16.60)$$

The second line follows from (16.44). A similar relation between the $\mathbf{\Gamma}$ -matrix and the velocity matrix for the left lead can be shown to exist,

$$\mathbf{V}_L(+) = \mathbf{U}_L^{a\dagger}(-) \mathbf{\Gamma}_L \mathbf{U}_L^a(-), \quad (16.61)$$

by using (16.42) and an equivalent expression for the velocity matrix, Eq. (16.93). Eqs. (16.60) and (16.61) imply that the $\mathbf{\Gamma}$ -matrices project onto the space spanned by the propagating states.

The third step is to introduce a matrix $\mathbf{P}$ that explicitly projects onto the propagating states of the left lead

$$\begin{aligned} \mathbf{P} &= \mathbf{U}_L(+) \mathbf{I}_p [\mathbf{U}_L^a(-)]^{-1} \\ &= \sum_{n=1}^{N_p} \mathbf{u}_{L,n}(+) \tilde{\mathbf{u}}_{L,n}^{a\dagger}(-), \end{aligned} \quad (16.62)$$

where the $\mathbf{I}_p$ -matrix contains 1 on the N_p diagonal elements that correspond to propagating states, and 0 at all other positions. Given this projector matrix, it is possible to prove that

$$\mathbf{Q}_0 \mathbf{P} = i \mathbf{\Gamma}_L. \quad (16.63)$$

The proof is given in Appendix B. Using this property one has

$$\begin{aligned} \mathbf{Q}_0 \mathbf{U}_L(+) \tilde{\mathbf{V}}_L(+) &= \mathbf{Q}_0 \mathbf{P} \mathbf{U}_L^a(-) \tilde{\mathbf{V}}_L(+) \\ &= i \mathbf{\Gamma}_L \mathbf{U}_L^a(-) \tilde{\mathbf{V}}_L(+) = i [\mathbf{U}_L^{a\dagger}(-)]^{-1} \mathbf{V}_L(+) \tilde{\mathbf{V}}_L(+) \\ &= i [\mathbf{U}_L^{a\dagger}(-)]^{-1} \mathbf{I}_p. \end{aligned} \quad (16.64)$$

Substituting (16.58), (16.60) and (16.64) into (16.52) leads directly to the Caroli expression, Eq. (16.53).

16.6.5 Transmission matrix: a compact expression

Using the results of the previous section it is possible to derive a more compact expression for the transmission matrix. Combining (16.51), (16.58) and (16.62) one has

$$\mathbf{t} = \mathbf{V}_R^{\frac{1}{2}}(+) \mathbf{U}_R^{-1}(+) \mathbf{G}^r \mathbf{Q}_0 \mathbf{P} \mathbf{U}_L^a(-) \tilde{\mathbf{V}}_L^{\frac{1}{2}}(+). \quad (16.65)$$

By following the same steps as in Eq. (16.64) and using $\mathbf{V}_L \tilde{\mathbf{V}}_L^{\frac{1}{2}} = \mathbf{V}_L^{\frac{1}{2}}$ this can be simplified to

$$\mathbf{t} = i \mathbf{V}_R^{\frac{1}{2}}(+) \mathbf{U}_R^{-1}(+) \mathbf{G}^r [\mathbf{U}_L^{a\dagger}(-)]^{-1} \mathbf{V}_L^{\frac{1}{2}}(+). \quad (16.66)$$

Writing out the transmission matrix elements gives the compact expression

$$t_{n,m} = i \hbar \sqrt{\frac{v_{R,n} v_{L,m}}{a_R a_L}} \tilde{\mathbf{u}}_{R,n}^\dagger(+) \mathbf{G}_{S+1,0}(E) \tilde{\mathbf{u}}_{L,m}^a(-). \quad (16.67)$$

This is the tight-binding equivalent of the Fisher-Lee expression relating transmission and Green function matrices.[37]

16.7 Examples

16.7.1 A simple analytical model

We consider a system consisting of a single impurity in a one-dimensional chain and treat this within a one-band nearest neighbor tight-binding model. The parameters of this model are given in Fig. 82. This model can be solved analytically,[23] so it can serve as a simple test to illustrate the equivalence of the different approaches.

We will first solve the problem using the mode matching approach of Sec. 16.4. It is convenient to define a scaled energy by

$$\omega \equiv \frac{E - \epsilon}{2\beta}. \quad (16.68)$$

The model involves only one channel and with $\mathbf{u}_m(\pm) = 1$ Eq. (16.4) reduces to

$$-\beta + (E - \epsilon)\lambda(\pm) - \beta\lambda(\pm)^2 = 0, \quad (16.69)$$

The roots $\lambda(\pm)$ can be given a more familiar form. For $|\omega| \leq 1$ we define a wave number k by

$$\cos(ka) = \omega, \quad (16.70)$$

where a is the lattice parameter. From Eqs. (16.69), (16.70) one then obtains

$$\lambda(\pm) = e^{\pm ika}, \quad (16.71)$$

which describes propagating states. For $|\omega| > 1$ one defines κ by

$$\cosh(\kappa a) = |\omega|. \quad (16.72)$$

One obtains $\lambda(\pm) = \exp(\mp \kappa a)$ if $\omega > 1$ and $\lambda(\pm) = -\exp(\mp \kappa a)$ if $\omega < -1$; both cases describe evanescent states.

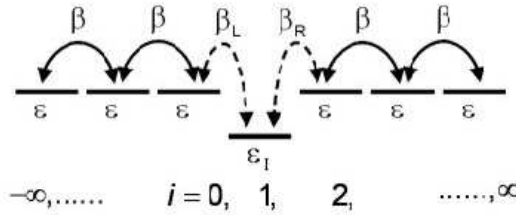


Figure 82: The parameters of a one-band nearest neighbor tight-binding model for a single impurity in a onedimensional chain.

Since the scattering region consists of a single impurity, $S = 1$, Eqs. (16.12)-(16.14) give three linear equations with three unknowns describing the scattering problem. In a one-channel model one has $\mathbf{F}(\pm) = \lambda(\pm)$ and $\lambda(\pm)^{-1} = \lambda(\mp)$. There is only one possible incoming wave, so $\mathbf{c}_0(+) = 1$. The linear equations then become in matrix form

$$\mathbf{A} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} \beta[\lambda(-) - \lambda(+)] \\ 0 \\ 0 \end{pmatrix}, \quad (16.73)$$

with

$$\mathbf{A} = \begin{pmatrix} E - \epsilon - \beta\lambda(+) & -\beta_L & 0 \\ -\beta_L & E - \epsilon_1 & -\beta_R \\ 0 & -\beta_R & E - \epsilon - \beta\lambda(+) \end{pmatrix}. \quad (16.74)$$

Solving this set of equations and using Eqs. (16.70) and (16.71) we obtain the compact expression

$$c_2 = e^{2ika} \frac{-if \sin(ka)}{d + (1-b) \cos(ka) - ib \sin(ka)}, \quad (16.75)$$

defining the dimensionless parameters

$$b = \frac{\beta_L^2 + \beta_R^2}{2\beta^2}, \quad d = \frac{\epsilon - \epsilon_I}{2\beta}, \quad f = \frac{\beta_L \beta_R}{\beta^2}. \quad (16.76)$$

Applying Eqs. (16.15)-(16.17) then yields for the total transmission probability

$$T(E) = |c_2|^2. \quad (16.77)$$

Using Eqs. (16.70) and (16.75) it is easy to show that this transmission probability is identical to Eq. (15) of Ref. [23], which was obtained using a different technique.

It is instructive to solve the same problem using the Green function approach of Sec. 16.5.1. First one has to find the surface Green functions of the leads from Eqs. (16.24) and (16.25), which for the current model become

$$[E - \epsilon - \beta^2 g(E)] g(E) = 1, \quad (16.78)$$

where E is a real energy. This equation has the solutions

$$g^\pm(E) = \frac{e^{\pm ika}}{\beta}, \quad (16.79)$$

for both leads. The Green function matrix in the scattering region can then be found from Eqs. (16.21) and (16.22), which can be combined in the 3×3 matrix equation

$$\mathbf{A} \mathbf{G}(E) = \mathbf{I}, \quad (16.80)$$

where $G_{i,j}(E)$, $i, j = 0, \dots, 2$ are the matrix elements of $\mathbf{G}(E)$ and $\mathbf{A}$ is given by Eq. (16.74). Inverting $\mathbf{A}$ yields the matrix element

$$G_{2,0}(E) = \frac{f}{2\beta} \frac{e^{2ika}}{d + (1-b) \cos(ka) - ib \sin(ka)}, \quad (16.81)$$

with the parameters b , d and f defined by Eq. (16.76). The relevant Green function matrix element for the ideal lead is found from Eq. (16.34)

$$G_{0,0}^{(0)}(E) = \frac{i}{2\beta \sin(ka)}. \quad (16.82)$$

Using these results in Eq. (16.30) one observes that the expression for the (one channel) transmission matrix element becomes identical to Eq. (16.75).

Finally one can calculate the transmission probability from the Caroli expression given in Sec. 16.6.4, cf. Eq. (16.53). Using Eqs. (16.26), (16.54) and (16.79) one obtains

$$\Gamma_L = \Gamma_R = -2\beta \sin(ka), \quad (16.83)$$

for left and right leads. Using Eqs. (16.81), (16.83) and $G_{0,2}^a = (G_{2,0})^*$ in Eq. (16.53) then yields an expression for the transmission probability that is identical to Eq. (16.77). It illustrates the equivalence of the different approaches for calculating the transmission in this simple model.

In addition to providing a channel for propagating states, an impurity can also give rise to localized states, whose energy is outside the energy band of the chain, cf. Eq. (16.72).

Such a state does not contribute to the physical transmission, but the transmission amplitude has a pole at an energy that corresponds to a localized state.[41] Within the mode matching approach this corresponds to an energy at which c_{S+1} becomes infinite. For the present model the energies of localized states can be obtained by setting $k = i\kappa$ and setting the denominator to zero in Eq. (16.75). This leads to the equation

$$(\omega + d) \left(\omega + \text{sgn}(\omega) \sqrt{\omega^2 - 1} \right) - b = 0; \quad |\omega| > 1, \quad (16.84)$$

the roots of which give the energies of the localized states. Again these results are equivalent to the results obtained using the approach of Ref. [23].

Within the Green function approach the energies of the localized states are given by the poles of the Green function matrix. Via Eq. (16.81) this again leads to Eq. (16.84). Alternatively, since the Green function matrix is the inverse of the $\mathbf{A}$ -matrix, cf. Eq. (16.80), its poles are given by the roots of $\det(\mathbf{A}) = 0$. This equation is equivalent to Eq. (16.84), as is easily shown by setting $\lambda(+) = \pm \exp(-\kappa a)$ in the $\mathbf{A}$ -matrix.

16.7.2 Fe|vacuum|Fe tunnel junction

As an example of a more complex system, we consider an Fe|vacuum|Fe tunnel junction where the electronic structure is treated using the local density approximation of DFT.[42] The calculations are based upon a tight-binding muffin tin orbital (TB-MTO) atomic spheres approximation (ASA) implementation [8, 19, 20, 21, 25] of the formalism described in Sect. 16.4.

The first step in the calculation is the self-consistent determination of the electronic structure of the tunnel junction using the layer Green function approach of Ref. [35]. The Fe leads are oriented in the (001) direction and the atoms at the Fe(001) surfaces are kept at their unrelaxed bulk positions. For the bcc structure and TB-MTOs,[43] a principal layer in the (001) direction contains two monolayers of Fe with a thickness of 2.866 Å. The vacuum region is modeled by a number of such slices, of the same thickness, filled with “empty” spheres of the same size and packing as the Fe atomic spheres. The atomic sphere potentials of the vacuum region and four monolayers (two principal layers) of Fe on either side of the vacuum are calculated self-consistently while the potentials of more distant layers are kept at their bulk values. These potentials then form the input to a transmission calculation based on mode matching. [8, 20, 21] Further technical details can be found in Ref. [25].

A useful quantity to extract from the self-consistent layer calculation is the layer density of states (LDOS) $\rho_i(E)$. It is related to the retarded Green function matrix defined in Eq. (16.21) by

$$\rho_i(E + i\eta) = -\pi^{-1} \text{ImTr}[\mathbf{G}_{i,i}(E + i\eta)] \quad (16.85)$$

where the trace refers to the usual lm angular momentum indices characterizing MTOs. For reasons of numerical stability the retarded Green function matrix is calculated by adding a finite imaginary part to the energy; we have used $\eta = 0.0025$ Ry. In the mode matching approach, the LDOS can be directly expressed in terms of the wave functions or, alternatively, in terms of the Green function matrix of Eq. (16.18)

$$\rho_i(E) = -\pi^{-1} \text{ImTr}[\mathbf{G}'_{i,i}(E)]. \quad (16.86)$$

This Green function matrix can be calculated for a real energy, but in order to make a comparison to the results obtained with Eq. (16.85), we add an imaginary part, $\eta = 0.0025$ Ry.

In Fig. 83 we compare the LDOS obtained using (16.85) and (16.86) for the topmost Fe monolayer of (001) Fe|vacuum|Fe where the vacuum layer was so thick (four principal layers, corresponding to a thickness of 11.466 Å) that the LDOS corresponds closely to that of a free Fe(001) surface. The two curves, displaced vertically for clarity, are indistinguishable,

illustrating the equivalence of the two Green functions defined in (16.18) and (16.21). Moreover, the LDOS are in essential agreement with results found previously for a Fe(001) surface.[44]

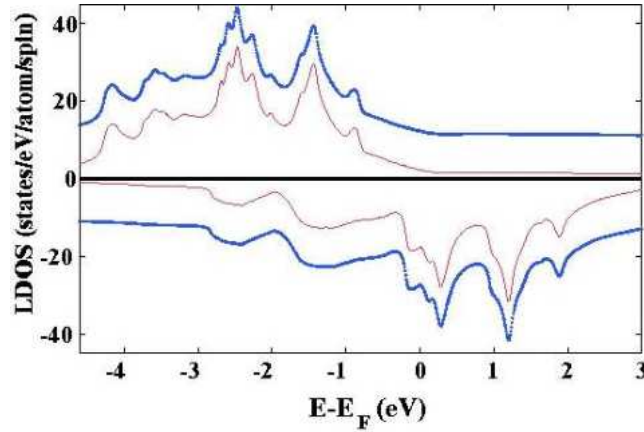


Figure 83: Layer density of states (LDOS) at the Fe(001) surface layer. Top panel: majority spin. Top curve (red, dashed): as calculated using the layer Green function method.[35] Bottom curve (blue, solid): as calculated using the mode matching approach. Both calculations use $\eta = 0.0025$ Ry. For clarity, the top curve has been displaced by 10 units along the y -axis. Bottom panel: minority spin, the LDOS is shown with a negative sign.

If one resolves the LDOS into contributions from different parts of the surface Brillouin zone, then the contribution at $\bar{\Gamma}$ ($\mathbf{k}_{\parallel} = 0$) exhibits sharp peaks near the Fermi level. These are associated with characteristic surface states found on (001) surfaces of bcc transition metals.[45] These surface states are mainly derived from $d_{3z^2-r^2}$ orbitals on the surface atoms projecting into the vacuum, and they have Δ_1 symmetry. [46, 47] They become most clearly visible if one filters out the contribution to the LDOS at $\bar{\Gamma}$ of the states with Δ_1 symmetry.

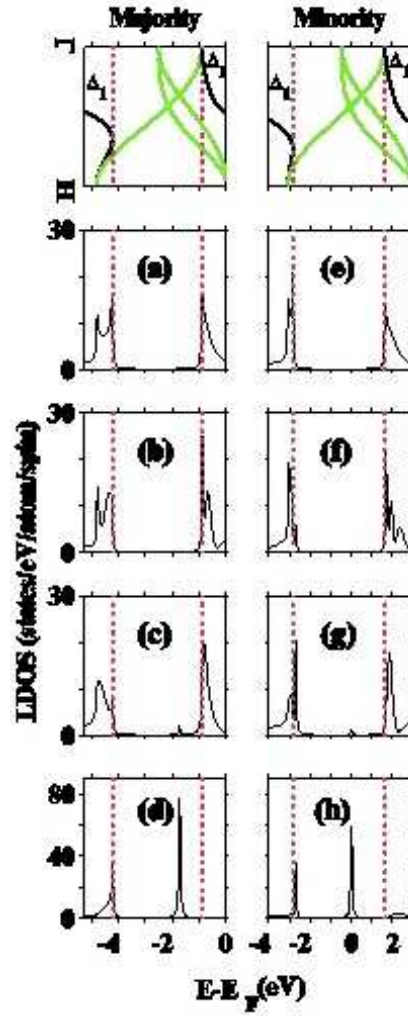


Figure 84: The top panels represent the band structures along Γ -H of bulk Fe for majority and minority spins. Panels (a-h) give the Δ_1 contribution to the LDOS $\rho_i(E)$ at $\bar{\Gamma}$ of the Fe|vacuum interface as a function of the layer position. (a) LDOS of an Fe layer at an infinite distance from the interface; the double peak structure below -4 eV and single peak (of a double peak structure) above -1 eV correspond to the Δ_1 bulk bands. (b) LDOS of the 6th Fe layer (the surface layer is layer 1). (c) LDOS of the 3rd Fe layer; the sharp peak in the bulk band gap corresponds to a surface state. (d) LDOS of the the surface Fe layer; the surface state peak has maximum amplitude. (e-h) Same for minority spin.

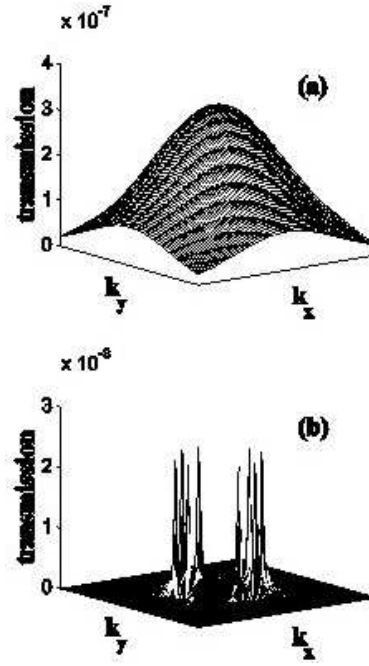


Figure 85: Calculated $\mathbf{k}_{\parallel}$ -resolved transmission in (a) the majority and (b) the minority spin channels through a Fe-vacuum-Fe tunnel junction. The vacuum region has a thickness of 5.733 Å and the Fe electrodes have a parallel magnetization. Only the central part of the Brillouin zone is shown, i.e. $-\pi/3a \leq k_x, k_y \leq \pi/3a$; the transmission in the other parts is close to zero.

This contribution, calculated using the mode matching approach, is shown in Fig. 84 as a function of the distance from the surface. In the majority spin LDOS, a sharp peak is found at $E_F - 1.8$ eV, which is very prominent in the surface Fe layer, cf. Fig. 84(d) and decays rapidly into the bulk, cf. Figs. 84(c)-(a). This evanescent state clearly represents a state that is localized at the surface. In the minority spin LDOS, a sharp peak with similar properties is found very close to E_F . It corresponds to the surface state that is observed experimentally by STM.[45] The positions of both these surface state peaks in the majority and minority LDOS are in excellent agreement with those obtained using a Green function (KKR) approach.[46, 47] Such surface states are good examples of evanescent states. Clearly, they are described properly within the mode matching approach which was, however, disputed in Ref. [22].

Using the mode matching approach, we next calculate the transmission through an Fe | vacuum | Fe tunnel junction in which the vacuum region is only 5.733 Å thick (corresponding to two principal layers) and the magnetizations are parallel. Fig. 85 shows the calculated $\mathbf{k}_{\parallel}$ -resolved transmission for majority- and minority-spin channels. The majority conductance has a maximum at $\bar{\Gamma}$ and it decreases smoothly going away from $\bar{\Gamma}$. This behavior is quite general for scattering from a potential barrier. At fixed energy, a wave travelling perpendicular to the vacuum barrier penetrates furthest into the vacuum and therefore has a maximum tunneling probability.[48]

In contrast, the minority conductance is dominated by sharp spikes of very high intensity, close to $0.13 \bar{\Gamma}X$. This behavior has been analyzed in terms of interface resonant states that extend relatively far into the vacuum. States originating from the two Fe|vacuum interfaces couple through the thin vacuum region, which enhances the transmission considerably.[48] These transmission spikes or “hot spots” are observed quite generally in calculations for perfect transition metal|insulator|transition metal tunnel junctions.[49, 50] A calculation using Eq. (16.30) gives the same transmission within numerical accuracy. It again illustrates the correct treatment of evanescent states within the mode matching approach. We have also tested expression (16.67)

numerically and found that it too gives the same transmission within numerical accuracy.

Finally, the invariance property of the transmission discussed in Sec. 16.6.3, has been tested numerically by moving the interfaces between the leads and the scattering region, where the matching is carried out, into the leads. In this way more and more slices of lead are treated as scattering region, see Fig. 81; the transmission should be invariant under this operation. Adding up to ten principal layers (20 monolayers) of bulk Fe to each side of the scattering region changes the transmission negligibly, by less than one part in 10^8 .

16.8 Velocity matrix

In this appendix we will prove that

$$\mathbf{V}(\pm)_{m,n} = \frac{\hbar}{a} v_m(\pm) \delta_{m,n}, \quad (16.87)$$

for the expression introduced for the velocity matrices in Sec. 16.6.4, Eq. (16.59). Here, a is the translation period along the wire and, for right- respectively left-propagating states, $v_n(\pm)$ is the Bloch velocity in the direction of the quantum wire. For evanescent states $v_n(\pm) = 0$.

For ease of notation we drop the index $\pm$ in the following. From Eq. (16.4) and its complex conjugate one has

$$\begin{aligned} -\mathbf{u}_m^\dagger \mathbf{B} \mathbf{u}_n + \lambda_n \mathbf{u}_m^\dagger (\mathbf{E} \mathbf{I} - \mathbf{H}) \mathbf{u}_n - (\lambda_n)^2 \mathbf{u}_m^\dagger \mathbf{B}^\dagger \mathbf{u}_n &= 0, \\ -\mathbf{u}_m^\dagger \mathbf{B}^\dagger \mathbf{u}_n + \lambda_m^* \mathbf{u}_m^\dagger (\mathbf{E} \mathbf{I} - \mathbf{H}) \mathbf{u}_n - (\lambda_m^*)^2 \mathbf{u}_m^\dagger \mathbf{B} \mathbf{u}_n &= 0. \end{aligned} \quad (16.88)$$

Multiplying the first equation by λ_m^* , the second by λ_n and subtracting the two gives

$$\begin{aligned} [\lambda_n \mathbf{u}_m^\dagger \mathbf{B}^\dagger \mathbf{u}_n - \lambda_m^* \mathbf{u}_m^\dagger \mathbf{B} \mathbf{u}_n] (1 - \lambda_m^* \lambda_n) &= 0 \Leftrightarrow \\ -i \mathbf{V}_{m,n} (1 - \lambda_m^* \lambda_n) &= 0, \end{aligned} \quad (16.89)$$

according to Eq. (16.59). So if $\lambda_m^* \lambda_n \neq 1$ then $\mathbf{V}_{m,n} = 0$.

The velocity matrices contain, by construction, either right- or left-going modes. For right-going evanescent modes one has $|\lambda| < 1$, for left-going evanescent modes $|\lambda| > 1$, so $\lambda_m^* \lambda_n \neq 1$ if n and/or m denotes an evanescent mode; thus $\mathbf{V}_{m,n} = 0$. For propagating states, $|\lambda_m| = 1$ and one can write $\lambda_m = \exp(ik_m a)$. It is obvious that if $k_m \neq k_n$, then $\lambda_m^* \lambda_n \neq 1$ and again $\mathbf{V}_{m,n} = 0$.

This argument does not hold for the diagonal matrix elements of propagating states, and also not for degenerate propagating states, i.e. if $\lambda_m = \lambda_n$. In the latter case we take the derivative d/dE of the first line of Eq. (16.88). The coefficients of the terms in $d\mathbf{u}_n/dE$ and $d\mathbf{u}_m^\dagger/dE$ vanish because $\mathbf{u}_n$ and $\mathbf{u}_m^\dagger$ satisfy the eigenvalue equation Eq. (16.4) and its complex conjugate, respectively, and $\lambda_n^* = 1/\lambda_n$ for propagating states. Collecting the remaining terms gives

$$\begin{aligned} -\frac{d\lambda_n}{dE} [\lambda_n \mathbf{u}_m^\dagger \mathbf{B}^\dagger \mathbf{u}_n - \lambda_n^* \mathbf{u}_m^\dagger \mathbf{B} \mathbf{u}_n] + \lambda_n \mathbf{u}_m^\dagger \mathbf{u}_n &= 0 \Leftrightarrow \\ i \frac{d\lambda_n}{dE} \mathbf{V}_{m,n} + \lambda_n \mathbf{u}_m^\dagger \mathbf{u}_n &= 0, \end{aligned} \quad (16.90)$$

where we emphasize that for $m \neq n$ this only holds for degenerate propagating states. Without loss of generality, degenerate states can be chosen orthogonal, i.e. $\mathbf{u}_m^\dagger \mathbf{u}_n = 0$ for $m \neq n$. Since $d\lambda_n/dE \neq 0$, it then follows that $\mathbf{V}_{m,n} = 0$ if $m \neq n$.

The diagonal matrix elements $m = n$ for propagating states are also given by Eq. (16.90). Setting $n = m$ and using the fact that the states are normalized, i.e. $\mathbf{u}_n^\dagger \mathbf{u}_n = 1$ gives the expression

$$\mathbf{V}_{n,n} = i\lambda_n \frac{dE}{d\lambda_n}. \quad (16.91)$$

Since for propagating states we can write $\lambda_n = \exp(ika)$ and the Bloch velocity is defined as $v_n = \hbar^{-1} dE/dk$, this then proves both Eq. (16.87) and Eq. (16.5).

It is straightforward to derive an alternative expression for the velocity matrices in terms of the advanced matrices

$$\mathbf{V}(\pm) = i \left[\mathbf{U}^{a\dagger}(\mp) \mathbf{B}^\dagger \mathbf{U}^a(\mp) \mathbf{\Lambda}^a(\mp) - \mathbf{\Lambda}^{a\dagger}(\mp) \mathbf{U}^{a\dagger}(\mp) \mathbf{B} \mathbf{U}^a(\mp) \right]. \quad (16.92)$$

Multiplying from the left by $[\mathbf{\Lambda}^{a\dagger}]^{-1}$ and from the right by $[\mathbf{\Lambda}^a]^{-1}$ this expression is seen to be equivalent to

$$\mathbf{V}(\pm) = i \left[\{\mathbf{\Lambda}^{a\dagger}(\mp)\}^{-1} \mathbf{U}^{a\dagger}(\mp) \mathbf{B}^\dagger \mathbf{U}^a(\mp) - \mathbf{U}^{a\dagger}(\mp) \mathbf{B} \mathbf{U}^a(\mp) \{\mathbf{\Lambda}^a(\mp)\}^{-1} \right]. \quad (16.93)$$

16.9 Projector matrix

In order to prove Eq. (16.63), we start from Eq. (16.14) and rewrite it using Eq. (16.42) as

$$\mathbf{Q}_0 = \mathbf{B}_L \mathbf{F}_L^{-1}(+) - [\mathbf{F}_L^{a\dagger}(-)]^{-1} \mathbf{B}_L^\dagger. \quad (16.94)$$

In the following we will drop the subscript L for ease of notation. Multiplying (16.94) on the left with the identity operator $\mathbf{I} = \sum \mathbf{u}_n(+) \tilde{\mathbf{u}}_n^\dagger(+)$ and on the right with $\mathbf{I} = \sum \tilde{\mathbf{u}}_m^a(-) \mathbf{u}_m^{a\dagger}(-)$ yields

$$\begin{aligned} \mathbf{Q}_0 &= \sum_{m,n=1}^N \tilde{\mathbf{u}}_m^a(-) \tilde{\mathbf{u}}_n^\dagger(+) \\ &\quad \left[\frac{1}{\lambda_n(+)} \mathbf{u}_m^{a\dagger}(-) \mathbf{B} \mathbf{u}_n(+) - \frac{1}{\lambda_m^{a*}(-)} \mathbf{u}_m^{a\dagger}(-) \mathbf{B}^\dagger \mathbf{u}_n(+) \right], \end{aligned} \quad (16.95)$$

where use has been made of Eqs. (16.6), (16.7) and (16.36).

By similar arguments as those leading to Eq. (16.89) it follows that $[\dots] = 0$, unless $\lambda_m^a(-) = 1/\lambda_n^*(+)$. From Eq. (16.36) it follows that this is true if $n = m$ and n denotes a propagating mode. In addition, for every (+) evanescent mode n with eigenvalue $\lambda_n(+)$, there is one (−) evanescent mode with eigenvalue $\lambda_n(-) = 1/\lambda_n^*(+)$. [31] Again it follows from Eq. (16.36) that only the $m = n$ terms are non-zero for evanescent modes in Eq. (16.95). This equation then simplifies to

$$\begin{aligned} \mathbf{Q}_0 &= \sum_{n=1}^{N_p} \tilde{\mathbf{u}}_n^a(-) \tilde{\mathbf{u}}_n^\dagger(+) \\ &\quad \left[\frac{1}{\lambda_n(+)} \mathbf{u}_n^\dagger(+) \mathbf{B} \mathbf{u}_n(+) - \lambda_n(+) \mathbf{u}_n^\dagger(+) \mathbf{B}^\dagger \mathbf{u}_n(+) \right] \\ &\quad + \sum_{n=N_p+1}^N \tilde{\mathbf{u}}_n^a(-) \tilde{\mathbf{u}}_n^\dagger(+) \\ &\quad \left[\frac{1}{\lambda_n(+)} \mathbf{u}_n^\dagger(-) \mathbf{B} \mathbf{u}_n(+) - \lambda_n(+) \mathbf{u}_n^\dagger(-) \mathbf{B}^\dagger \mathbf{u}_n(+) \right], \end{aligned} \quad (16.96)$$

where the first summation is over the N_p propagating modes and the second summation is over the $N - N_p$ evanescent modes.

The projector matrix $\mathbf{P}$ is defined by Eq. (16.62). In the product $\mathbf{Q}_0\mathbf{P}$ only the first summation in Eq. (16.96) survives. Moreover, since $[\dots] = i\mathbf{V}_{n,n}$ according to Eq. (16.89), we have

$$\mathbf{Q}_0\mathbf{P} = i \sum_{n=1}^{N_p} \tilde{\mathbf{u}}_n^a(-) \tilde{\mathbf{u}}_n^{a\dagger}(-) \mathbf{V}_{n,n}(+). \quad (16.97)$$

Together with Eq. (16.61) this then proves Eq. (16.63).

16.10 Easy illustrative examples

16.10.1 Example 1

16.10.2 Example 2

17 Thermal transport in one-dimensional spin heterostructures

Liliana Arrachea¹, Gustavo S. Lozano¹ and A. A. Aligia²

Abstract

We study heat transport in a one-dimensional inhomogeneous quantum spin 1/2 system. It consists of a finite-size XX spin chain coupled at its ends to semi-infinite XX and XY chains at different temperatures, which play the role of heat and spin reservoirs. After using the Jordan-Wigner transformation we map the original spin Hamiltonian into a fermionic Hamiltonian, which contains normal and pairing terms. We find the expressions for the heat currents and solve the problem with a non-equilibrium Green's function formalism. We analyze the behavior of the heat currents as functions of the model parameters. When finite magnetic fields are applied at the two reservoirs, the system exhibits rectifying effects in the heat flow.

17.1 Introduction

In the last decade there has been a renewed interest related to the research of thermal transport in one-dimensional magnetic systems [1, 2]. Most of these studies have been motivated by unusual high values of thermal conductance of some materials, as for example reported in Ref. [3]. From the theoretical side, several works calculating and discussing thermal conductivity in different model Hamiltonians have also appeared [4, 5]. On the other hand, spin Hamiltonians provide the natural scenario to implement quantum computation. This motivated interesting proposals of a variety of physical systems, like arrays of quantum dots, optical lattices and nuclear magnetic resonance experiments which are architected to effectively behave like one-dimensional spin systems. [6]

Most of the theoretical studies on heat transport in spin systems are performed within the linear response regime, assuming a very small temperature gradient ∇T , and using a Kubo formula. Although this formula is widely used in calculating thermal transport properties it has several conceptual difficulties, particularly if ∇T is not small [7, 2]. The thermal averaging assumes a constant temperature, and (as in all Kubo formulas), the time dependence is governed by the total Hamiltonian $H = H_0 + H'$, including the perturbation $H' \sim \nabla T$, while the thermodynamic averaging is done with H_0 , assuming a fast evolution. For finite ∇T the

separation of H into an unperturbed part H_0 and a perturbation H' is in principle ambiguous. On the other hand, real experiments are performed by coupling the system under study to macroscopic systems with well defined temperatures that act as reservoirs. Therefore, it seems desirable to develop alternative approaches designed to treat systems out of equilibrium. Important progress has been achieved by studying the properties of non-equilibrium steady states of XX and XY chains within an algebraic setting which allows one to obtain explicit analytical expressions for different quantities such as entropy production [8]. Recently, the usefulness of Kubo formula for the investigation of heat transport in quantum systems has been discussed [9] and the investigation of energy transport through spin systems beyond Kubo formula has been addressed on the basis of master equations and quantum Monte Carlo methods. [10] Nevertheless, it is not clear how to extend these ideas to more general situations.

In this work we address the problem of thermal transport in one dimensional magnetic systems from a different perspective. We study a problem that to the best of our knowledge has not been considered yet. We study the heat flow through a *spin chain heterostructure*, which we generically define as a set of finite or semi-infinite spin chains attached by their ends. Each piece of the heterostructure can be in principle described by a *different* Hamiltonian. For definiteness, we shall consider in this work a finite central system connected to two (left and right) semi-infinite chains, with a *finite* temperature difference between them [11, 10]. This problem is the thermal analog of electronic transport through mesoscopic structures, nanodevices or molecules connected to conducting leads with a finite applied bias voltage, a subject of intense research in recent years. In fact, this type of set-up is the common situation found in the study of charge transport in electronic systems, where a central system is connected to charge reservoirs that also act as thermal baths. This is the basis of the "Landauer" approach, which is one of the most common frameworks to study transport properties of meso- and nano- devices in the last years. [12]

In addition, as it is well known, one-dimensional spin 1/2 systems can be mapped, via the Jordan-Wigner transformation to fermionic systems. Thus, the model under investigation is equivalent to an electronic heterostructure where very well established techniques, as the Schwinger-Keldysh non equilibrium Green's functions method can be applied [13].

We focus on a simple device where the central and right parts of the system can be described by XX spin 1/2 chains, and the left part corresponds to an anisotropic XY chain. The Jordan-Wigner transformation maps these models into bilinear fermionic systems, rendering the theoretical study simpler. We show that this simple device presents an interesting physical effect: due to a mechanism reminiscent of Andreev reflections in superconductors, this device could act as thermal diode. This rectifying effect might be useful for applications. We argue that this is a generic feature, which remains valid for anisotropic XYZ models.

The paper is organized as follows: in Section II we present the model, and the non-equilibrium formalism based on Green's functions. Section III contains the numerical results. Section IV is a summary and discussion.

17.2 The formalism

17.2.1 Model

We consider a system of three one-dimensional spin 1/2 chains coupled through their ends as shown in Fig. 86. The system is described by the following Hamiltonian $H = H_L + H_C +$

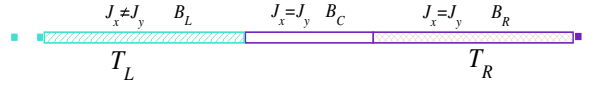


Figure 86: (Color online) Schematic digram showing the set up of the system. The right and central chains correspond to XX models while the left chain is described by an XY model. Left and right chains are taken semi-infinite at temperatures T_L and T_R respectively.

$H_R + H_{coup}$, where

$$\begin{aligned}
 H_\alpha &= J_\alpha \sum_{i=1}^{N_\alpha-1} [S_{\alpha,i}^x S_{\alpha,i+1}^x + \lambda_\alpha S_{\alpha,i}^y S_{\alpha,i+1}^y] \\
 &\quad - B_\alpha \sum_{i=1}^{N_\alpha} S_{\alpha,i}^z, \\
 H_{coup} &= J' [S_{L,1}^x S_{C,1}^x + S_{L,1}^y S_{C,1}^y] + \\
 &\quad J' [S_{R,1}^x S_{C,N_C}^x + S_{R,1}^y S_{C,N_C}^y],
 \end{aligned} \tag{17.1}$$

where in the first line, the index $\alpha = L, C, R$ (left, center, right) labels the different chains. Each chain is characterized by its nearest-neighbor exchange J_α , with anisotropy λ_α for the ratio of the interaction along the y direction with respect to that along x , and the magnetic field B_α applied along the z direction. The isotropic case for a given chain corresponds to $\lambda_\alpha = 1$. The Hamiltonian H_{coup} describes the coupling between the central chain, and the left and right ones. In terms of the representation for the spin operators: $S^\pm = S^x \pm iS^y$, the above Hamiltonian reads

$$\begin{aligned}
 H_\alpha &= \frac{J_\alpha}{4} \sum_{l=1}^{N_\alpha-1} [(1 - \lambda_\alpha)(S_{\alpha,l}^+ S_{\alpha,l+1}^+ + S_{\alpha,l}^- S_{\alpha,l+1}^-) \\
 &\quad + (1 + \lambda_\alpha)(S_{\alpha,l}^+ S_{\alpha,l+1}^- + S_{\alpha,l}^- S_{\alpha,l+1}^+)] \\
 &\quad - B_\alpha \sum_{l=1}^{N_\alpha} S_{\alpha,l}^z, \\
 H_{coup} &= \frac{J'}{2} [S_{L,1}^+ S_{C,1}^- + S_{L,1}^- S_{C,1}^+] + \\
 &\quad \frac{J'}{2} [S_{R,1}^- S_{C,N_C}^+ + S_{R,1}^+ S_{C,N_C}^-].
 \end{aligned} \tag{17.2}$$

For the isotropic case, $\lambda_\alpha = 1$ only the flip-flop terms with products of one raising and one lowering spin operators survive.

We now introduce the Jordan-Wigner transformation to map the spin 1/2 Hamiltonian into a fermionic Hamiltonian through: $S_{\alpha,l}^+ = f_{\alpha,l}^\dagger \exp(i\pi \sum_j' f_{\alpha,j}^\dagger f_{\alpha,j})$, where $\sum_j'$ denotes a sum over all the positions located at the left of the position α, l . Similarly, the other spin operators transform as $S_{\alpha,l}^- = \exp(-i\pi \sum_j' f_{\alpha,j}^\dagger f_{\alpha,j}) f_{\alpha,l}$ and $S_{\alpha,l}^z = f_{\alpha,l}^\dagger f_{\alpha,l} - 1/2$, where the operators $f_l, f_l^\dagger$

obey fermionic commutation rules: $\{f_{\alpha,l}, f_{\alpha',l'}^\dagger\} = \delta_{l,l'}\delta_{\alpha,\alpha'}$, and $\{f_{\alpha,l}^\dagger, f_{\alpha',l'}^\dagger\} = \{f_{\alpha,l}, f_{\alpha',l'}\} = 0$. Substituting in the Hamiltonian (17.2), we get

$$\begin{aligned}
H_\alpha &= \sum_{l=1}^{N_\alpha-1} \{w_\alpha [f_{\alpha,l}^\dagger f_{\alpha,l+1} + f_{\alpha,l+1}^\dagger f_{\alpha,l}] \\
&\quad + \Delta_\alpha [f_{\alpha,l}^\dagger f_{\alpha,l+1}^\dagger + f_{\alpha,l+1} f_{\alpha,l}]\} \\
&\quad - \mu_\alpha \sum_{l=1}^{N_\alpha} f_{\alpha,l}^\dagger f_{\alpha,l}, \\
H_{coup} &= w' [f_{L,1}^\dagger f_{C,1} + f_{C,1}^\dagger f_{L,1}] + \\
&\quad w' [f_{C,N_C}^\dagger f_{R,1} + f_{R,1}^\dagger f_{C,N_C}],
\end{aligned} \tag{17.3}$$

where $w_\alpha = J_\alpha(1 + \lambda_\alpha)/4$, $\Delta_\alpha = J_\alpha(1 - \lambda_\alpha)/4$, $\mu_\alpha = B_\alpha^z$, and $w' = J'/2$. Therefore, in the language of fermionic operators, the Hamiltonian contains “normal” terms, with a creation and a destruction operator, as well as “anomalous” terms, with two creation or two destruction operators. The normal ones are a hopping term between nearest neighbors (w_α), which is originated in the flip-flop spin terms and a chemical potential (μ_α) coupled to the fermionic density, which comes from the magnetic field pointing along the z -direction. The anomalous terms (Δ_α) are similar to those of a one-dimensional Hamiltonian with a gap function with p -wave symmetry, decoupled in the Bardeen-Cooper-Schrieffer (BCS) approximation and are originated by the anisotropy between the X and Y exchange interaction.

Inspired in this analogy, we focus our study on a junction between a chain with isotropic interactions (XX spin chain) and an anisotropic one (XY spin chain), which in the fermionic language is similar to a normal-superconductor junction. Such a situation is realized in a configuration with $\lambda_R = \lambda_C = 0$ and $\lambda_L \neq 0$. We also assume that the left and right chains are at temperatures T_L and T_R respectively and they are both of infinite length ($N_L \rightarrow \infty$ and $N_R \rightarrow \infty$).

17.2.2 Energy balance

A consistent procedure to define an expression for the heat current from first principles, is to analyze the evolution of the energy stored in a small volume of the system and derive the corresponding equation for the conservation of the energy. [14] For the present Hamiltonian we choose an elementary volume containing two nearest-neighbor positions of the chain. We place the volume enclosing the sites $l, l+1$ within the central (XX) chain, which in the fermionic language contains only normal terms. We work in units where $\hbar = 1$. The equation for the conservation of the energy enclosed by this volume is

$$\begin{aligned}
\frac{dE_{l,l+1}}{dt} &= \frac{J_C}{2} \frac{d}{dt} \langle S_l^+ S_{l+1}^- + S_l^- S_{l+1}^+ \rangle - B_C^z \frac{d\langle S_l^z \rangle}{dt} \\
&= w_C \frac{d}{dt} \langle f_{C,l}^\dagger f_{C,l+1} + f_{C,l+1}^\dagger f_{C,l} \rangle - \\
&\quad B_C^z \frac{d}{dt} \langle f_{C,l}^\dagger f_{C,l} \rangle \\
&= -iw_C \langle [H, f_{C,l}^\dagger f_{C,l+1} + f_{C,l+1}^\dagger f_{C,l}] \rangle \\
&\quad + i\mu_C \langle [H, f_{C,l}^\dagger f_{C,l}] \rangle = J_{l+1,l+2}^Q - J_{l-1,l}^Q,
\end{aligned} \tag{17.4}$$

where $J_{l,l+1}^Q$ is the heat current flowing from l to $l+1$, which in the present setup coincides with the energy current. Its explicit expression is obtained from the evaluation of the above

commutator, which gives

$$\begin{aligned} J^Q &= J_{l,l+1}^Q = J_{l-1,l}^Q \\ &= i\varepsilon_{l,l+1}^C (\varepsilon_{l-1,l}^C \langle f_{C,l-1}^\dagger f_{C,l+1} - f_{C,l+1}^\dagger f_{C,l-1} \rangle \\ &\quad + \varepsilon_{l+1,l+1}^C \langle f_{C,l+2}^\dagger f_{C,l+1} - f_{C,l+1}^\dagger f_{C,l+2} \rangle), \end{aligned} \quad (17.5)$$

where $\varepsilon_{l,l'}^C$ denotes the matrix element l, l' of the Hamiltonian H_C . In order to evaluate the above current it is convenient to introduce the lesser Green's functions

$$G_{\alpha l, \beta l'}^<(t, t') = i \langle f_{\beta l'}^\dagger(t') f_{\alpha l}(t) \rangle, \quad (17.6)$$

thus

$$\begin{aligned} J^Q &= 2\text{Re}\{ \varepsilon_{l,l+1}^C G_{Cl+1, Cl-1}^<(t, t) \varepsilon_{l-1,l}^C \\ &\quad + \varepsilon_{l+1,l+1}^C G_{Cl+1, Cl+2}^<(t, t) \varepsilon_{l+2,l+1}^C \}. \end{aligned} \quad (17.7)$$

The lesser Green's functions are one of the basic elements within Keldysh non-equilibrium Green's function formalism [13]. They are evaluated by solving the equations of motion (Dyson's equations), which for our model can be written as follows

$$\sum_k G_{Cl, Ck}^<(\omega) [\delta_{k,l'}\omega - \varepsilon_{k,l'}^C] = 0 \quad (17.8)$$

for coordinates l, l' lying within the central chain. We have used the stationary property of the system, as a consequence of which the Green's functions depend on the difference $t - t'$, which allows us to transform: $G_{j,j'}^<(t - t') = \int_{-\infty}^{+\infty} d\omega / (2\pi) e^{-i\omega(t-t')} G_{j,j'}^<(\omega)$. Thus, using the above equation in Eq. (17.7) the energy current can be also expressed in the following way

$$\begin{aligned} J^Q &= 2\text{Re}\{ \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \omega [\varepsilon_{l,l+1}^C G_{Cl+1, Cl}^<(\omega) \\ &\quad + \varepsilon_{l+1,l+1}^C G_{Cl+1, Cl+1}^<(\omega)] \}. \end{aligned} \quad (17.9)$$

However $\text{Re}\{G_{j+1,j+1}^<(\omega)\} = 0$. Thus, the heat current reduces to

$$J^Q = 2\text{Re}\{ \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \omega \varepsilon_{l,l+1}^C G_{Cl+1, Cl}^<(\omega) \}. \quad (17.10)$$

Similarly, if we evaluate the heat current through the contacts $L - C$ and $C - R$, we find

$$\begin{aligned} J^Q &= 2w' \text{Re}\{ \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \omega G_{L1, C1}^<(\omega) \}, \\ &= 2w' \text{Re}\{ \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \omega G_{CN, R1}^<(\omega) \}. \end{aligned} \quad (17.11)$$

17.2.3 Solving Dyson's equations

In order to evaluate $G^<$ and the heat current we follow a treatment close to that presented in Ref. [15]. We define the retarded “normal” and “Gorkov” Green's functions

$$\begin{aligned} G_{\alpha l, \beta l'}^R(t, t') &= -i\Theta(t - t') \langle \{f_{\alpha l}(t), f_{\beta l'}^\dagger(t')\} \rangle, \\ F_{\alpha l, \beta l'}^R(t, t') &= -i\Theta(t - t') \langle \{f_{\alpha l}^\dagger(t), f_{\beta l'}(t')\} \rangle. \end{aligned} \quad (17.12)$$

Before writing down the Dyson's equations satisfied the these “full” Green's function let us define the following “free” particle $\hat{g}_\alpha^0(\omega)$ and hole $\hat{\bar{g}}_\alpha^0(\omega)$ Green's functions

$$\begin{aligned} [\hat{g}_\alpha^0(\omega)]_{\alpha l, \alpha l'}^{-1} &= \delta_{l, l'}(\omega + i\eta) - \varepsilon_{l, l'}^\alpha \\ [\hat{\bar{g}}_\alpha^0(\omega)]_{\alpha l, \alpha l'}^{-1} &= \delta_{l, l'}(\omega + i\eta) + \varepsilon_{l, l'}^\alpha, \end{aligned} \quad (17.13)$$

with $\eta = 0^+$.

From now on we will work in Fourier space and we will not write explicitly the ω dependence of Green's functions unless necessary. For the left chain, we also define the functions $\hat{G}_L^0, \hat{\bar{G}}_L^0$ containing the pairing term contribution, through the relations

$$\begin{aligned} [\hat{G}_L^0]^{-1} &= [\hat{g}_L^0]^{-1} - \hat{\Delta}_L \hat{\bar{g}}_L^0 \hat{\Delta}_L, \\ [\hat{\bar{G}}_L^0]^{-1} &= [\hat{\bar{g}}_L^0]^{-1} - \hat{\Delta}_L \hat{g}_L^0 \hat{\Delta}_L, \end{aligned} \quad (17.14)$$

We also introduce

$$\begin{aligned} \hat{g}^0 &= \sum_{\alpha=L, C, R} \hat{g}_\alpha^0, \quad \hat{\bar{g}}^0 = \sum_{\alpha=L, C, R} \hat{\bar{g}}_\alpha^0, \\ \hat{G}^0 &= \hat{G}_L^0 + \hat{g}_C^0 + \hat{g}_R^0, \\ \hat{\bar{G}}^0 &= \hat{\bar{G}}_L^0 + \hat{\bar{g}}_C^0 + \hat{\bar{g}}_R^0, \\ \hat{F}_L^0 &= \hat{\bar{G}}^0 \hat{\Delta}_L \hat{g}_L^0, \\ \hat{\bar{F}}_L^0 &= \hat{G}^0 \hat{\Delta}_L \hat{\bar{g}}_L^0. \end{aligned} \quad (17.15)$$

Here $\varepsilon_{l, l'}^\alpha$ and $\Delta_{\alpha l, \alpha l'}$ are matrices defined on the coordinates of the chain $\alpha = L, C$, or R , containing respectively, the normal and anomalous elements of the Hamiltonian. In the case we are studying only $\hat{\Delta}_L$ is non-vanishing.

To obtain these Green's functions, we write the following Dyson's equation which relate them with Green's functions of the “disconnected” chains $\hat{g}_\alpha^0$ and $\hat{\bar{g}}_\alpha^0$ (see below) and the matrix elements of the contacts

$$\{[\hat{g}_L^0]^{-1} + [\hat{g}_C^0]^{-1} + [\hat{g}_R^0]^{-1} - [\hat{W}]\} \hat{G}^R - \hat{\Delta}_L \hat{F}^R = \hat{1}, \quad (17.16)$$

$$\{[\hat{\bar{g}}_L^0]^{-1} + [\hat{\bar{g}}_C^0]^{-1} + [\hat{\bar{g}}_R^0]^{-1} + [\hat{W}]\} \hat{\bar{F}}^R - \hat{\Delta}_L \hat{G}^R = \hat{0}. \quad (17.17)$$

The matrix $\hat{W} = \hat{W}_L + \hat{W}_R$ contains the matrix elements of H_{cont} describing the connections between the central and left parts and between the central and right parts.

The above equations can be rewritten in a more convenient form by recourse to the following procedure. From Eq. (17.17)

$$\hat{F}^R = \hat{\bar{g}}^0 [\hat{\Delta}_L \hat{G}^R - \hat{W} \hat{F}^R], \quad (17.18)$$

$$\hat{G}^R = \hat{g}^0 + \hat{g}^0 [\hat{\Delta}_L \hat{F}^R + \hat{W} \hat{G}^R], \quad (17.19)$$

Substituting (17.18) in (17.16) and (17.19) in (17.17) one obtains

$$\hat{G}^R = \hat{G}^0 (1 + \hat{W} \hat{G}^R) + \hat{\bar{F}}_L^0 \hat{W} \hat{F}^R, \quad (17.20)$$

$$\hat{F}^R = \hat{F}_L^0 (1 + \hat{W} \hat{G}^R) - \hat{\bar{G}}^0 \hat{W} \hat{F}^R, \quad (17.21)$$

Let us now consider Eq. (17.20) for the following particular coordinates

$$G_{Cl,C\ell'}^R = g_{C,l,\ell'}^0 + g_{C,l,1}^0 w' G_{L1,C\ell'}^R + g_{C,l,N}^0 w' G_{R1,C\ell'}^R, \quad (17.22)$$

$$G_{L1,C\ell'}^R = G_{L,1,1}^0 w' G_{C1,C\ell'}^R + \bar{F}_{L,1,1}^0 w' F_{C1,C\ell'}^R, \quad (17.23)$$

$$G_{R1,C\ell'}^R = g_{R,1,1}^0 w' G_{CN,C\ell'}^R. \quad (17.24)$$

Substituting Eqs. (17.23) and (17.24) in Eq. (17.22) it is easy to see that the Dyson's equation for the two indices corresponding to coordinates of C can be written as follows

$$\{[\hat{g}_C^0]^{-1} - \hat{\Sigma}^{R,gg}\} \hat{G}_C^R + \hat{\Sigma}^{R,gf} \hat{F}_C^R = \hat{1}, \quad (17.25)$$

$$\{[\hat{g}_L^0]^{-1} - \hat{\Sigma}^{R,ff}\} \hat{F}_C^R + \hat{\Sigma}^{R,fg} \hat{G}_C^R = \hat{0}, \quad (17.26)$$

where the matrices of the above equations have sizes $N_C \times N_C$ and elements corresponding to the coordinates of the central chain. The “self-energy” matrices are

$$\begin{aligned} \Sigma_{l,\ell'}^{R,ff} &= \delta_{l,\ell'} |w'|^2 [\delta_{l,1} \bar{G}_{L,1,1}^0 + \delta_{l,N_C} \bar{g}_{R,1,1}^0] \\ \Sigma_{l,\ell'}^{R,gg} &= \delta_{l,\ell'} |w'|^2 [\delta_{l,1} G_{L,1,1}^0 + \delta_{l,N_C} g_{R,1,1}^0] \\ \Sigma_{l,\ell'}^{R,gf} &= \delta_{l,\ell'} |w'|^2 \delta_{l,1} \bar{F}_{L,1,1}^0, \\ \Sigma_{l,\ell'}^{R,fg} &= \delta_{l,\ell'} |w'|^2 \delta_{l,1} F_{L,1,1}^0. \end{aligned} \quad (17.27)$$

The explicit expressions for these functions imply the evaluation of all the functions appearing in the right hand sides of (17.27). Notice that these functions have been defined from manipulations of the Dyson's equations corresponding to H_L or H_R *isolated* from the central chain. We indicate a procedure for the calculation of these functions in appendix A. Note also that since the right and left parts of the system are held at two different but constant temperatures, these Green's functions can be calculated at equilibrium. The advantage of the above representation becomes clear by writing (17.26) as

$$\begin{aligned} \hat{F}_C^R &= \hat{g}_C \hat{\Sigma}^{R,fg} \hat{G}_C^R, \\ [\hat{g}_C]^{-1} &= [\hat{g}_C^0]^{-1} - \hat{\Sigma}^{R,ff}, \end{aligned} \quad (17.28)$$

and substituting it in (17.25). The result leads to the solution of the retarded normal Green's function within C

$$\begin{aligned} \hat{G}_C^R &= \{[\hat{g}_C^0]^{-1} - \hat{\Sigma}_{\text{eff}}^{-1}\}^{-1}, \\ \hat{\Sigma}_{\text{eff}}^R &= \hat{\Sigma}^{R,gg} + \hat{\Sigma}^{R,gf} \hat{g}_C \hat{\Sigma}^{R,fg}. \end{aligned} \quad (17.29)$$

The results obtained so far correspond to the retarded Green's functions and self energies. The lesser Green's function with coordinates within C can be easily obtained by recourse to Langreth rules [16, 17]. In particular one obtains [15]

$$\hat{G}_C^< = \hat{G}_C^R \hat{\Sigma}_{\text{eff}}^< \hat{G}_C^A, \quad (17.30)$$

where the advanced Green's function is obtained from the retarded one, by means of the relation $\hat{G}_C^A(\omega) = [\hat{G}_C^R(\omega)]^\dagger$ and the lesser component of the “self-energy” is

$$\begin{aligned} \hat{\Sigma}_{\text{eff}}^< &= \hat{\Sigma}^{<,gg} + \hat{\Lambda}^R \hat{\Sigma}^{<,fg} + \\ &\hat{\Sigma}^{<,gf} \hat{\Lambda}^A + \hat{\Lambda}^R \hat{\Sigma}^{<,ff} \hat{\Lambda}^A, \end{aligned} \quad (17.31)$$

with $\hat{\Lambda}^R(\omega) = \hat{\Sigma}_\alpha^{R,gf}(\omega)\hat{g}_C(\omega)$ and $\hat{\Lambda}^A(\omega) = [\hat{\Lambda}^R(\omega)]^\dagger$. The self-energies have components

$$\Sigma_{l,l'}^{<,\nu,\nu'}(\omega) = i\delta_{l,l'}[\delta_{l,1}\Gamma_L^{\nu,\nu'}(\omega)f_L(\omega) + \delta_{l,N}\Gamma_R^{\nu,\nu'}(\omega)f_R(\omega)]$$

being $\Gamma_L^{\nu,\nu'}(\omega) = -2\text{Im}[\Sigma^{R,\nu\nu'}(\omega)_{1,1}]$ and $\Gamma_R^{\nu,\nu'}(\omega) = -2\text{Im}[\Sigma^{R,\nu\nu'}(\omega)_{N,N}]$ with $\nu, \nu' = g, f$. The Fermi functions $f_\alpha(\omega)$, with $\alpha = L, R$ depend on the temperatures T_L and T_R of the left and right chains respectively: $f_\alpha(\omega) = 1/(1 + e^{\omega/T_\alpha})$, in units where $k_B = 1$.

Finally, the lesser counterparts of Eqs. (17.23) and (17.24), which correspond to Green's functions with one of the coordinates in the central (C) chain and the other one in the left (L) or right (R) chain, can be calculated by recourse again to Langreth rules [16, 17]

$$G_{L1,C\nu'}^{<} = G_{L,1,1}^{0,<}w'G_{C1,C\nu'}^A + G_{L,1,1}^{0,R}w'G_{C1,C\nu'}^{<} + \bar{F}_{L,1,1}^{0,<}w'F_{C1,C\nu'}^A + \bar{F}_{L,1,1}^{0,R}w'F_{C1,C\nu'}^{<}, \quad (17.32)$$

$$G_{R1,C\nu'}^{<} = g_{R,1,1}^{0,<}w'G_{CN,C\nu'}^A + g_{R,1,1}^{0,R}w'G_{CN,C\nu'}^{<}. \quad (17.33)$$

17.2.4 Heat currents and transmission functions

We focus on the expression for the heat current evaluated in the contact between the central chain C and R given in Eq. (17.11). Using Eq. (17.33) one obtains

$$J^Q = -2|w'|^2 \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \omega \text{Re}[G_{CN,CN}^{<}(\omega)g_{R,1,1}^{0,A}(\omega) + G_{CN,CN}^R(\omega)g_{R,1,1}^{0,<}(\omega)]. \quad (17.34)$$

Using (17.30) and after some algebra (see Ref. [15]), it is found

$$\begin{aligned} J^Q &= \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \omega [f_L(\omega) - f_R(\omega)] [T^n(\omega) - T^a(\omega)], \\ T^n(\omega) &= \Gamma_R^{gg}(\omega) |G_{C,N,1}^R(\omega)|^2 \Gamma_{L,\text{eff}}^{gg}(\omega), \\ T^a(\omega) &= \Gamma_R^{gg}(\omega) |\bar{\Lambda}_{N,N}(\omega)|^2 \Gamma_R^{ff}(\omega), \end{aligned} \quad (17.35)$$

where

$$\begin{aligned} \Gamma_{L,\text{eff}}^{gg} &= \Gamma_L^{gg} + 2\text{Re}[\Gamma_L^{gf}\Lambda_{1,N}^A] + |\Lambda_{1,1}^R|^2 \Gamma_L^{ff} + |\Lambda_{1,N}^R|^2 \Gamma_R^{ff}, \\ \bar{\Lambda}_{N,N}^R &= G_{C,N,1}^R \Lambda_{1,N}^R. \end{aligned} \quad (17.36)$$

The difference of Fermi functions in the expression of J^Q , reflects the fact that the existence of a non-vanishing heat current through the central system depends on the existence of a difference of temperatures between the left and right chains. The details of the model are enclosed in the behavior of the “normal” and “anomalous” transmission functions $T^n(\omega)$ and $T^a(\omega)$, which are analogous to those defined in Ref. [15] in the context of particle transport in a setup with normal and superconducting wires. The first function has, in fact, the structure of a transmission. Notice that it depends on the densities of states of the right and left chains through the functions Γ_R and $\Gamma_{L,\text{eff}}^{gg}$, and one the particle propagator between the first and last points of the central chain. Instead, $T^a(\omega)$ actually has the structure of a reflection process. Notice that it depends on the density of states for particles and holes of the right reservoir and on a multiparticle propagator $\bar{\Lambda}_{N,N}^R$ at the last point of the central chain. Typical plots for these functions are shown in Fig. 87. These functions do not depend on the temperatures T_L and T_R and are non-vanishing only within a finite range of energies of a width that is set by

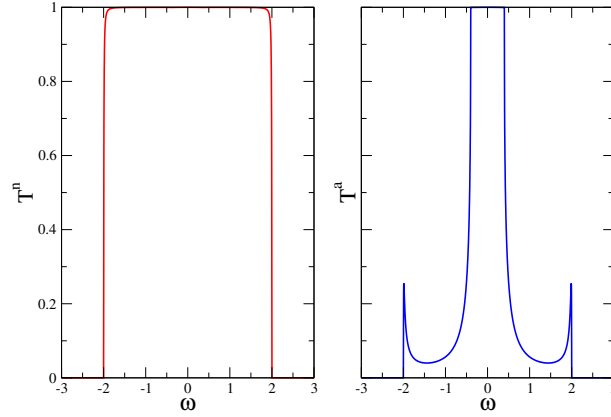


Figure 87: (Color online) “normal” (left) and “anomalous” (right) transmission functions. The latter is the counterpart to Andreev transmission function in electronic systems. Parameters are $\Delta_L = 0.2$, $J_L = J_R = J_C = J' = 1$ and all chemical potentials set to zero.

the largest exchange parameter between the left, right and central chains. These functions are symmetric with respect to $\omega = 0$ for $B_L = B_R = 0$ (see Fig. 87). This symmetry is broken for finite B_α , since the effect of a finite magnetic field in one of the side chains is to shift the corresponding function as $\Gamma_\alpha^{\nu\nu'}(\omega) \rightarrow \Gamma_\alpha^{\nu\nu'}(\omega - B_\alpha)$.

In the language of fermionic systems, two different kinds of processes take place in a normal-superconductor junction. For energies higher than the gap, the transport is due to the tunneling of normal single particle high-energy excitations. This mechanism contributes to the electronic transmission function $T^n(\omega)$. Instead, for low energies, below the gap, the transport is due to the mechanism known as “Andreev reflection”, which implies the combination of two fermions of the normal side into a Cooper pair within the superconducting one, leaving a hole that is reflected back from the junction into the normal side. Because of this mechanism, $T^n(\omega) \sim T^a(\omega) \sim 1$ for energies within the superconducting gap, i.e. $|\omega| \leq \Delta_L$. The effective conversion of electrons into Cooper pairs taking place in the mechanism of Andreev reflection helps to particle transport. Mathematically, this is reflected by the fact that the total particle transmission function is $T^n(\omega) + T^a(\omega)$. [15, 18] Instead, in the case of heat transport, $T^a(\omega)$ and $T^n(\omega)$ contribute with opposite sign, as explicitly shown in Eq. (17.35), i.e. the mechanism of Andreev reflection, plays a negative role regarding the heat transport. The consequence is a vanishing heat transport due to excitations within the energy window defined by the superconducting gap.

In the original language of interacting spins, the above picture translates as follows. Low energy spin excitations traveling from the isotropic chain via flip-flop processes in the z direction meet an energy gap at the other side of junction due to the anisotropic interaction which tends to favor flip-flop processes in a different direction. This favors the simultaneous raising or lowering of two spins at two neighboring positions of the chain and causes multiscattering processes in which a portion of the incident spin wave packet manages to twist and propagate into the other side, at the same time that a portion becomes reflected and propagates back.

We can describe the behavior of J^Q for low T and small temperature gradients δT as follows. Writing $T_R = T$ and $T_L = T_R + \delta T$ we can approximate the difference of Fermi functions in Eq. (17.35) as

$$f_L(\omega) - f_R(\omega) \sim \frac{\partial f_R(\omega)}{\partial T} \delta T. \quad (17.37)$$

On the other hand, from Fig. 87, we can write,

$$T^n(\omega) - T^a(\omega) \sim 0 \quad \omega \leq 2\Delta_L \quad (17.38)$$

$$T^n(\omega) - T^a(\omega) \sim 1 \quad 2\Delta_L \leq \omega \leq 2J \quad (17.39)$$

leading to,

$$J^Q = \frac{\delta T}{\pi} \int_{2\Delta_L}^{+2J} d\omega \omega \frac{\partial f_R(\omega)}{\partial T} \quad (17.40)$$

For low enough temperatures, $T \ll 2\Delta_L \ll 2J$, this expression can be further approximated as,

$$J^Q \sim \frac{\delta T}{\pi} \frac{\partial}{\partial T} \int_{2\Delta_L}^{\infty} d\omega \omega e^{-\beta\omega} \quad (17.41)$$

$$\sim \frac{4}{\pi} \delta T \frac{\Delta_L^2}{T} e^{-\frac{2\Delta_L}{T}} \quad (17.42)$$

Therefore, for $T < \Delta_L$, the heat current is exponentially small.

On the other hand, for $\Delta_L = 0$, the behavior of the J^Q is fully due to normal tunneling. For low T we can perform a Sommerfeld expansion on the Fermi function to get

$$\begin{aligned} J^Q &= \frac{2}{T^2} \delta T \frac{\partial}{\partial \beta} \left(\int_0^{+\infty} \frac{d\omega}{2\pi} \omega f(\omega) T^n(\omega) \right) \\ &\sim \frac{\pi}{3} T \delta T, \quad T \ll J_\alpha, \quad \Delta_L = 0. \end{aligned} \quad (17.43)$$

17.3 Results

In this section we discuss the behavior of the heat current as a function of the different ingredients of the spin system. For simplicity, we consider identical exchange parameters along the left, central and right chains: $J_L = J_C = J_R = J' = J$. Without loss of generality we set $J = 1$. Thus, we focus on a spin heterostructure with a single junction between a semi-infinite XX and a semi-infinite XY chain, which in the fermionic language translates to a single S-N junction. For this particular configuration, our results do not depend on the length of the central chain.

As discussed in the previous subsection, the structure of the expression (17.35) for the heat current clearly reflects the fact that for small temperature differences, we obtain a behavior of the form

$$J^Q = -K \delta T, \quad (17.44)$$

where the coefficient K can be interpreted as a *thermal conductance*. It is tempting to relate this coefficient with the conductivity κ evaluated in several works on the basis of linear response theory. If we assume that the relation between the two coefficients is similar to the one between electrical conductance and electrical conductivity, K and κ should differ just by a geometrical factor. However, to the best of our knowledge, a rigorous relation between these two coefficients has not been presented in the literature. Nevertheless, the behavior of K as a function of T shown in the left upper panel of Fig. 88 for the case of two connected XX chains (see the plot corresponding to $\Delta_L = 0$) is similar to the one reported in the literature for homogeneous and isotropic chains [3], [4],[5]. In this case, the anomalous component is zero and K increases linearly in T for low temperature [see Eq. (17.43)], as discussed at the end of the previous section. The conductance reaches a maximum at $T \sim J$. and decreases at higher temperatures, as a consequence of the finite bandwidth (energy window) for the spin excitations amenable to

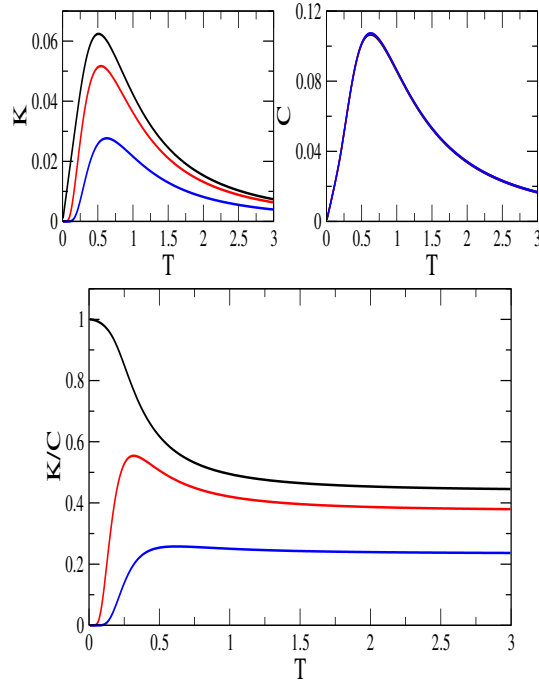


Figure 88: (Color online) Top left: Conductance, defined as the ratio of heat current to temperature gradient between right and left chains. Top right: specific heat of the central system. Bottom: ratio of the above magnitudes. Parameters are $T_L = T + \delta_T$, $T_R = T$ where $\delta_T = 0.005$, all chemical potentials set to zero, and (from top to bottom) $\Delta_L = 0, 0.25, 0.5$.

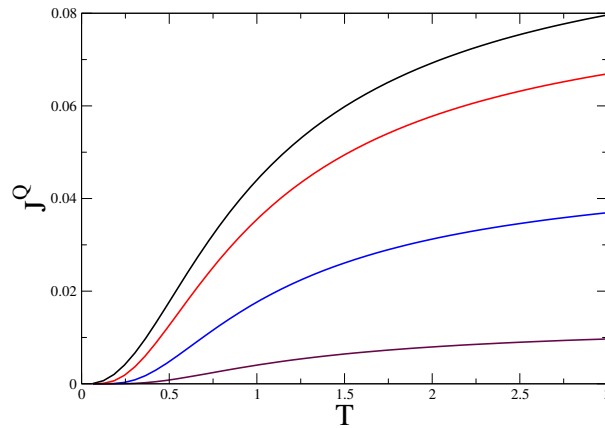


Figure 89: (Color online) Heat current as a function of the temperature of the left-chain (right chain at fixed $T_R = 0$) for all chemical potentials set to zero, all hopping parameters set to 1 and different values of the pairing parameter Δ_L , from top to bottom $\Delta_L = 0.1, 0.25, 0.5, 0.75$.

cross the central chain transporting energy from one side to the other one. As expected, for a fixed temperature K decreases for increasing values of Δ_L . In agreement with the behavior discussed in the previous section, K is exponentially small for $T < \Delta_L$ [see Eq. (17.40)]. For higher temperatures, the high energy excitations are allowed to perform tunneling above the energy gap, with the concomitant increase of K . As in the case with $\Delta_L = 0$, the maximum is achieved at $T \sim J$.

We also evaluate the specific heat for the *equilibrium* central system in contact to the side chains at the same temperature T as follows:

$$C(T) = -\frac{2}{N} \sum_{l=1}^N \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \frac{\partial f(\omega)}{\partial T} \omega \text{Im}[G_{C,l,l}^R(\omega)] \quad (17.45)$$

In a normal metallic system as described by Drude model, this quantity is related to the thermal conductivity through $\kappa = vlC/3$, being v the Fermi velocity of the electrons and l their mean free path [19]. We plot this quantity in the right upper panel of Fig. 88. This physical quantity is almost insensitive to the opening of the energy gap and the different plots, corresponding to different values of Δ_L almost coincide within the scale of the figure. From the lower panel of Fig. 88 we see that while for high temperatures there is a linear relation between K and C , this is not the case at lower temperatures where Andreev type processes are relevant.

As stressed before, our calculation is not restricted to small temperature gradients. We show in Fig. 89 a plot of the heat current for several values of the anisotropy parameter Δ_L as a function of the temperature of the left chain while the temperature of the right chain T is set fixed to zero. The figure clearly shows the suppression of the current as a consequence of the “Andreev reflection” phenomena mentioned before. In fact for $T < \Delta_L$ the current is exponentially small, while it grows for higher temperatures.

Finally, in Fig. 90 we illustrate the behavior of the heat current when finite different magnetic fields are applied at the two side chains. The effect of applying magnetic fields at both sides of the junction leads to a interesting effect which we name “thermal diode effect”. As discussed in the previous section, a finite magnetic field originates a shift in the arguments of the functions $\Gamma_{\alpha}^{\nu\nu'}(\omega)$, which leads to asymmetries in the transmission functions $T^n(\omega)$ and $T^a(\omega)$. For $\Delta_L = 0$, only the normal transmission function and the functions $\Gamma_{\alpha}^{gg}(\omega)$ are non-vanishing. Furthermore, these functions are identical and gapless for $\alpha = L, R$. Therefore, the heat flow is perfectly antisymmetrical (the sign of the current is reversed preserving the absolute value) under the simultaneous change $T_L \leftrightarrow T_R$ and $B_L \rightarrow B_R$. Instead, for a finite Δ_L , the situation changes. A gap opens for the excitations of the left chain and the functions $\Gamma_L^{\nu\nu'}(\omega) = 0$ vanish for $|\omega| < \Delta_L$, while the functions $\Gamma_R^{gg}(\omega) = \Gamma_R^{ff}(-\omega)$ remain finite. The consequence is an asymmetry in the behavior of the transmission functions under the change $B_L \rightarrow B_R$. The result is an effect of thermal rectification. That is, the magnitude of the current J^Q when $(T_L = T, \mu_L = \mu)$ and $(T_R = T', \mu_R = \mu')$ is different to $J^{Q'}$ when $(T_L = T', \mu_L = \mu')$ and $(T_R = T, \mu_R = \mu)$, which means that the device is more likely to conduct heat when the temperature difference is applied in one direction than in the other. We display the phenomena for two different values of Δ_L . We show the current when $(\mu_L = 0.3, T_L = T, \mu_R = 0, T_R = 0)$ with dots and the current when $(\mu_R = 0.3, T_R = T, \mu_L = 0, T_L = 0)$ with a full line. When the value of $\Delta_L = 0.2$ both currents are rather large and similar but when $\Delta_L = 0.75$ the currents are smaller and clearly different.

17.4 Summary and discussion

We have presented a theoretical framework to study heat transport in one-dimensional spin heterostructures.

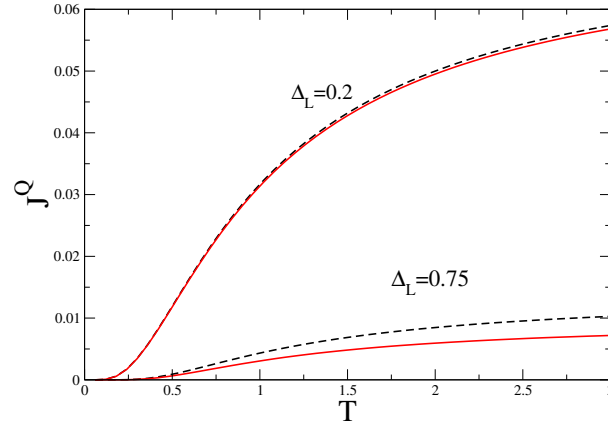


Figure 90: (Color online) Thermal Diode effect. We show the heat current when $(\mu_L = 0.3, T_L = T, \mu_R = 0, T_R = 0)$ (dotted line) and when $(\mu_R = 0.3, T_R = T, \mu_L = 0, T_L = 0)$ for two values of the pairing parameter $\Delta_L = 0.2, 0.75$

In the present work we have focussed on a simple system composed of a junction between an anisotropic (XY) and an isotropic (XX) chain under the effect of an inhomogeneous magnetic field along the z direction. Using the Jordan-Wigner transformation to map the problem into a fermionic system and using the non equilibrium Keldysh-Schwinger formalism we have obtained exact expression for the heat current in terms of Green's functions of the "disconnected" spin chain components. The resulting expressions can be evaluated numerically in a simple way. In the limits $\Delta_L \ll T \ll J$ and $T \ll \Delta_L \ll J$ explicit analytic expressions can also be given. We have studied the heat transport as a function of the different parameters of the model and we have shown that when different magnetic fields are applied at the end chains, a rectifying effect in the heat current occurs. This effect might be of interest for applications. Its origin can be traced back to the appearance of pairing terms induced by the anisotropy parameter, which are in turn responsible for an Andreev reflection type mechanism.

In this work we have analyzed a simple model. However this methodology can be straightforwardly extended to more complex structures with many junctions and disorder. Our treatment relies on the Jordan-Wigner transformation which maps the original spin Hamiltonians into fermionic ones. In the case we have considered, the latter are bilinear. In more generic models, although we expect the rectifying effect still to be present, the technical analysis could be more complicated. For instance, in the isotropic Heisenberg model, which in addition to the exchange interaction along x and y directions, contains an additional exchange term along the z direction, the Jordan-Wigner transformation translates such a term into a many-body fermionic interaction, which does not enable a straightforward analytical solution of the problem, as in the case we considered here. Nevertheless, the Green's function formalism offers a framework for the construction of systematic approximations to treat those terms. Numerical methods could also be useful to deal with models containing many-body terms. [22] As in electronic systems, many-body terms are expected to introduce further inelastic scattering processes, which could add further ingredients in addition to the transport mechanisms we have discussed here. We hope to report on some of these issues in future work.

17.5 Green's functions for an open chain with p -wave superconductivity

In this appendix we show a derivation of the Green's functions $G_{L,1,1}^0(\omega)$, $\bar{G}_{L,1,1}^0(\omega)$, $F_{L,1,1}^0(\omega)$ and $\bar{F}_{L,1,1}^0(\omega)$, entering Eq. (17.27), which correspond to the end of a half infinite chain with p -wave superconductivity in the BCS approximation. Making the superconducting parameter $\Delta_L = 0$, the first two Green's functions give the corresponding result for the normal chain $g_{R,1,1}^0(\omega)$ and $\bar{g}_{R,1,1}^0(\omega)$ respectively.

The Green's functions of the open chain can be solved considering a ring of N sites, periodic except for the fact that the energy at one site (which we label as site 0) is increased by an energy A , and then taking the limit $N, A \rightarrow +\infty$. The Hamiltonian is

$$\begin{aligned} H = & \sum_{l=0}^{N-1} \{w[f_l^\dagger f_{l+1} + f_{l+1}^\dagger f_l] \\ & + \Delta[f_l^\dagger f_{l+1}^\dagger + f_{l+1} f_l]\} \\ & - \mu \sum_{l=0}^{N-1} f_l^\dagger f_l + A f_0^\dagger f_0. \end{aligned} \quad (17.46)$$

We have solved the problem using two different methods: i) solving the equations of motion in Fourier space, and ii) solving a Dyson's equation that relates the above Green's functions to those of the periodic chain ($A = 0$) which can be obtained easily using Bloch theorem. Both results of course coincide, but the latter method involves a simpler algebra. We define a matrix

$$\tilde{G} = \begin{pmatrix} G_{L,1,1}^0(\omega) & F_{L,1,1}^0(\omega) \\ \bar{F}_{L,1,1}^0(\omega) & \bar{G}_{L,1,1}^0(\omega) \end{pmatrix}, \quad (17.47)$$

with the Green's functions for $A \neq 0$, and a corresponding matrix $\tilde{g}$ for $A = 0$. These matrices are equivalent to the ones obtained by using Nambu's representation for the Hamiltonian and the Green's functions. [20] From the equations of motion of these Green's functions, one obtains

$$\tilde{G} = \tilde{g} + \tilde{g} \tilde{A} \tilde{G}, \quad (17.48)$$

where $\tilde{A}$ is proportional to A .

Solving Eq. (17.48) for $\tilde{G}$ and taking the limit $A \rightarrow +\infty$ the following expressions result

$$G_{L,1,1}^0 = h_0(\omega) - \frac{h_1^2(\omega)}{h_0(\omega)} - \frac{h_2^2(\omega)}{h_0^*(-\omega)}. \quad (17.49)$$

$$F_{L,1,1}^0 = h_2(\omega) \left(\frac{h_1^*(-\omega)}{h_0^*(-\omega)} - \frac{h_1(\omega)}{h_0(\omega)} \right). \quad (17.50)$$

$$\bar{G}_{L,1,1}^0(\omega) = -(G_{L,1,1}^0)^*(-\omega) \quad \bar{F}_{L,1,1}^0 = -(F_{L,1,1}^0)^*(-\omega) \quad (17.51)$$

The h -functions entering the second members of Eqs. (17.49) and (17.50) are Green's functions of the periodic chain and can be calculated easily in Fourier space. The result is

$$\begin{aligned}
h_0(\omega) &= \frac{1}{N} \sum_k \frac{\omega + \epsilon_k}{\omega^2 - \epsilon_k^2 - \Delta_k^2}, \\
h_1(\omega) &= \frac{1}{N} \sum_k \frac{(\omega + \epsilon_k) \cos k}{\omega^2 - \epsilon_k^2 - \Delta_k^2}, \\
h_2(\omega) &= \frac{1}{N} \sum_k \frac{2\Delta \sin^2 k}{\omega^2 - \epsilon_k^2 - \Delta_k^2},
\end{aligned} \tag{17.52}$$

where in the second members ω includes an infinitesimally small imaginary part, $\epsilon_k = 2w \cos k - \mu$, and $\Delta_k = 2\Delta \sin k$.

For $N \rightarrow +\infty$, the sums can be replaced by integrals. Decomposing the integrands into a sum of simple fractions with denominators linear in $\cos k$ and numerators independent of k , the integrals can be evaluated analytically using [21]

$$I(b) = \frac{1}{\pi} \int_0^\pi \frac{dk}{\cos k + b} = \frac{1}{\sqrt{b^2 - 1}}, \tag{17.53}$$

where the sign of the root is determined by the sign of the imaginary part of the second member.

Defining

$$\begin{aligned}
\tilde{\omega} &= \frac{\omega}{2w}, \quad \tilde{\Delta} = \frac{\Delta}{w}, \quad \tilde{\mu} = \frac{\mu}{2w}, \quad d = 1 - \tilde{\Delta}^2, \\
r &= \left[(\tilde{\omega}^2 - \tilde{\Delta}^2)d + (\tilde{\Delta}\tilde{\mu})^2 \right]^{1/2} / d, \\
b_1 &= \tilde{\mu}/d - r, \quad b_2 = \tilde{\mu}/d + r, \\
d_1 &= [4wrd]^{-1},
\end{aligned} \tag{17.54}$$

the result takes the form

$$\begin{aligned}
h_0(\omega) &= d_1 [(r - \tilde{\omega} - \tilde{\mu}\tilde{\Delta}^2/b)I(b_1) \\
&\quad + (r + \tilde{\omega} + \tilde{\mu}\tilde{\Delta}^2/b)I(b_2)], \\
h_1(\omega) &= -d_1 \{ 2r - [(r - \tilde{\omega} - \tilde{\mu}\tilde{\Delta}^2/b)b_1 I(b_1) \\
&\quad + (r + \tilde{\omega} + \tilde{\mu}\tilde{\Delta}^2/b)b_2 I(b_2)] \}, \\
h_2(\omega) &= d_1 \tilde{\Delta} \{ 2r + [I(b_1)]^{-1} - [I(b_2)]^{-1} \}.
\end{aligned} \tag{17.55}$$

17.6 Easy illustrative examples

17.6.1 Example 1

17.6.2 Example 2

18 Equivalence of wave function matching and Green's functions methods for quantum transport: generalized Fisher-Lee relation Hocine by Boumrar, Hamidi, Zenia, Lounis

Abstract

We present a proof of an exact equivalence of the two approaches that are most used in computing conductance in quantum electron and phonon transport: the wave function matching and Green's functions methods. We can obtain all the quantities defined in one method starting from those obtained in the other. This completes and illuminates the work started Ando[Ando T 1991 *Phys. Rev. B* **44** 8017.] and continued later by Khomyakov *et al.*[Khomyakov P A, Brocks G, Karpan V, Zwierzycki M and Kelly P J 2005 *Phys. Rev. B* **72** 035450.]. The aim is to allow for solving the transport problem with whichever approach fits most the system at hand. One major corollary of the proven equivalence is our derivation of a generalized Fisher-Lee formula for resolving the transmission function into individual phonon mode contributions. As an illustration, we applied our method to a simple model to highlight its accuracy and simplicity.

18.1 Conclusion

In conclusion, we have reformulated the conventional WFM method and completed its equivalence with the AGF approach. Through a clever choice of the evanescent eigenmodes entering the calculation, we need only use four Bloch matrices, getting rid of the redundant four others introduced in earlier literature, and which have no counterparts in the AGF method. This choice also means that the injection matrix selects from the outset the propagating modes as the sole that contribute to the transmission. In order to arrive at the equivalence, we showed how to obtain the group velocity matrices used in the WFM method from the “escape rates” defined in the AGF formalism. This then allowed us to proceed to a more physical spectral decomposition of the transmission amplitude. The “escape rates” are related to the self-energies, and these are computed differently in the WFM and in the AGF formalisms. In the first, they are obtained from the Bloch matrices, and in the second from the surface Green's functions of the leads. The findings of this study intimate that when proper care is taken in the choice of the phonon modes that enter the machinery of the WFM method, its equivalence with the AGF technique is complete. As an illustration, we applied our method to a simple model and, additionally, we made a thorough comparison with earlier methods. Our results demonstrate the equivalence of the two methods to a very high accuracy, limited only by the machine precision. The comparison with the Fisher-Lee relation sheds a light on its failure to properly decompose the transmission into individual contributions of the phonon modes.

We have succeeded in completing the equivalence of the two most widely used methods in ballistic transport at the nanoscale. Besides, we have devised a procedure that provides a number of advantages when compared to earlier approaches. The main advantages of our method are as follows:

1. We used the inverse of the Bloch factor in the eigenvalue problem of the left lead, conventionally chosen as the side from which the incident phonons arrive. This choice precludes many costly matrix inversions in the following steps of the procedure.
2. We defined four Bloch matrices, instead of eight as in Refs. [2, 32], which are enough to carry out all the subsequent calculations.
3. Our injection matrix selects out the evanescent modes from the outset. In prior methods [2], the latter are eliminated “by hand” or by using a projector operator to project out their unwanted contribution (Eq. 2.62 in [2]).
4. Each of the four Bloch matrices defined in this work has its equivalent quantity in the AGF method, contrary to the eight matrices defined in prior methods, where four of them have no corresponding quantities in the AGF method.

5. The transmission function in our method is given in terms of the bulk group velocities of the leads. These are guaranteed to be diagonal and positive, and their square roots are clearly defined. However, in prior methods, the transmission function is written in terms of the “escape rates”. These being defined as the difference between a self-energy matrix and its conjugate, are not necessarily real as is assumed in most publications [17, 45, 44, 46].
6. And of utmost importance, we derived a generalized Fisher-Lee formula to properly decompose the transmission function into individual phonon mode contributions, which is valid regardless of the orthogonality of the eigenvectors.

These findings add to a growing body of literature on the theoretical investigation of transport, thermal as well as electronic, through interfaces at the nanoscale. The complete equivalence of the two most widely used methods allows for a better understanding of the underlying physics, especially in the AGF method, where, despite its obvious mathematical superiority, lacks somewhat the transparency inherent to the WFM method.

18.2 Introduction

Electron and phonon transport in mesoscopic systems has gained much interest in the last three decades, due mainly to advances in experimental and computational techniques, as well as the continued improvements in computing resources[3, 4]. Indeed, experimental techniques are nowadays so precise, it has become feasible to wire a nanometric-sized molecule to two metallic leads and measure its transport properties[5]. The technological applications are also varied as a result of the trend to miniaturization, but also the push to exploit the novel effects observed only at the atomic scale. Heat dissipation, Josephson junctions, thermoelectricity, spintronics are among the most prominent applications that rely on such novel phenomena [6, 7, 8, 9]. In general, the setup consists of a central part, the device, considered small, connected to much larger leads. The leads act as heat and particle reservoirs each at a constant temperature and chemical potential. When the leads are kept at slightly different temperatures or chemical potentials, a heat or particle current ensues that passes through the central region. The first attempts at a theoretical understanding of the phenomenon were made in the second half of the last century. We note that when interactions are involved, the heat current in the electronic case is harder to describe. However, in this work we are concerned with the heat transport via harmonic phonons, whose treatment is also similar to the electronic charge transport. Using a phenomenological approach, Landauer[10, 11] expressed the conductance in a two-terminal setup in terms of a transmission probability. There are mainly two methods that are used to compute the latter. The first relies on using wave functions in the electronic case or displacements in the phononic case, and is sometimes known as the wave function matching (WFM) method[1, 12, 13, 14, 15]. The second uses the Green’s function as the main object, and is known as the atomistic Green’s functions method (AGF)[3, 16, 17, 18, 19, 20, 21].

The AGF was first applied to transport by Caroli *et al.*[22] using a Hamiltonian with a basis of localized wave functions. Their aim was to circumvent the inherently difficult problem of matching correctly the wave functions at the interfaces. Moreover, Green’s functions are the most convenient tool to include correlations in a more realistic treatment of the problem when these are important[9]. The method is also “unique” in that studying systems far from thermal equilibrium relies exclusively on the use of Green’s functions as introduced by the Keldysh formalism [4, 23, 24]. An alternative to the Green’s function method, consists of studying the scattering problem using wave functions. This requires that appropriate boundary conditions be imposed at the interfaces between the scattering central region and the leads. Ando[1] applied the wave function matching method to a tight-binding Hamiltonian, obtaining directly

the transmission as well as reflexion probability amplitudes. Ando's approach was criticized by Krstić *et al.*[25] for not yielding the correct expression of the conductance, due mainly, as their argument goes, to an inconsistent treatment of the evanescent modes. It is indeed the case that these modes do not participate in the transport, but their inclusion is necessary for a correct matching of the wave functions at the interfaces. Khomyakov *et al.*[2] later attempted to demonstrate the equivalence between the wave function matching and Green's functions methods and show that Ando's approach was correct. Their attempt was successful at deriving the Green's functions from quantities obtained using the wave functions method. However, finding the complete equivalence remained an open question. Beyond the mere challenge of showing that an equivalence exists, the fallout from such a demonstration is substantial in that one can readily check whether a calculation carried out using either method is corroborated by the other. This is indeed a much needed way to reduce errors, given that the problems tackled are way beyond reach of analytical tools. Moreover, for more realistic calculations one has to appeal to first-principles methods based mostly on the density functional theory(DFT). Most computer codes that implement the DFT make use of the wave function formalism, whereby the appropriate Khon-Sham equations are solved by expanding the electronic wave function on a basis of either localized or extended wave functions[26]. A few other codes rely on the Green's function formalism[27]. It is, therefore, convenient to implement transport calculations on top of the existing codes using either the WFM or AGF method, depending on the nature of the code, wave-function-based or Green's-functions-based, and knowing that the results will be identical in either case.

Besides the total transmission function, one major aspect of the whole framework is obtaining mode-resolved transmission probabilities, i.e., individual phonon contributions to the total transmission function [28, 29, 30, 31, 32, 33]. These are important for several reasons. For instance, such detailed information is required as input to other methods used at longer length scales, such as, non-gray Boltzmann equation [28]. One can also consider the interaction between phonons and light, where only the optical branches intervene, hence the need to disentangle their contribution from the rest of the phonon spectrum. The first attempt at deriving these mode-resolved transmissions was contained in the Fisher-Lee relation [34]. Several attempts have been made later on at obtaining such information, but most were misled by the invalid assumption that for a fixed frequency, the eigenvectors are orthogonal [3, 35]. We derive, in the current work, the correct and physically sound generalized Fisher-Lee relation for the mode-resolved transmission probabilities. Aside from the ease of implementation, the final expression which we arrived at has a clear physical interpretation. Previous derivations were obtained through rather involved mathematical machinery[1, 2, 36, 37, 38], and the final expressions, though amenable to our result, still require further manipulations.

In this work we provide a rigorous demonstration of the equivalence of the WFM and the AGF methods by showing that all the quantities calculated within one approach can be readily obtained from the other. These include, for instance, group velocities, transmission probabilities, and the individual phonon contributions to these probabilities. In order to illustrate the concordance, we compare various quantities computed within the two frameworks applied to a simple model. Moreover, we also argue that in order to describe correctly the problem at hand the usual region defined as the central part must include one more atomic plane from each lead.

This paper is organized as follows: In the next section our WFM formalism is introduced and the transmission function is derived within the formalism. In section 3 we provide the equivalence between the WFM and the AGF methods, and derive a generalized Fisher-Lee formula for the mode-resolved transmission amplitudes. In section 4 we apply our method to a simple model and make comparisons with earlier approaches. We finally summarize our main results and discussions in the conclusion.

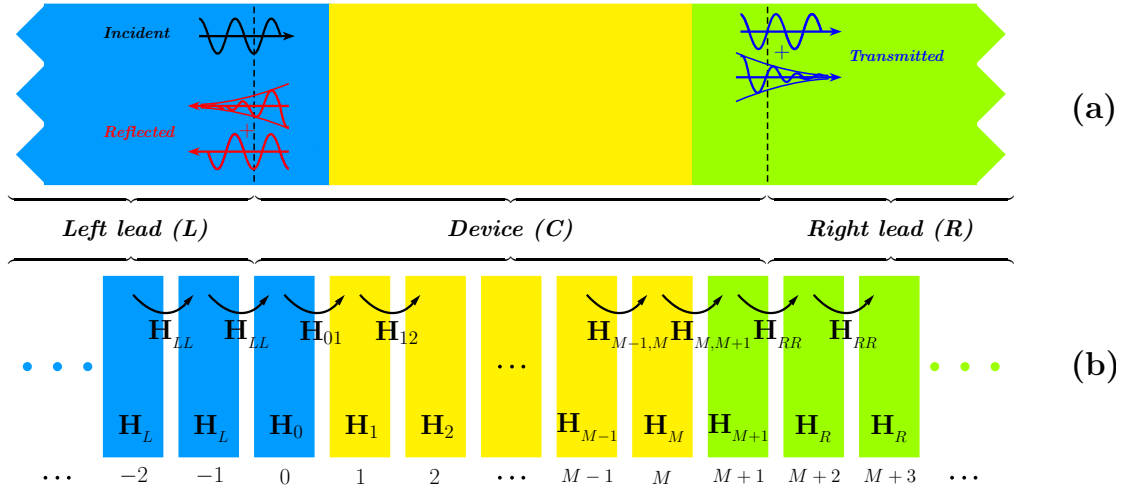


Figure 91: (a) Schematic representation and (b) Hamiltonian matrix divided into slices for the bi-dimensional nano-scale system studied within the WFM approach.

18.3 Wave function matching approach

Our main objective is to provide a full correspondence of the two approaches mentioned earlier. We think that the Green's functions formalism is the more robust, and as such we posit that all the quantities defined in this formalism must be expressible in terms of variables defined in the wave function matching approach. Indeed, the AGF method suffers no ambiguity in defining and computing the various quantities involved in the transport problem. The WFM approach, on the other hand, has been subject to debate as to its correct implementation in the present situation. One chief contention regards the treatment of the decaying modes that result from the presence of the scattering region. In the AGF approach the two main quantities are the retarded and advanced Green's functions. The first describes outgoing waves from the device, while the second is related to incoming waves. From the mathematical standpoint, in order to obtain these two quantities from the WFM formalism, we must keep only the evanescent modes that decay away from the device. There is also a compelling physical argument in favor of this choice: the evanescent modes always originate from the device, and can only decay away from it. Besides, if we were to reproduce the asymptotic behavior of the transmission probability, we must not include incoming evanescent modes since they would never reach the device if we inject them far away from it, as is clearly illustrated in Fig. 91.

We begin by reminding the formalism used in studying the scattering problem using matching of the wave functions at the interfaces of a bi-dimensional nano-scale system composed of semi-infinite right and left leads separated with a central device (see Figure 1-a). The leads are considered to be ideal with a periodic potential. The incoming propagating state from the left lead and normal to the interface (say in x direction) is scattered in the device and results in reflected and transmitted propagating as well as evanescent states. In the transverse direction (normal to x) the periodicity is assumed to hold. Consequently, the real space components in the transverse direction can be replaced by the normal component of the momentum ($\mathbf{k}_\perp$). As is usual we first derive quantities proper to the ideal left and right leads, to be used later to express the transmission probabilities.

The dynamical matrix of the whole system is given by the following infinite tridiagonal matrix:

$$\mathbf{H} = \begin{bmatrix} \ddots & \ddots & 0 & 0 & 0 & 0 & 0 \\ \ddots & \mathbf{H}_L & \mathbf{H}_{LL} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{H}_{LL}^\dagger & \mathbf{H}_L & \mathbf{H}_{LL} & 0 & 0 & 0 \\ 0 & 0 & \mathbf{H}_{LL}^\dagger & \mathbf{H}_C & \mathbf{H}_{RR} & 0 & 0 \\ 0 & 0 & 0 & \mathbf{H}_{RR}^\dagger & \mathbf{H}_R & \mathbf{H}_{RR} & 0 \\ 0 & 0 & 0 & 0 & \mathbf{H}_{RR}^\dagger & \mathbf{H}_R & \ddots \\ 0 & 0 & 0 & 0 & 0 & \ddots & \ddots \end{bmatrix}, \quad (18.1)$$

where the indices L and R stand for the left and right leads, respectively, as shown in Fig. 91(b). The part $\mathbf{H}_C$ represents the scattering region for which the Hamiltonian matrix is given as

$$\mathbf{H}_C = \begin{bmatrix} \mathbf{H}_0 & \mathbf{H}_{01} & 0 & 0 & 0 & 0 & 0 \\ \mathbf{H}_{01}^\dagger & \mathbf{H}_1 & \mathbf{H}_{12} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{H}_{12}^\dagger & \mathbf{H}_2 & \ddots & 0 & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & 0 & 0 & \ddots & \mathbf{H}_{M-1} & \mathbf{H}_{M-1,M} & 0 \\ 0 & 0 & 0 & 0 & \mathbf{H}_{M-1,M}^\dagger & \mathbf{H}_M & \mathbf{H}_{M,M+1} \\ 0 & 0 & 0 & 0 & 0 & \mathbf{H}_{M,M+1}^\dagger & \mathbf{H}_{M+1} \end{bmatrix}. \quad (18.2)$$

Here the slices 0 and $M + 1$ are included in the scattering region, although their respective Hamiltonians are seemingly similar to the left and right leads. Indeed, when either the force constants or the masses of the three regions are different, the Hamiltonians $\mathbf{H}_0$ and $\mathbf{H}_{M+1}$ are different than $\mathbf{H}_L$ and $\mathbf{H}_R$, respectively. Because of the assumption that any given slice is connected only to its immediate neighboring slices, the dynamical matrix takes on a tri-diagonal form. As such one atomic slice can be made up of one or more atomic planes, depending on whether interactions are first-neighbor only or are of a longer range (see Fig. 91(b)). The left and right leads are semi-infinite and extend all the way to $-\infty$ and to $+\infty$, respectively.

18.3.1 Ideal leads

From now on we restrict ourselves to studying the phonon transport, although all the ensuing discussions can be easily carried on to the electronic charge transport. In order to study transport of phonons from one side of the device to the other, we need to first study the motion of phonons in the ideal leads, that is in each lead as if it were infinite in both directions. The eigenmodes of these ideal leads will then be used to study how a phonon sent from the left lead scatters off the central part and results in transmitted and reflected phonons. We begin, therefore, by obtaining the eigenmodes of the ideal leads.

As shown in Figure 1-b, the studied system consists of three parts of slices normal to the direction of propagation and indexed from $-\infty$ to -1 for the left lead, from 0 to $M + 1$ for the scattering region, and from $M + 2$ to $+\infty$ for the right lead. Here we note that one slice from each lead is included as part of the scattering region. The slices are chosen such that only the interactions between neighboring slices are considered.

The equations of motion for the atoms in each ideal lead are given by

$$\omega^2 \mathbf{u}_{S,i} = \mathbf{H}_{SS}^\dagger \mathbf{u}_{S,i-1} + \mathbf{H}_S \mathbf{u}_{S,i} + \mathbf{H}_{SS} \mathbf{u}_{S,i+1}, \quad (18.3)$$

where $\mathbf{H}_S$ and $\mathbf{H}_{SS}$ ($S \equiv R$ for the right lead and $S \equiv L$ for the left lead) represent the on-slice and the inter-slice parts of the Hamiltonian, respectively. In this case i runs from $-\infty$ to $+\infty$. It is important to bear in mind that all the Hamiltonian and force constant matrices depend explicitly on $k_\perp$, i.e the momentum component which is transverse to the propagation direction x . We choose not to show this dependence for the sake of uncluttering the formulae.

They are $N \times N$ matrices, where N is the number of “vibrational degrees of freedom” per slice. In the electronic problem, N is the number of orbitals per slice used as a basis to represent the electronic wave functions. The component of a vibrational mode in slice i is represented by a vector $\mathbf{u}_{S,n}$ of size N . The equations of motion can be cast in matrix form as

$$\omega^2 \begin{pmatrix} \vdots \\ \mathbf{u}_{S,i-1} \\ \mathbf{u}_{S,i} \\ \mathbf{u}_{S,i+1} \\ \vdots \end{pmatrix} = \begin{bmatrix} \ddots & \ddots & 0 & 0 & 0 \\ \ddots & \mathbf{H}_S & \mathbf{H}_{SS} & 0 & 0 \\ 0 & \mathbf{H}_{SS}^\dagger & \mathbf{H}_S & \mathbf{H}_{SS} & 0 \\ 0 & 0 & \mathbf{H}_{SS}^\dagger & \mathbf{H}_S & \ddots \\ 0 & 0 & 0 & \ddots & \ddots \end{bmatrix} \begin{pmatrix} \vdots \\ \mathbf{u}_{S,i-1} \\ \mathbf{u}_{S,i} \\ \mathbf{u}_{S,i+1} \\ \vdots \end{pmatrix}, \quad (18.4)$$

For the ideal right lead, where the same unit cell or slice is repeated indefinitely in the x direction of space, we make use of Bloch’s theorem and write $\mathbf{u}_{R,i} = e^{iq_x a_R} \mathbf{u}_{R,i-1}$, where q_x and a_R are the wave vector and the size of the unit cell in the x direction, respectively. It is important to stress, at this point, that q_x can be either a purely real or a complex, having both real and imaginary parts, number. In the latter case the use of the expression “Bloch’s theorem” is a misnomer given that the corresponding modes are not propagating. Defining $\lambda_R = e^{iq_x a_R}$, we have $\mathbf{u}_{R,i} = \lambda_R \mathbf{u}_{R,i-1}$. We can now rewrite the equations of motion (18.3) as

$$\omega^2 \mathbf{u}_{R,i} = \mathbf{H}_{RR}^\dagger \lambda_R^{-1} \mathbf{u}_{R,i} + \mathbf{H}_R \mathbf{u}_{R,i} + \mathbf{H}_{RR} \lambda_R \mathbf{u}_{R,i}. \quad (18.5)$$

That is

$$(\omega^2 - \mathbf{H}_R - \lambda_R^{-1} \mathbf{H}_{RR}^\dagger - \lambda_R \mathbf{H}_{RR}) \mathbf{u}_{R,i} = 0. \quad (18.6)$$

This is a quadratic eigenvalue problem which can be recast as a generalized eigenvalue problem by defining $\lambda_R^{-1} \mathbf{u}_R$ as a new “independent” variable. The new generalized eigenvalue problem then reads

$$\begin{bmatrix} \omega^2 \mathbf{I} - \mathbf{H}_R & -\mathbf{H}_{RR}^\dagger \\ \mathbf{I} & 0 \end{bmatrix} \begin{pmatrix} \mathbf{u}_R \\ \lambda_R^{-1} \mathbf{u}_R \end{pmatrix} = \lambda_R \begin{bmatrix} \mathbf{H}_{RR} & 0 \\ 0 & \mathbf{I} \end{bmatrix} \begin{pmatrix} \mathbf{u}_R \\ \lambda_R^{-1} \mathbf{u}_R \end{pmatrix}. \quad (18.7)$$

Similarly, for the ideal periodic left lead, we have chosen to write, using Bloch’s theorem, $\mathbf{u}_{L,i-1} = \tilde{\lambda}_L \mathbf{u}_{L,i}$, where $\tilde{\lambda}_L = e^{-iq_x a_L}$ and a_L is the size of the left unit cell. Let us note that the choice of $\tilde{\lambda}_L$ as the Bloch factor being different from that of the right lead is only motivated by the aim to obtain results that are homogeneous with those obtained through the AGF formalism, as it can be seen later. We can now rewrite Eq. (18.3) as

$$\omega^2 \mathbf{u}_{L,n} = \mathbf{H}_{LL}^\dagger \tilde{\lambda}_L \mathbf{u}_{L,i} + \mathbf{H}_L \mathbf{u}_{L,i} + \mathbf{H}_{LL} \tilde{\lambda}_L^{-1} \mathbf{u}_{L,i}, \quad (18.8)$$

and then

$$(\omega^2 - \mathbf{H}_L - \tilde{\lambda}_L \mathbf{H}_{LL}^\dagger - \tilde{\lambda}_L^{-1} \mathbf{H}_{LL}) \mathbf{u}_{L,n} = 0. \quad (18.9)$$

By defining the new “independent” variable $\tilde{\lambda}_L^{-1} \mathbf{u}_L$, the quadratic eigenvalue problem becomes a generalized eigenvalue problem and reads

$$\begin{bmatrix} \omega^2 \mathbf{I} - \mathbf{H}_L & -\mathbf{H}_{LL}^\dagger \\ \mathbf{I} & 0 \end{bmatrix} \begin{pmatrix} \mathbf{u}_L \\ \tilde{\lambda}_L^{-1} \mathbf{u}_L \end{pmatrix} = \tilde{\lambda}_L \begin{bmatrix} \mathbf{H}_{LL}^\dagger & 0 \\ 0 & \mathbf{I} \end{bmatrix} \begin{pmatrix} \mathbf{u}_L \\ \tilde{\lambda}_L^{-1} \mathbf{u}_L \end{pmatrix}. \quad (18.10)$$

From (18.7) and (18.10), for each value of ω , we obtain $2N$ eigenvalues and eigenvectors for each ideal lead. The eigenmodes are split into two sets of size N each:

1. Right-moving propagating ($|\lambda_{R,n}| = 1$ or $|\tilde{\lambda}_{L,n}| = 1$ with $v_{R/L,n} > 0$) and evanescent ($|\lambda_{R,n}| < 1$ or $|\tilde{\lambda}_{L,n}| > 1$) modes, where $v_{R/L,n}$ is the group velocity of the n^{th} mode along the x -direction.
2. Left-moving propagating ($|\lambda_{R,n}| = 1$ or $|\tilde{\lambda}_{L,n}| = 1$ with $v_{R/L,n} < 0$) and evanescent ($|\lambda_{R,n}| > 1$ or $|\tilde{\lambda}_{L,n}| < 1$) modes.

In view of attaching later the leads to the scattering region and having in sight the equivalence with the AGF formalism, we depart from the usual treatments [1, 12, 13, 17] and discard the evanescent modes that decay towards the central region. As mentioned above, the physical justification resides in that evanescent modes are localized around the “defect” and always decay away from it. It does not make sense to include evanescent modes moving towards the “defect”. We also know for a fact that this is indeed the case when using the Green’s functions formalism, and by this clever trick, we will show how to move seamlessly between the two formalisms. And in keeping with the terminology used in the Green’s function formalism, we define quantities built out of these selected evanescent modes and propagating modes moving either towards the defect or away from it. We call the latter “retarded” and the first “advanced”, and in order to differentiate them we append a superscript “a” to the advanced quantities. We note that in what follows the advanced quantities for the right lead are defined only for the sake of completeness. So long as the incoming phonon is injected from the left lead, the boundary conditions are imposed such that there are only transmitted waves in the right lead. These waves are represented by the retarded quantities. As such the advanced quantities are not needed in the right lead. However, if one were to inject phonons from the right lead, then the advanced quantities in the left lead would be the ones to be unused. We begin by reminding the definition of the group velocity:

$$v_S = \frac{\partial \omega}{\partial q_x} = \frac{1}{2\omega} \frac{\partial \omega^2}{\partial q_x}. \quad (18.11)$$

From the equations of motion in the leads and using the Hellmann-Feynman theorem [39], we obtain the following expressions for the various group velocities defined in the right and left leads:

$$v_{g,R} = \frac{ia_R}{2\omega} \left[\mathbf{u}_R^\dagger \mathbf{H}_{RR} \mathbf{u}_R \lambda_R - \mathbf{u}_R^\dagger \mathbf{H}_{RR}^\dagger \mathbf{u}_R \lambda_R^\dagger \right], \quad (18.12)$$

$$v_{g,R}^a = \frac{ia_L}{2\omega} \left[\mathbf{u}_R^{a\dagger} \mathbf{H}_{RR} \mathbf{u}_R^a \lambda_R^a - \mathbf{u}_R^{a\dagger} \mathbf{H}_{RR}^\dagger \mathbf{u}_R^a \lambda_R^{a\dagger} \right], \quad (18.13)$$

$$v_{g,L} = -\frac{ia_L}{2\omega} \left[\mathbf{u}_L^\dagger \mathbf{H}_{LL}^\dagger \mathbf{u}_L \tilde{\lambda}_L - \mathbf{u}_L^\dagger \mathbf{H}_{LL} \mathbf{u}_L \tilde{\lambda}_L^\dagger \right], \quad (18.14)$$

$$v_{g,L}^a = -\frac{ia_L}{2\omega} \left[\mathbf{u}_L^{a\dagger} \mathbf{H}_{LL}^\dagger \mathbf{u}_L^a \tilde{\lambda}_L^a - \mathbf{u}_L^{a\dagger} \mathbf{H}_{LL} \mathbf{u}_L^a \tilde{\lambda}_L^{a\dagger} \right], \quad (18.15)$$

where a_L and a_R are the unit cell size in the left and right leads, respectively. We also define new quantities which are constructed out of the group velocities, with the coefficients $\frac{a_{L/R}}{2\omega}$ omitted for convenience when we will be comparing the WFM with the AGF formalisms:

$$\mathbf{V}_R = i \left[\mathbf{C}_R^\dagger \mathbf{H}_{RR} \mathbf{C}_R \boldsymbol{\Lambda}_R - \boldsymbol{\Lambda}_R^\dagger \mathbf{C}_R^\dagger \mathbf{H}_{RR}^\dagger \mathbf{C}_R \right], \quad (18.16)$$

$$\mathbf{V}_R^a = i \left[\mathbf{C}_R^{a\dagger} \mathbf{H}_{RR} \mathbf{C}_R^a \boldsymbol{\Lambda}_R^a - \boldsymbol{\Lambda}_R^{a\dagger} \mathbf{C}_R^{a\dagger} \mathbf{H}_{RR}^\dagger \mathbf{C}_R^a \right], \quad (18.17)$$

$$\mathbf{V}_L = -i \left[\mathbf{C}_L^\dagger \mathbf{H}_{LL}^\dagger \mathbf{C}_L \mathbf{\Lambda}_L - \mathbf{\Lambda}_L^\dagger \mathbf{C}_L^\dagger \mathbf{H}_{LL} \mathbf{C}_L \right], \quad (18.18)$$

$$\mathbf{V}_L^a = -i \left[\mathbf{C}_L^{a\dagger} \mathbf{H}_{LL}^\dagger \mathbf{C}_L^a \mathbf{\Lambda}_L^a - \mathbf{\Lambda}_L^{a\dagger} \mathbf{C}_L^{a\dagger} \mathbf{H}_{LL} \mathbf{C}_L^a \right]. \quad (18.19)$$

These matrices are diagonal, with the diagonal elements being the group velocities for the propagating modes, and vanishing for the decaying modes.

We now introduce the matrices built from the eigenvectors and eigenmodes of the ideal leads that will be used in the WFM formalism to obtain the vibration modes in the central region. For the right lead, we define

$$\mathbf{C}_R = (\mathbf{u}_{R,1}, \mathbf{u}_{R,2}, \dots, \mathbf{u}_{R,n}), \quad (18.20)$$

$$\mathbf{C}_R^a = (\mathbf{u}_{R,1}^a, \mathbf{u}_{R,2}^a, \dots, \mathbf{u}_{R,n}^a), \quad (18.21)$$

$$\mathbf{\Lambda}_R = \text{diag}(\lambda_{R,1}, \lambda_{R,2}, \dots, \lambda_{R,n}), \quad (18.22)$$

and

$$\mathbf{\Lambda}_R^a = \text{diag}(\lambda_{R,1}^a, \lambda_{R,2}^a, \dots, \lambda_{R,n}^a), \quad (18.23)$$

and we define similar quantities for the left lead with λ_R replaced by $\tilde{\lambda}_L$, and $\mathbf{u}_R$ by $\mathbf{u}_L$.

18.3.2 The scattering region

We now proceed to writing the effective Hamiltonian of the scattering region, replacing the effects of the ideal leads by “self-energy” terms to be added to the original Hamiltonian. Writing the equation of motion for slice 0,

$$-\mathbf{H}_{LL}^\dagger \mathbf{u}_{-1} + (\omega^2 \mathbf{I} - \mathbf{H}_0) \mathbf{u}_0 - \mathbf{H}_{01} \mathbf{u}_1 = 0,$$

and moving the first term to the right-hand side, we have

$$(\omega^2 \mathbf{I} - \mathbf{H}_0) \mathbf{u}_0 - \mathbf{H}_{01} \mathbf{u}_1 = \mathbf{H}_{LL}^\dagger \mathbf{u}_{-1}.$$

In a similar fashion, we find for slice $M + 1$,

$$(\omega^2 \mathbf{I} - \mathbf{H}_{M+1}) \mathbf{u}_{M+1} - \mathbf{H}_{M,M+1} \mathbf{u}_M = \mathbf{H}_{RR} \mathbf{u}_{M+2}.$$

The equations of motion for the central part are then given in matrix form as

$$[\omega^2 \mathbf{I} - \mathbf{H}_C] \begin{bmatrix} \mathbf{u}_0 \\ \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_M \\ \mathbf{u}_{M+1} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{LL}^\dagger \mathbf{u}_{-1} \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \mathbf{H}_{RR} \mathbf{u}_{M+2} \end{bmatrix}, \quad (18.24)$$

with the full matrix form of $\mathbf{H}_C$ given in Eq. 18.2.

This means that if we know or impose the values of $\mathbf{u}_{-1}$ and $\mathbf{u}_{M+2}$, we can invert this matrix equation to obtain all the displacement $\mathbf{u}_i$, with $i = 0, 1, \dots, M + 1$.

To accomplish this, we make use of Bloch's theorem which holds in the leads. We write

$$\begin{cases} \mathbf{u}_{-1} = \tilde{\lambda}_L \mathbf{u}_0 & \text{with } \tilde{\lambda} = e^{-iq_x a_L}, \\ \mathbf{u}_{M+2} = \lambda_R \mathbf{u}_{M+1} & \text{with } \lambda_R = e^{iq_x a_R}, \end{cases} \quad (18.25)$$

Imposing an incoming mode from the left lead, with an amplitude $\mathbf{a}_{in}$, the displacement $\mathbf{u}_0$ will be a linear combination of the incoming and reflected waves. In the right lead we impose that only the transmitted wave exist. We have then

$$\begin{cases} \mathbf{u}_0 = \mathbf{C}_L^a \mathbf{a}_{in} + \mathbf{C}_L \mathbf{a}_{ref}, \\ \mathbf{u}_{M+1} = \mathbf{C}_R \mathbf{a}_{tr} \end{cases}, \quad (18.26)$$

where the vector $\mathbf{a}_{in}$ is the amplitude of the right-going incoming wave, $\mathbf{a}_{ref}$ is the amplitude of the reflected wave, and $\mathbf{a}_{tr}$ is the amplitude of the transmitted wave. The latter two vectors are to be calculated given the boundary conditions. The first boundary condition is given by setting the vector $\mathbf{a}_{in}$. This is usually taken as a unit vector, whose components are all zero, except for the component, corresponding to one mode, which is set to unity. The second "boundary condition" imposed on the wave function is that the left-going part in the right lead vanishes.

Making use of Bloch's theorem, we have

$$\begin{cases} \mathbf{u}_{-1} = \mathbf{C}_L^a \Lambda_L^a \mathbf{a}_{in} + \mathbf{C}_L \Lambda_L \mathbf{a}_{ref}, \\ \mathbf{u}_{M+2} = \mathbf{C}_R \Lambda_R \mathbf{a}_{tr} \end{cases}. \quad (18.27)$$

We introduce the pseudo-inverse matrices $\tilde{\mathbf{C}}_L$, $\tilde{\mathbf{C}}_L^a$, and $\tilde{\mathbf{C}}_R$ such that $\tilde{\mathbf{C}}_L \mathbf{C}_L = \mathbf{I}$, $\tilde{\mathbf{C}}_L^a \mathbf{C}_L^a = \mathbf{I}$, and $\tilde{\mathbf{C}}_R \mathbf{C}_R = \mathbf{I}$, where $\mathbf{I}$ is the $n \times n$ identity matrix. Using these matrices, the wave function component $\mathbf{u}_{-1}$ is expressed as

$$\mathbf{u}_{-1} = \mathbf{C}_L^a \Lambda_L^a \tilde{\mathbf{C}}_L^a \mathbf{C}_L^a \mathbf{a}_{in} + \mathbf{C}_L \Lambda_L \tilde{\mathbf{C}}_L \mathbf{C}_L \mathbf{a}_{ref},$$

and given that $\mathbf{C}_L \mathbf{a}_{ref} = \mathbf{u}_0 - \mathbf{C}_L^a \mathbf{a}_{in}$, we have

$$\mathbf{u}_{-1} = \mathbf{C}_L^a \Lambda_L^a \tilde{\mathbf{C}}_L^a \mathbf{C}_L^a \mathbf{a}_{in} + \mathbf{C}_L \Lambda_L \tilde{\mathbf{C}}_L (\mathbf{u}_0 - \mathbf{C}_L^a \mathbf{a}_{in}).$$

Introducing the Bloch matrices

$$\mathbf{B}_L = \mathbf{C}_L \Lambda_L \tilde{\mathbf{C}}_L \quad (18.28)$$

and

$$\mathbf{B}_L^a = \mathbf{C}_L^a \Lambda_L^a \tilde{\mathbf{C}}_L^a, \quad (18.29)$$

we obtain

$$\mathbf{u}_{-1} = \mathbf{B}_L^a \mathbf{C}_L^a \mathbf{a}_{in} + \mathbf{B}_L \mathbf{u}_0 - \mathbf{B}_L \mathbf{C}_L^a \mathbf{a}_{in}, \quad (18.30)$$

and finally

$$\mathbf{u}_{-1} = (\mathbf{B}_L^a - \mathbf{B}_L) \mathbf{C}_L^a \mathbf{a}_{in} + \mathbf{B}_L \mathbf{u}_0. \quad (18.31)$$

Similarly for the right lead, we introduce the Bloch matrix

$$\mathbf{B}_R = \mathbf{C}_R \Lambda_R \tilde{\mathbf{C}}_R, \quad (18.32)$$

and write

$$\mathbf{u}_{M+2} = \mathbf{C}_R \Lambda_R \tilde{\mathbf{C}}_R \mathbf{C}_R \mathbf{a}_{tr} = \mathbf{B}_R \mathbf{C}_R \mathbf{a}_{tr} = \mathbf{B}_R \mathbf{u}_{M+1}. \quad (18.33)$$

Using the above, we finally obtain the following system of equations for the central part

$$(\omega^2 \mathbf{I} - \mathbf{H}_{eff}^W) \begin{bmatrix} \mathbf{u}_0 \\ \vdots \\ \mathbf{u}_{M+1} \end{bmatrix} = \begin{bmatrix} \mathbf{q}_{inj} \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

with

$$\mathbf{H}_{eff}^W = \begin{bmatrix} \mathbf{H}_{LS} & \mathbf{H}_{01} & 0 & 0 & 0 & 0 & 0 \\ \mathbf{H}_{01}^\dagger & \mathbf{H}_1 & \mathbf{H}_{12} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{H}_{12}^\dagger & \mathbf{H}_2 & \ddots & 0 & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & 0 & 0 & \ddots & \mathbf{H}_{M-1} & \mathbf{H}_{M-1,M} & 0 \\ 0 & 0 & 0 & 0 & \mathbf{H}_{M-1,M}^\dagger & \mathbf{H}_M & \mathbf{H}_{M,M+1} \\ 0 & 0 & 0 & 0 & 0 & \mathbf{H}_{M,M+1}^\dagger & \mathbf{H}_{RS} \end{bmatrix}, \quad (18.34)$$

$$\mathbf{H}_{LS} = \mathbf{H}_0 + \mathbf{H}_{LL}^\dagger \mathbf{B}_L,$$

$$\mathbf{H}_{RS} = \mathbf{H}_{M+1} + \mathbf{H}_{RR} \mathbf{B}_R,$$

and

$$\mathbf{q}_{inj} = \mathbf{H}_{LL}^\dagger (\mathbf{B}_L^a - \mathbf{B}_L) \mathbf{C}_L^a \mathbf{a}_{in}.$$

The quantities

$$\Sigma_R^W = \mathbf{H}_{RR} \mathbf{B}_R \quad (18.35)$$

and

$$\Sigma_L^W = \mathbf{H}_{LL}^\dagger \mathbf{B}_L \quad (18.36)$$

are known, respectively, as the self-energies for the right and left leads, and are formally introduced later in the AGF formalism. These are the defining quantities that will allow us to establish a connection with the AGF method and, consequently, the sought after equivalence. Defining $\mathbf{G}_C^W = [\omega^2 \mathbf{I} - \mathbf{H}_{eff}^W]^{-1}$, the above system can be inverted as

$$\begin{bmatrix} \mathbf{u}_0 \\ \vdots \\ \vdots \\ \mathbf{u}_{M+1} \end{bmatrix} = \mathbf{G}_C^W \begin{bmatrix} \mathbf{q}_{inj} \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (18.37)$$

We remark at this point that only the three Bloch matrices ($\mathbf{B}_L$, $\mathbf{B}_L^a$, and $\mathbf{B}_R$) we defined above are needed to carry out all the subsequent analyses and computations. As we will see below, these matrices can all be obtained from the AGF approach. However, it is customary in prior literature[2, 30] to define eight such matrices, with four of them having no connection at all to the AGF formalism.

18.3.3 Transmission

For a particular incoming mode m given by the m^{th} column of the matrix $\mathbf{C}_L^a$, the expression $\mathbf{u}_{M+1} = \mathbf{C}_R \mathbf{a}_{tr}$ reads:

$$\mathbf{C}_R \boldsymbol{\tau}_m = \mathbf{u}_{M+1,m}. \quad (18.38)$$

We define the matrix $\boldsymbol{\tau}$ as composed of the columns $\tau_m (m = 1, 2, \dots, N)$, where N is the number of modes:

$$\boldsymbol{\tau} = (\boldsymbol{\tau}_1 \boldsymbol{\tau}_2 \dots \boldsymbol{\tau}_N).$$

We generalize the earlier relation and obtain

$$\mathbf{C}_R \boldsymbol{\tau} = (\mathbf{u}_{M+1,1} \mathbf{u}_{M+1,2} \dots \mathbf{u}_{M+1,N}).$$

From Eq. 18.37, we have

$$\mathbf{u}_{M+1,m} = \mathbf{G}_C^W(M+1, 0) \mathbf{q}_{inj}. \quad (18.39)$$

It is a common practice to fix $\mathbf{a}_{in}$ as having a unit component only along the m^{th} mode, with all the other components vanishing. Instead of working with the vectors $\mathbf{a}_{in}$ and $\mathbf{q}_{inj}$, we define the respective matrices $\mathbf{A}_{in}$ and $\mathbf{Q}_{inj}$. Here $\mathbf{A}_{in}$ is just the $n \times n$ identity matrix, and as a consequence, $\mathbf{Q}_{inj}$ is given as

$$\mathbf{Q}_{inj} = \mathbf{H}_{LL}^\dagger (\mathbf{B}_L^a - \mathbf{B}_L) \mathbf{C}_L^a \mathbf{A}_{in} = \mathbf{H}_{LL}^\dagger (\mathbf{B}_L^a - \mathbf{B}_L) \mathbf{C}_L^a, \quad (18.40)$$

which follows from $\mathbf{A}_{in}$ being the identity matrix. We then obtain

$$\tau = \mathbf{C}_R^{-1} \mathbf{G}_C^W (M+1, 0) \mathbf{Q}_{inj}.$$

The transmission amplitude is therefore found by normalizing with respect to the current and reads

$$\mathbf{t} = \mathbf{V}_R^{1/2} \mathbf{C}_R^{-1} \mathbf{G}_{C(M+1,0)}^W \mathbf{Q}_{inj} (\mathbf{V}_L^a)^{-1/2}.$$

Replacing $\mathbf{Q}_{inj}$, defined in Eq. (eq. 18.40), with its alternative expression, $\mathbf{Q}_{inj} = i (\mathbf{C}_L^{a\dagger})^{-1} \mathbf{V}_L^a$ (cf. 18.7), we arrive at the final expression for the transmission amplitude:

$$\mathbf{t} = i \mathbf{V}_R^{1/2} \mathbf{C}_R^{-1} \mathbf{G}_{C(M+1,0)}^W (\mathbf{C}_L^{a\dagger})^{-1} (\mathbf{V}_L^a)^{1/2}. \quad (18.41)$$

This is similar to the expression arrived at by Komyakov *et al.*[2], starting from the AGF formalism. The difference lies in the way they define the Bloch matrix describing the incoming wave. As a result, in their case, even the evanescent modes yield non-vanishing elements in $\mathbf{Q}_{inj}$, which they then proceed to eliminate “by hand”, by considering the group velocities. The evanescent modes, having a vanishing group velocity, the corresponding entries in $\mathbf{Q}_{inj}$ are set to zero. However, in our case the evanescent modes are eliminated “naturally” thanks to the clever choice of the modes entering our definition of the Bloch matrix $\mathbf{B}_L^a$. By naturally we mean that there is no need to scan over the modes in search of the decaying ones and then zero out the corresponding entries in $\mathbf{Q}_{inj}$. We must indicate that Komyakov *et al.*[2] define a matrix $\mathbf{Q}_0$ (their eq. 14), which when multiplied by their $\mathbf{U}_L(+)$ (their eq. 55) on the right, yields a matrix $\mathbf{Q}_{inj}^K$, to which we compare our $\mathbf{Q}_{inj}$.

18.4 AGF formalism

18.4.1 A brief review of the conventional AGF formalism

We now provide a brief overview of the well-known AGF method with the aim to demonstrate its equivalence to the WFM formalism described earlier. And more importantly, once the complete equivalence is shown, we will exploit it to extract the individual, i.e. mode-resolved transmission amplitudes. For a more detailed treatment of the AGF method, we refer the reader to prior literature [40, 41].

In the single particle formalism the retarded Green’s function $\mathbf{G}$, also known as the resolvent, is defined as the operator inverse of the Hamiltonian $\mathbf{H}$ (given in Eq. 18.1):

$$(\omega^2 \mathbf{I} + i\eta - \mathbf{H}) \mathbf{G}(\omega, \mathbf{k}_\perp) = \mathbf{I},$$

where η is a positive infinitesimal. A similar equation with $-i\eta$ (instead of $+i\eta$) defines the advanced Green’s function $\mathbf{G}^a$, which can also be viewed as the conjugate of the retarded function $\mathbf{G}$. The element $\mathbf{G}_{ij}$ represents the response of slice j to a small vibration of slice i .

It is therefore related to the transmission probability between slices i and j . In matrix form, the above equation reads

$$\begin{bmatrix} \omega^2 \mathbf{I} - \mathbf{H}_L & -\mathbf{H}_{LL} & 0 \\ -\mathbf{H}_{LL}^\dagger & \omega^2 \mathbf{I} - \mathbf{H}_C & -\mathbf{H}_{RR} \\ 0 & -\mathbf{H}_{RR}^\dagger & \omega^2 \mathbf{I} - \mathbf{H}_R \end{bmatrix} \begin{bmatrix} \mathbf{G}_L & \mathbf{G}_{LC} & \mathbf{G}_{LR} \\ \mathbf{G}_{CL} & \mathbf{G}_C & \mathbf{G}_{CR} \\ \mathbf{G}_{RL} & \mathbf{G}_{RC} & \mathbf{G}_R \end{bmatrix} = \begin{bmatrix} \mathbf{I} & 0 & 0 \\ 0 & \mathbf{I} & 0 \\ 0 & 0 & \mathbf{I} \end{bmatrix}.$$

We can then solve for the central part alone and obtain

$$\mathbf{G}_C = [\omega^2 \mathbf{I} - \mathbf{H}_C - \boldsymbol{\Sigma}_L - \boldsymbol{\Sigma}_R]^{-1},$$

where $\boldsymbol{\Sigma}_L$ and $\boldsymbol{\Sigma}_R$ are the self-energies containing the effects due to connecting the central region to the left and right leads, respectively. They are given by

$$\boldsymbol{\Sigma}_L = \mathbf{H}_{LL}^\dagger \mathbf{g}_L \mathbf{H}_{LL}, \quad (18.42)$$

and

$$\boldsymbol{\Sigma}_R = \mathbf{H}_{RR} \mathbf{g}_R \mathbf{H}_{RR}^\dagger, \quad (18.43)$$

where $\mathbf{g}_L$ and $\mathbf{g}_R$ are the surface Green's functions of the left and right leads, respectively [40, 41]. In this work they are computed using a numerically efficient algorithm devised by Lopez-Sancho [42, 43, 24]. The central part is as such represented by an effective Hamiltonian

$$\mathbf{H}_{eff}^G = \begin{bmatrix} \mathbf{H}_0 + \boldsymbol{\Sigma}_L & \mathbf{H}_{01} & 0 & 0 & 0 & 0 & 0 \\ \mathbf{H}_{01}^\dagger & \mathbf{H}_1 & \mathbf{H}_{12} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{H}_{12}^\dagger & \mathbf{H}_2 & \ddots & 0 & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & 0 & 0 & \ddots & \mathbf{H}_{M-1} & \mathbf{H}_{M-1,M} & 0 \\ 0 & 0 & 0 & 0 & \mathbf{H}_{M-1,M}^\dagger & \mathbf{H}_M & \mathbf{H}_{M,M+1} \\ 0 & 0 & 0 & 0 & 0 & \mathbf{H}_{M,M+1}^\dagger & \mathbf{H}_{M+1} + \boldsymbol{\Sigma}_R \end{bmatrix}, \quad (18.44)$$

such that

$$(\omega^2 \mathbf{I} - \mathbf{H}_{eff}^G) \mathbf{G}_C = \mathbf{I}.$$

The device Green's functions $\mathbf{G}_C$ and $\mathbf{G}_C^a$ are then used to compute the total transmission via the Caroli formula [22]:

$$T = \text{trace}(\mathbf{G}_C \boldsymbol{\Gamma}_L \mathbf{G}_C^a \boldsymbol{\Gamma}_R), \quad (18.45)$$

where

$$\boldsymbol{\Gamma}_{R(L)} = i(\boldsymbol{\Sigma}_{R(L)} - \boldsymbol{\Sigma}_{R(L)}^\dagger).$$

18.4.2 Equivalence of WFM and AGF

The effective Hamiltonian $\mathbf{H}_{eff}^G$ defined in Eq. 18.44 is to be compared to the effective Hamiltonian $\mathbf{H}_{eff}^W$ obtained in the WFM formalism (Eq. 18.34). The two must indeed be identical. As a result, we can now proceed to identifying the elements appearing in the two expressions of the effective Hamiltonian. The right and left self-energies pertaining to the effects of the right and left leads, respectively, in the WFM formalism are given in Eqs. 18.35 and 18.36. The equivalent quantities given in the AGF formalism are written in Eqs. 18.42 and 18.43. Equating these quantities term by term leads to the following equivalences

$$\mathbf{B}_R = \mathbf{g}_R \mathbf{H}_{RR}^\dagger, \quad (18.46)$$

$$\mathbf{B}_L = \mathbf{g}_L \mathbf{H}_{LL}. \quad (18.47)$$

WFM	AGF
H_L, H_R	H_L, H_R
B_L, B_L^a, B_R	g_L, g_R
Σ_L^W, Σ_R^W	Σ_L, Σ_R
H_C^W	H_C
$(H_C^W)^{-1} \equiv G_C^W$	$H_C^{-1} \equiv G_C$
T	T

Table 3: A comparative list of the quantities defined in the two formalisms, WFM and AGF, and in the same order of their calculations.

The surface Green's functions are then related to the Bloch matrices as

$$\mathbf{g}_R = \mathbf{B}_R (\mathbf{H}_{RR}^\dagger)^{-1}, \quad (18.48)$$

$$\mathbf{g}_L = \mathbf{B}_L (\mathbf{H}_{LL})^{-1}, \quad (18.49)$$

for the right and left leads, respectively. Similarly, for the advanced quantities, the self-energies are given in the WFM formalism as

$$\Sigma_R^{a(W)} = \mathbf{H}_{RR} \mathbf{B}_R^a, \quad (18.50)$$

$$\Sigma_L^{a(W)} = \mathbf{H}_{LL}^\dagger \mathbf{B}_L^a, \quad (18.51)$$

and in the AGF formalism as

$$\Sigma_R^a = \mathbf{H}_{RR} \mathbf{g}_R^a \mathbf{H}_{RR}^\dagger, \quad (18.52)$$

$$\Sigma_L^a = \mathbf{H}_{LL}^\dagger \mathbf{g}_L^a \mathbf{H}_{LL}. \quad (18.53)$$

Equating these quantities leads to the following equivalences

$$\mathbf{B}_R^a = \mathbf{g}_R^a \mathbf{H}_{RR}^\dagger, \quad (18.54)$$

$$\mathbf{B}_L^a = \mathbf{g}_L^a \mathbf{H}_{LL}. \quad (18.55)$$

The advanced surface GFs are then related to the advanced Bloch matrices as

$$\mathbf{g}_R^a = \mathbf{B}_R^a (\mathbf{H}_{RR}^\dagger)^{-1}, \quad (18.56)$$

$$\mathbf{g}_L^a = \mathbf{B}_L^a (\mathbf{H}_{LL})^{-1}. \quad (18.57)$$

Finally, we arrive at the following expressions for the normalized group velocity matrices needed in the expression of the transmission amplitudes:

$$\begin{aligned} \mathbf{V}_R &= \mathbf{C}_R^\dagger \Gamma_R \mathbf{C}_R, \\ \mathbf{V}_L &= \mathbf{C}_L^{a\dagger} \Gamma_L \mathbf{C}_L^a. \end{aligned} \quad (18.58)$$

with the details given in 18.6. These are identical to the expressions found by Komyakov *et al.*[12] who used a rather complicated route. We note that these are the key ingredients to resolving, in the next sub-section, the total transmission function in terms of individual phonon contributions. But, before moving on, we think it is important to keep track of the most important quantities introduced so far within the two approaches. These are shown in Table 3 and in the order that they are computed in each method.

18.4.3 Mode-resolved transmissions

As mentioned earlier, a number of attempts were made aiming at deriving expressions [2, 28, 29, 30, 32, 37] for the mode-resolved transmission probabilities. However, these reports were based on involved mathematical derivations and it is not yet obvious whether all agree in the end. In our case, we will exploit the identities in Eq. 18.58 for a straightforward derivation of the desired relation. One advantage of our derivation is that it yields to a straightforward physical interpretation.

Indeed, we can obtain Γ_R and Γ_L in terms of the normalized group velocity matrices from Eq. 18.58 as:

$$\Gamma_R = (\mathbf{C}_R^\dagger)^{-1} \mathbf{V}_R (\mathbf{C}_R)^{-1} \quad (18.59)$$

and

$$\Gamma_L = (\mathbf{C}_L^{a\dagger})^{-1} \mathbf{V}_L^a (\mathbf{C}_L^a)^{-1}. \quad (18.60)$$

Next, we define the following operators:

$$\Pi_R = (\mathbf{C}_R^\dagger)^{-1} (\mathbf{V}_R)^{1/2} \quad (18.61)$$

and

$$\Pi_L = (\mathbf{C}_L^{a\dagger})^{-1} (\mathbf{V}_L^a)^{1/2}, \quad (18.62)$$

such that

$$\Pi_R \Pi_R^\dagger = (\mathbf{C}_R^\dagger)^{-1} \mathbf{V}_R (\mathbf{C}_R)^{-1} = \Gamma_R,$$

and

$$\Pi_L \Pi_L^\dagger = (\mathbf{C}_L^{a\dagger})^{-1} \mathbf{V}_L^a (\mathbf{C}_L^a)^{-1} = \Gamma_L.$$

The Caroli formula (Eq. 18.45) can now be recast as

$$\begin{aligned} T &= \text{trace}(\mathbf{G}^r \Gamma_L \mathbf{G}^a \Gamma_R) \\ &= \text{trace}(\mathbf{G}^r \Pi_L \Pi_L^\dagger \mathbf{G}^a \Pi_R \Pi_R^\dagger) \\ &= \text{trace}(\Pi_R^\dagger \mathbf{G}^r \Pi_L \Pi_L^\dagger \mathbf{G}^a \Pi_R) \\ &= \text{trace}(\mathbf{t} \mathbf{t}^\dagger), \end{aligned} \quad (18.63)$$

where we made use of the cyclic property of the trace. We can identify the matrices appearing in the last line of the equality as

$$\mathbf{t} = i \Pi_R^\dagger \mathbf{G}_C \Pi_L = i (\mathbf{V}_R)^{1/2} (\mathbf{C}_R)^{-1} \mathbf{G}_C (\mathbf{C}_L^{a\dagger})^{-1} (\mathbf{V}_L^a)^{1/2}, \quad (18.64)$$

and

$$\mathbf{t}^\dagger = -i \Pi_L^\dagger \mathbf{G}_C^a \Pi_R = -i (\mathbf{V}_L^a)^{1/2} (\mathbf{C}_L^a)^{-1} \mathbf{G}_C^a (\mathbf{C}_R^\dagger)^{-1} (\mathbf{V}_R)^{1/2}.$$

We note that the matrices $\mathbf{C}_R$ and $\mathbf{C}_L^a$ can be readily obtained from the Green's functions as shown in 18.8. We can clearly see that the transmission matrix given by Eq. 18.64 is identical to its definition in the WFM formalism (Eq. 18.41), with the identity $\mathbf{G}_C^W \equiv \mathbf{G}_C$. The mode-resolved transmissions, say from mode m in the left lead to mode k in the right lead, are given as the matrix elements

$$t_{k,m} = \left\{ i \mathbf{V}_R^{1/2} \mathbf{C}_R^{-1} \mathbf{G}_C (\mathbf{C}_L^{a\dagger})^{-1} (\mathbf{V}_L^a)^{1/2} \right\}_{k,m}. \quad (18.65)$$

This expression gives the correct mode-resolved transmission functions, but one has to perform all the matrix operations involved in obtaining t . It would be convenient to derive a simpler expression where matrix elements could be computed without the unnecessary computational

overhead of these matrix operations. In order to derive the desired expression, we multiply both sides of the identities given in Eqs. 18.59 and 18.60, by $\mathbf{V}_R^{-1/2} \mathbf{C}_R^\dagger$ on the right and by $\mathbf{C}_L^a (\mathbf{V}_L^a)^{-1/2}$ on the left, respectively, and obtain

$$\begin{aligned} \mathbf{V}_R^{-1/2} \mathbf{C}_R^\dagger \mathbf{\Gamma}_R &= \mathbf{V}_R^{1/2} \mathbf{C}_R^{-1}, \\ \mathbf{\Gamma}_L \mathbf{C}_L^a (\mathbf{V}_L^a)^{-1/2} &= (\mathbf{C}_L^{a\dagger})^{-1} (\mathbf{V}_L^a)^{1/2}. \end{aligned}$$

Using these identities, we can, finally, rewrite the transmission matrix (Eq. 18.64) in a useful form as

$$\mathbf{t} = i \mathbf{V}_R^{-1/2} \mathbf{C}_R^\dagger \mathbf{\Gamma}_R \mathbf{G}_C \mathbf{\Gamma}_L \mathbf{C}_L^a (\mathbf{V}_L^a)^{-1/2}. \quad (18.66)$$

Since the above expression contains the matrices $\mathbf{C}_R$ and $\mathbf{C}_L^a$ instead of their inverses, we can extract straightaway the mode-resolved transmissions as

$$t_{k,m} = \frac{i}{\sqrt{v_{R,k} v_{L,m}^a}} \mathbf{u}_{R,k}^\dagger \mathbf{\Gamma}_R \mathbf{G}_C \mathbf{\Gamma}_L \mathbf{u}_{L,m}^a. \quad (18.67)$$

Although similar expressions have been derived in previous works[2, 36, 37, 38], by means of mathematically sound manipulations, our derivation is much simpler and, moreover, the final expression has a very straightforward physical interpretation. Indeed, reading the expression from right to left, we have the complete picture of a phonon travelling from the left to the right leads, through the device region: $\mathbf{u}_{L,m}^a$ is the incoming wave or phonon in the mode m from the left lead, $\mathbf{\Gamma}_L$ gives the part of the incoming wave that is transmitted into the device, $\mathbf{G}_C$ is the propagator that “transports” that part to the other side of the device, $\mathbf{\Gamma}_R$ gives the part that is then transmitted to the right lead, and finally, the multiplication by $\mathbf{u}_{R,k}^\dagger$ is merely a projection that yields the probability amplitude to end up in the mode k of the right lead. We note that the group velocities serve only as normalizing factors for the eigenvectors $\mathbf{u}_{L,m}^a$ and $\mathbf{u}_{R,k}^\dagger$. However, unlike in Eq. 18.65, care must be taken when using Eq. 18.67, as a result of the presence of the group velocities in the denominator. It is to be understood that when either of the modes (m or k) is evanescent, the corresponding $t_{k,m}$ vanishes, and no calculation is needed. The above expression is to be compared to the similar formula, obtained earlier using the WFM formalism (Eq. 18.41), which, in hindsight and using the equivalence of the two formalisms, can readily be transformed into Eq. 18.66. Our formula (Eq. 18.67) is also to be compared to the Fisher-Lee expression, which is valid only if the eigenvectors are orthogonal. Our result is therefore more general than that of Fisher-Lee which has been, for years, the text-book expression[3, 35].

The question of whether the two methods are equivalent has remained unsolved. The reason is that it has been widely held that the AGF includes both evanescent and propagating modes, whereas in the WFM the first are discarded and hence do not contribute to the transmission function[25, 38]. We do, however, disagree with this analysis and assert that even in the AGF method the evanescent modes are implicitly discarded through the “escape rate” matrices. Indeed, in Eq. 18.67 the term $\mathbf{\Gamma}_L \mathbf{u}_{L,m}^a$, we refer to as the injection vector in the WFM formalism (Eq. 18.70), selects out the evanescent modes, as can be seen below (next section) where we compute this term.

18.5 Applications

In this part, we demonstrate the strength of our method by illustrating its application with a simple model, and also by comparing it to previous common methods. We consider the case of the 2D square lattice, of lattice constant a , with the central part made up of three atomic planes as shown in Fig 92. The force constants are identical throughout the system, but we

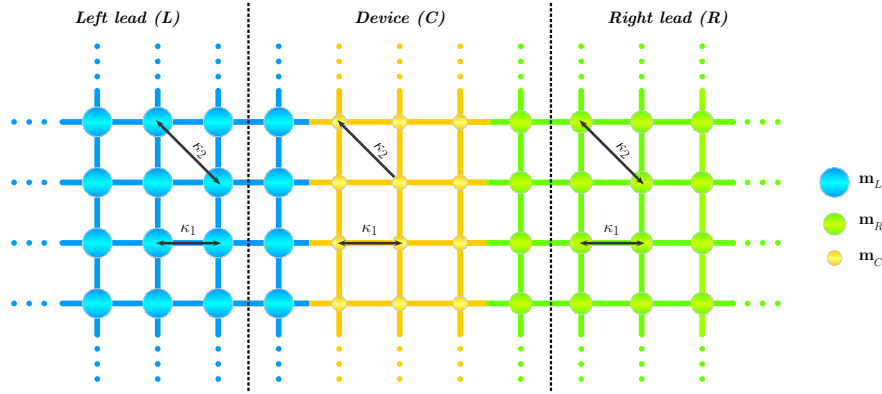


Figure 92: A schematic representation the system consisting of a square lattice with the same lattice constant and force constants (κ_1 for first-neighbor and κ_2 for second-neighbor) throughout. The masses are different in the three regions: they are m_L in the left lead, m_C in the central region, and m_R in the right lead.

	$\omega = 1.55$	$\omega = 2.20$
$\Delta \mathbf{g}_L$	$\begin{bmatrix} 2.1156 & 8.4644 \\ 8.4688 & 30.4458 \end{bmatrix} 10^{-12}$	$\begin{bmatrix} 2.1717 & 0.7383 \\ 0.7402 & 0.3862 \end{bmatrix} 10^{-12}$
$\Delta \mathbf{g}_R$	$\begin{bmatrix} 0.9315 & 3.2897 \\ 3.2895 & 12.0360 \end{bmatrix} 10^{-12}$	$\begin{bmatrix} 0.7092 & 0.9069 \\ 0.9066 & 2.3428 \end{bmatrix} 10^{-12}$
$\Delta \mathbf{B}_L$	$\begin{bmatrix} 2.0924 & 3.0534 \\ 8.9079 & 11.0897 \end{bmatrix} 10^{-12}$	$\begin{bmatrix} 2.5303 & 1.0733 \\ 1.1024 & 0.2075 \end{bmatrix} 10^{-12}$
$\Delta \mathbf{B}_R$	$\begin{bmatrix} 2.3381 & 1.8484 \\ 7.8625 & 6.8557 \end{bmatrix} 10^{-12}$	$\begin{bmatrix} 1.5516 & 0.6837 \\ 2.4664 & 1.2571 \end{bmatrix} 10^{-12}$
$\Delta \mathbf{\Gamma}_L$	$\begin{bmatrix} 3.0993 & 6.0824 \\ 6.0837 & 3.0986 \end{bmatrix} 10^{-12}$	$\begin{bmatrix} 3.3946 & 2.0217 \\ 2.0255 & 0.8398 \end{bmatrix} 10^{-12}$
$\Delta \mathbf{\Gamma}_R$	$\begin{bmatrix} 3.1228 & 9.1446 \\ 9.1450 & 3.1148 \end{bmatrix} 10^{-12}$	$\begin{bmatrix} 5.1781 & 3.0598 \\ 3.0615 & 0.6453 \end{bmatrix} 10^{-12}$
$\Delta \mathbf{V}_L^a$	$\begin{bmatrix} 3.0985 & 5.1429 \\ 5.1422 & 3.1002 \end{bmatrix} 10^{-12}$	$\begin{bmatrix} 0.2061 & 0.5931 \\ 0.5897 & 3.2494 \end{bmatrix} 10^{-12}$
$\Delta \mathbf{V}_R$	$\begin{bmatrix} 3.1173 & 6.5236 \\ 6.5243 & 3.1151 \end{bmatrix} 10^{-12}$	$\begin{bmatrix} 1.6979 & 1.6577 \\ 1.6571 & 3.1965 \end{bmatrix} 10^{-12}$
ΔT	7.9869×10^{-12}	0.7000×10^{-12}

Table 4: The absolute value of the differences between various quantities computed using the WFM and the AGF formalisms for $k_y = 1$. The quantities are the left ($\mathbf{g}_L$) and right ($\mathbf{g}_R$) surface Green's functions, the left ($\mathbf{B}_L$) and right ($\mathbf{B}_R$) Bloch matrices, the left ($\mathbf{\Gamma}_L$) and right ($\mathbf{\Gamma}_R$) escape rates and the transmission T . In the AGF formalism we used $\eta = 10^{-12}$.

	$\mathbf{Q}_{inj}$	$\mathbf{Q}_{inj}^K$
$\omega = 0.50$	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} -0.7730 - 0.0022i & -0.0017 - 0.5730i \\ -0.0018 + 0.6537i & -1.0162 + 0.0030i \end{bmatrix}$
$\omega = 1.00$	$\begin{bmatrix} -0.0072 - 1.1453i & -0.0000 - 0.0000i \\ -0.6058 + 0.0038i & 0.0000 + 0.0000i \end{bmatrix}$	$\begin{bmatrix} -0.9992 - 0.5597i & -0.4055 - 0.6913i \\ -0.2960 + 0.5285i & -0.3641 - 0.6208i \end{bmatrix}$
$\omega = 1.55$	$\begin{bmatrix} 0.2629 - 0.1315i & -2.0993 + 0.8788i \\ -0.2748 + 0.1374i & -0.1968 + 0.0824i \end{bmatrix}$	$\begin{bmatrix} 0.2914 - 0.0385i & -1.2456 - 1.9046i \\ -0.3046 + 0.0403i & -0.1168 - 0.1785i \end{bmatrix}$
$\omega = 2.20$	$\begin{bmatrix} -0.0000 - 0.0000i & -1.4994 - 0.5448i \\ -0.0000 - 0.0000i & -0.2194 + 0.6038i \end{bmatrix}$	$\begin{bmatrix} 0.6107 + 0.0604i & 1.5771 - 0.2403i \\ -2.4851 - 0.2459i & -0.0968 - 0.6351i \end{bmatrix}$

Table 5: The injection matrices computed with our WFM method($\mathbf{Q}_{inj}$) and with the method used in Komyakov *et al.* [2] ($\mathbf{Q}_{inj}^K$), for $k_y = 1$. We consider three frequencies: $\omega = 0.50$, $\omega = 1.00$, $\omega = 1.55$ and $\omega = 2.20$.

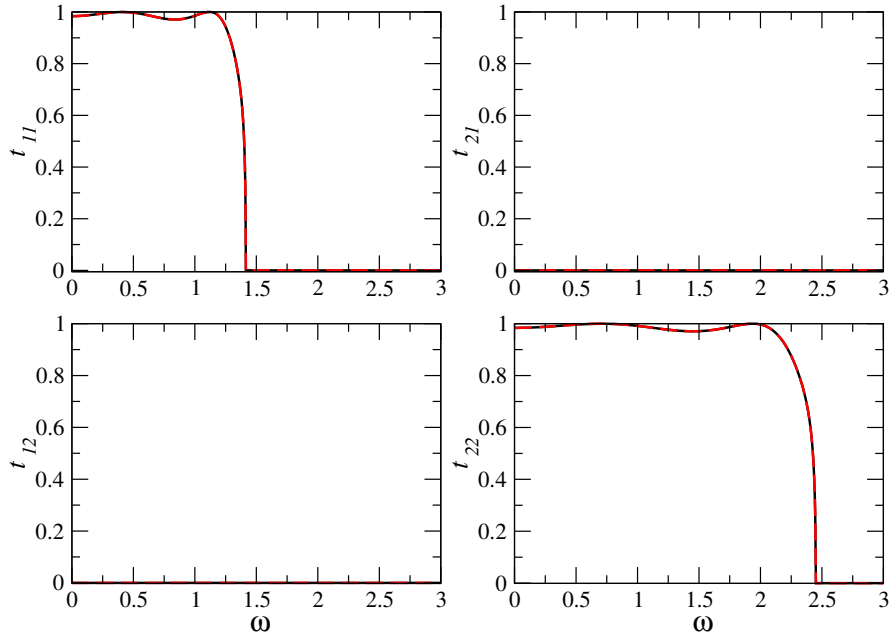


Figure 93: Mode-resolved transmission amplitudes for $k_y = 0$. Our results are shown in black continued line, and the results obtained using the Fisher-Lee formula are shown in red dashed line.

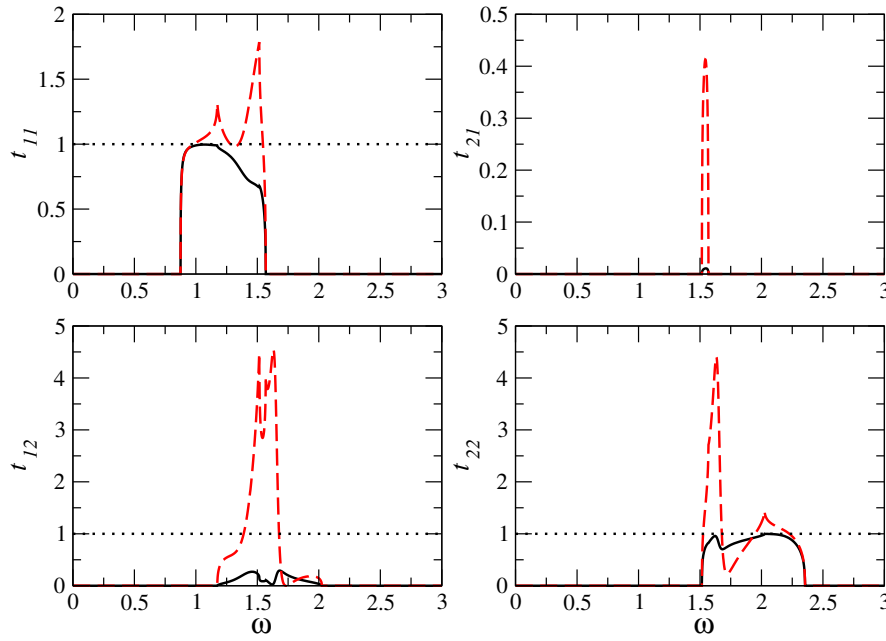


Figure 94: Mode-resolved transmission amplitudes for $k_y = 1$. Our results are shown in black continued line, and the results obtained using the Fisher-Lee formula are shown in red dashed line. Where relevant a horizontal dotted line is added to mark the maximum allowed value of unity for the transmission amplitude.

choose to use a different mass for each part of the device. The force constants are κ_1 and $\kappa_2 = 0.5 \kappa_1$ for the first and second-neighbor interactions, respectively. The masses are m_L for the left lead, $m_R = 0.6 m_L$ for the right lead, and $m_C = 0.8 m_L$ for the central part. We define the frequency unit $\omega_0 = \sqrt{\kappa_1/m_L}$. Henceforth the frequency ω is given in units of ω_0 .

In Table 4 we report the differences between various quantities computed both with our WFM method and with the AGF method for a wave vector $k_y = 1$ and for two frequencies: $\omega = 1.55$ and $\omega = 2.20$. The latter are chosen so as to include both propagating and evanescent or decaying modes. In the AGF method we chose the parameter $\eta = 10^{-12}$. In theory the parameter η is infinitesimal and one is to take the limit $\eta \rightarrow 0^+$ to recover the correct result. In the current calculation, however, we have no access to a closed-form final formula, and if a limit is to be taken, it has to be numerical in nature. Our choice of a fixed value is only meant to serve as an illustration. It is clear from Table 4 that indeed all the relevant quantities agree perfectly when computed with either of the methods. The differences are, indeed, due only to the size of the parameter η . This is a clear demonstration of the complete equivalence of our WFM method and the AGF method as shown above.

In Table 5 we report the values of the injection matrix computed both with our WFM method ($\mathbf{Q}_{inj}$) and with the “classical” WFM by Komyakov *et al.*[2] ($\mathbf{Q}_{inj}^K$) for the wave vector $k_y = 1$. We considered four frequencies: $\omega = 0.50$, $\omega = 1.00$, $\omega = 1.55$ and $\omega = 2.20$. These are chosen so as to include both propagating and evanescent modes. It is clear that, whenever one or both of the modes are decaying, our method yields a vanishing matrix element corresponding to the transmission from that mode. However, the classical method always gives finite matrix elements. As a result, our method selects from the outset only the transmission from a propagating mode. In the classical method this is not the case, and to obtain the final transmission amplitudes one has to zero out by hand the transmission amplitude whenever “incoming” mode is decaying.

We show in Fig. 93 the mode-resolved transmission amplitudes for the case of $k_y = 0$ as obtained by our method and as obtained using the Fisher-Lee formula. It is evident that our results coincide perfectly with those from the Fisher-Lee formula. In Fig. 94 we show the same

quantities computed for $k_y = 1$. We now can see that the results are very different than one another. Moreover, it is striking to see that the transmission amplitudes computed using the Fisher-Lee relation are in some cases greater than the maximum allowed value of unity. The difference, as discussed above, appears only when the wave vector matrices $\mathbf{C}_L^a$ and $\mathbf{C}_R$ are not orthogonal. This is, indeed, the case for the choice $k_y = 1$. We surmise that the failure of the Fisher-Lee relation to give proper mode-resolved transmission amplitudes has eluded researchers because they probably checked their results only in the particular situations where the wave vector matrices are orthogonal; this is the case when $k_y = 0$ as shown above. And, more importantly, many such applications were carried out in the one-dimensional case where such situation is never encountered.

18.6 Velocity matrices

Starting from the definition of the Bloch matrix $\mathbf{B}_R = \mathbf{C}_R \mathbf{\Lambda}_R \tilde{\mathbf{C}}_R$, we have $\mathbf{B}_R \mathbf{C}_R = \mathbf{C}_R \mathbf{\Lambda}_R$ and $\mathbf{C}_R^\dagger \mathbf{B}_R^\dagger = \mathbf{\Lambda}_R^\dagger \mathbf{C}_R^\dagger$. We make use of these equalities in the definition of the velocity matrix (Eq. 18.16) and obtain

$$\begin{aligned}
 \mathbf{V}_R &= i \left[\mathbf{C}_R^\dagger \mathbf{H}_{RR} \mathbf{C}_R \mathbf{\Lambda}_R - \mathbf{\Lambda}_R^\dagger \mathbf{C}_R^\dagger \mathbf{H}_{RR}^\dagger \mathbf{C}_R \right] \\
 &= i \left[\mathbf{C}_R^\dagger \mathbf{H}_{RR} \mathbf{B}_R \mathbf{C}_R - \mathbf{C}_R^\dagger \mathbf{B}_R^\dagger \mathbf{H}_{RR}^\dagger \mathbf{C}_R \right] \\
 &= i \mathbf{C}_R^\dagger \left[\mathbf{H}_{RR} \mathbf{B}_R - \mathbf{B}_R^\dagger \mathbf{H}_{RR}^\dagger \right] \mathbf{C}_R \\
 &= i \mathbf{C}_R^\dagger \left[\mathbf{\Sigma}_R - \mathbf{\Sigma}_R^\dagger \right] \mathbf{C}_R \\
 &= i \mathbf{C}_R^\dagger \left[\mathbf{\Sigma}_R - \mathbf{\Sigma}_R^a \right] \mathbf{C}_R \\
 &= \mathbf{C}_R^\dagger \mathbf{\Gamma}_R \mathbf{C}_R,
 \end{aligned} \tag{18.68}$$

where we used the expressions of the self-energy matrices in terms of the Bloch matrices as given in Eq. 18.35.

Starting again from the definition of the Bloch matrix $\mathbf{B}_L^a = \mathbf{C}_L^a \mathbf{\Lambda}_L^a \tilde{\mathbf{C}}_L^a$, we have $\mathbf{B}_L^a \mathbf{C}_L^a = \mathbf{C}_L^a \mathbf{\Lambda}_L^a$ and $\mathbf{C}_L^{a\dagger} \mathbf{B}_L^{a\dagger} = \mathbf{\Lambda}_L^{a\dagger} \mathbf{C}_L^{a\dagger}$. Similarly, we make use of these equalities in the definition of the velocity matrix (Eq. 18.19) and obtain

$$\begin{aligned}
 \mathbf{V}_L^a &= -i \left[\mathbf{C}_L^{a\dagger} \mathbf{H}_{LL}^\dagger \mathbf{C}_L^a \mathbf{\Lambda}_L^a - \mathbf{\Lambda}_L^{a\dagger} \mathbf{C}_L^{a\dagger} \mathbf{H}_{LL} \mathbf{C}_L^a \right] \\
 &= -i \left[\mathbf{C}_L^{a\dagger} \mathbf{H}_{LL}^\dagger \mathbf{B}_L^a \mathbf{C}_L^a - \mathbf{C}_L^{a\dagger} \mathbf{B}_L^{a\dagger} \mathbf{H}_{LL} \mathbf{C}_L^a \right] \\
 &= -i \mathbf{C}_L^{a\dagger} \left[\mathbf{H}_{LL}^\dagger \mathbf{B}_L^a - \mathbf{B}_L^{a\dagger} \mathbf{H}_{LL} \right] \mathbf{C}_L^a \\
 &= -i \mathbf{C}_L^{a\dagger} \left[\mathbf{\Sigma}_L^a - \mathbf{\Sigma}_L^{a\dagger} \right] \mathbf{C}_L^a \\
 &= -i \mathbf{C}_L^{a\dagger} \left[\mathbf{\Sigma}_L^a - \mathbf{\Sigma}_L \right] \mathbf{C}_L^a \\
 &= \mathbf{C}_L^{a\dagger} \mathbf{\Gamma}_L^a \mathbf{C}_L^a,
 \end{aligned} \tag{18.69}$$

where we used the expressions of the self-energy matrices in terms of the Bloch matrices as given in Eq. 18.51.

18.7 Injection matrix

From Eq. 18.40, we have $\mathbf{Q}_{\text{inj}} = \mathbf{H}_{LL}^\dagger (\mathbf{B}_L^a - \mathbf{B}_L) \mathbf{C}_L^a$ and from Eqs. 18.36 and 18.51, we have

$$\mathbf{Q}_{\text{inj}} = (\mathbf{\Sigma}_L^a - \mathbf{\Sigma}_L) \mathbf{C}_L^a = i \mathbf{\Gamma}_L \mathbf{C}_L^a. \tag{18.70}$$

Then, from Eq. 18.60, the injection matrix reads

$$\mathbf{Q}_{\text{inj}} = i \left(\mathbf{C}_L^{a\dagger} \right)^{-1} \mathbf{V}_L^a.$$

18.8 Eigenvector matrices from the Green's functions

From Eqs. 18.29 and 18.32, we multiply by $\mathbf{C}_L^a$ and $\mathbf{C}_R$ on the right and obtain

$$\mathbf{B}_L^a \mathbf{C}_L^a = \mathbf{C}_L^a \mathbf{\Lambda}_L^a$$

and

$$\mathbf{B}_R \mathbf{C}_R = \mathbf{C}_R \mathbf{\Lambda}_R.$$

Also from the expressions of $\mathbf{B}_L^a$ and $\mathbf{B}_R$ in terms of the surface Green's functions (Eqs. 18.46 and 18.55), we obtain the following right eigenvalue equation for the eigenvector matrices $\mathbf{C}_L^a$ and $\mathbf{C}_R$:

$$\begin{aligned} \mathbf{g}_L^a \mathbf{H}_{LL} \mathbf{C}_L^a &= \mathbf{C}_L^a \mathbf{\Lambda}_L^a, \\ \mathbf{g}_R \mathbf{H}_{RR}^\dagger \mathbf{C}_R &= \mathbf{C}_R \mathbf{\Lambda}_R. \end{aligned}$$

These are similar to the expressions found by Ong [32].

19 Green's function by Wimmer

19.1 Chapter 2 Green's function formalism for transport

19.1.1 2.1 Introduction

Calculating transport properties of a system involves a complicated many-body problem that as such cannot be hoped to be solved exactly. Therefore, any transport calculation involves approximations motivated by physical arguments. Here, we will consider systems with a considerable number of electrons, i.e. high density, and strong coupling to external leads. In this limit, the system can be modeled usually in terms of non-interacting particles in an effective potential, resulting from external and meanfield potentials [58]. It will not be the subject of this work to derive such an effective Hamiltonian from first principles, but we will instead use well-established models in order to derive new physical phenomena.

A common approach to transport is based on the Green's function formalism [58-63]. In particular, the non-equilibrium Green's function (NEGF) formalism pioneered by Kadanoff and Baym [64] and Keldysh [65] has gained a lot of popularity in the recent years. The non-equilibrium Green's function formalism is capable to describe systems including a finite bias and interactions. In addition, the central results of the NEGF formalism can be summarized in a few equations that can be conveniently implemented on a computer. Due to this fact, the NEGF formalism is more and more applied also to non-interacting systems, and so will we do in this work.

However, the generality and power of the non-equilibrium Green's function formalism comes at the cost of complicating the physical understanding, as an intuitive picture of the underlying physics is not obvious there, and a rather involved derivation of the central results. Especially for the beginner, the NEGF formalism is thus hard to understand. Being a formalism it can still be applied to a given problem without full understanding, but the results obtained thereby lack physical intuition. In fact, an intuitive picture may not be possible for an interacting system. However for noninteracting systems, such an intuitive understanding is still possible, as we discuss below. Applying the non-equilibrium Green's function formalism to non-interacting systems is justified because of the numerical convenience it provides, but may comprise the danger of obscuring the physical intuition, if not discussed properly.

Below, we will therefore discuss the non-equilibrium Green's function formalism in detail. Being an introduction to the subject, this chapter tries to be as self-contained as possible. We

start with some basic definitions and introduce the concept of Green's functions. In order to give a physical meaning to these Green's functions, we will explicitly calculate them for the case of a non-interacting system. We then derive the main results of the non-equilibrium Green's function formalism in the form of a perturbation theory. From these results we develop a set of transport equations in tight-binding approximation, suitable for a computer implementation. By showing connections between the non-equilibrium Green's function and the scattering formalism, we will then provide an intuitive picture of transport valid for non-interacting systems, that to some extent also carries over to interacting systems. We conclude this chapter by identifying problems that remain to be solved in order to apply the non-equilibrium formalism to arbitrary systems. Solutions to these problems will then be presented in the following two chapters.

19.1.2 2.2 Basic definitions

We start with a very general description of a many-body system with the Hamiltonian

$$\mathcal{H}(t) = \mathcal{H}_0 + \mathcal{H}_{\text{int}} + \mathcal{H}_{\text{ext}}(t) = \mathcal{H}_0 + \mathcal{H}'(t) \quad (2.1)$$

where $\mathcal{H}_0$ is the Hamiltonian of a non-interacting system, $\mathcal{H}_{\text{int}}$ describes interactions and $\mathcal{H}_{\text{ext}}(t)$ is a (possibly) time-dependent external perturbation which is assumed to vanish before some time t_0 , $\mathcal{H}_{\text{ext}}(t) = 0$ for $t < t_0$. The time-dependence of a quantum mechanical (many-body) state $|\psi\rangle_S$ is then governed by the Schrödinger equation,

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle_S = \mathcal{H}(t) |\psi(t)\rangle_S \quad (2.2)$$

Note that we use here the Schrödinger picture of quantum mechanics, where all the time-dependence is carried by the quantum mechanical state. Since any theory of Green's functions in many-body systems inevitably involves using different quantum mechanical pictures, we will carefully distinguish these by appropriate subscripts whenever such a distinction is necessary.

The Schrödinger equation can be formally solved by some time evolution operator $U(t, t_0)$, that evolves the state from time t_0 to t ,

$$|\psi(t)\rangle_S = U(t, t_0) |\psi(t_0)\rangle_S \quad (2.3)$$

The time evolution operator itself is formally obtained by integrating the Schrödinger equation,

$$i\hbar \frac{\partial}{\partial t} U(t, t_0) = \mathcal{H}(t) U(t, t_0) \quad (2.4)$$

as

$$U(t, t_0) = \mathcal{T} e^{-\frac{i}{\hbar} \int_{t_0}^t dt' \mathcal{H}(t')} \quad (2.5)$$

where $\mathcal{T}$ is the time ordering operator that orders operators according to their time argument, with the latest times to the left.

The time-dependent expectation value of an observable A with respect to some pure quantum mechanical state is given as

$$\bar{A}(t) = \langle \psi(t) |_S A_S | \psi(t) \rangle_S \quad (2.6)$$

where A_S is an operator in the Schrödinger picture. However, commonly we will consider the more general situation that the system is not in a pure state, but in a statistical mixture, such that

$$\bar{A}(t) = \sum_i p_i \langle \psi_i(t) |_S A_S | \psi_i(t) \rangle_S, \quad (2.7)$$

where p_i is the probability of being in state ψ_i . Introducing the density matrix $\rho_S(t) = \sum_i p_i |\psi_i(t)\rangle_S \langle \psi_i(t)|_S$, this expectation value can be written in a compact form,

$$\bar{A}(t) = \text{Tr}(\rho_S(t) A_S) \quad (2.8)$$

Note that the density matrix $\rho_S(t)$ is not an operator in the usual sense - contrary to the usual Schrödinger operators it is time-dependent. This is due to the time-dependence of the quantum mechanical states, and the density matrix obeys the von Neumann equation

$$i\hbar \frac{\partial}{\partial t} \rho_S(t) = [H(t), \rho_S(t)] \quad (2.9)$$

as can be seen easily from the Schrödinger equation (2.2). Here, $[A, B] = AB - BA$ is the commutator.

Contrary to the Schrödinger picture, in the Heisenberg picture all the time dependence is carried by the operators, whereas the states are time-independent. The Schrödinger and the Heisenberg picture are supposed to coincide at some time t_0 , such that

$$|\psi\rangle_{\mathcal{H}} = |\psi(t_0)\rangle_S \text{ and } A_{\mathcal{H}}(t_0) = A_S \quad (2.10)$$

Operators in the Heisenberg picture are then related to operators in the Schrödinger picture via the time evolution operator $U(t, t_0)$, Eq. (2.5),

$$A_{\mathcal{H}}(t) = U^\dagger(t, t_0) A_S U(t, t_0), \quad (2.11)$$

and they obey the equation of motion ¹

$$i\hbar \frac{\partial}{\partial t} A_{\mathcal{H}}(t) = [A_{\mathcal{H}}(t), \mathcal{H}_{\mathcal{H}}(t)] \quad (2.12)$$

Since the Heisenberg states are time-independent, so is the density matrix $\rho_{\mathcal{H}}$, and the expectation value of an observable is given as

$$\bar{A}(t) = \text{Tr}(\rho_{\mathcal{H}} A_{\mathcal{H}}(t)) \quad (2.13)$$

In this work, we will assume that the system is initially in thermal equilibrium, for $t < t_0$. In order to simplify the discussion, we take the limit $t_0 \rightarrow -\infty$ and assume that the interaction Hamiltonian $\mathcal{H}_{\text{int}}$ is switched off adiabatically when going back in time. Then the grand canonical density matrix depends only on $\mathcal{H}_0$,²

$$\rho_{\mathcal{H}} = \frac{e^{-\beta(\mathcal{H}_0 - y_0)}}{\text{Tr}(e^{-\beta(\mathcal{H}_0 - y_0)})} = \frac{1}{Z} e^{-\beta(\mathcal{H}_0 - y_0)}. \quad (2.14)$$

Here, $\beta = 1/k_B T$, where T is the temperature, k_B the Boltzmann constant, and $Z = \text{Tr}(e^{-\beta(\mathcal{H}_0 - y_0)})$ the partition function. In general, we will consider systems consisting of parts with different electrochemical potential. Consequently, we assume y_0 to be of the general form $y_0 = \sum_i \mu_i \mathcal{N}_i$, where μ_i is the electrochemical potential and $\mathcal{N}_i$ the number operator for the respective subsystem [69]. A particular example is a system consisting of leads with different electrochemical potentials, as considered in Section 2.5.2. With respect to the grand canonical ensemble, the expectation value of an observable is given as

$$\bar{A}(t) = \langle A_{\mathcal{H}}(t) \rangle = \frac{\text{Tr}(e^{-\beta(\mathcal{H}_0 - y_0)} A_{\mathcal{H}}(t))}{\text{Tr}(e^{-\beta(\mathcal{H}_0 - \dagger_0)})} \quad (2.15)$$

Many-body operators are most conveniently written in terms of creation and annihilation operators in second quantization (for an introduction see, e.g. [70]). The

operator c_n ($c_n^\dagger$) removes (creates) a particle in state $|n\rangle$, with corresponding singleparticle wave function $\varphi_n(\mathbf{x})$. Fermion statistics is imposed by the anti-commutation relations

$$\{c_n^\dagger, c_m\} = \delta_{nm}, \quad (2.16a)$$

$$\{c_n^\dagger, c_m^\dagger\} = \{c_n, c_m\} = 0 \quad (2.16b)$$

where $\{A, B\} = AB + BA$ is the anti-commutator. If the $\varphi_n(\mathbf{x})$ form a complete set, we can define field operators as

$$\begin{aligned} \psi(\mathbf{x}) &= \sum_n \varphi_n(\mathbf{x}) c_n, \text{ and} \\ \psi^\dagger(\mathbf{x}) &= \sum_n \varphi_n^*(\mathbf{x}) c_n^\dagger \end{aligned} \quad (2.17)$$

that remove and add a particle at point $\mathbf{x}$, respectively. The field operators obey the anti-commutation relations

$$\{\psi(\mathbf{x})^\dagger, \psi(\mathbf{x}')\} = \delta(\mathbf{x} - \mathbf{x}') \quad (2.18a)$$

$$\{\psi(\mathbf{x})^\dagger, \psi^\dagger(\mathbf{x}')\} = \{\psi(\mathbf{x}), \psi(\mathbf{x}')\} = 0 \quad (2.18b)$$

These commutation relations are valid for both the Schrödinger and the Heisenberg picture. It should be noted though that in the Heisenberg picture both operators have to be evaluated at the same time t . Otherwise, the complicated time-evolution leads to non-trivial relations.

Green's functions are defined as an expectation value of two or more field operators. There is a number of different definitions for Green's functions, each of them useful in a specific situation. In this work, we will consider in particular the retarded, the advanced, the lesser and the greater Green's function:

$$G^r(\mathbf{x}, t, \mathbf{x}', t') = -\frac{i}{\hbar} \Theta(t - t') \left\langle \left\{ \psi_{\mathcal{H}}(\mathbf{x}, t), \psi_{\mathcal{H}}^\dagger(\mathbf{x}', t') \right\} \right\rangle, \quad (2.19a)$$

$$G^a(\mathbf{x}, t, \mathbf{x}', t') = \frac{i}{\hbar} \Theta(t' - t) \left\langle \left\{ \psi_{\mathcal{H}}(\mathbf{x}, t), \psi_{\mathcal{H}}^\dagger(\mathbf{x}', t') \right\} \right\rangle, \quad (2.19b)$$

$$G^<(\mathbf{x}, t, \mathbf{x}', t') = \frac{i}{\hbar} \left\langle \psi_{\mathcal{H}}^\dagger(\mathbf{x}', t') \psi_{\mathcal{H}}(\mathbf{x}, t) \right\rangle, \quad (2.19c)$$

$$G^>(\mathbf{x}, t, \mathbf{x}', t') = -\frac{i}{\hbar} \left\langle \psi_{\mathcal{H}}(\mathbf{x}, t) \psi_{\mathcal{H}}^\dagger(\mathbf{x}', t') \right\rangle, \quad (2.19d)$$

¹²¹ Note that the Hamiltonian in the Heisenberg picture is defined as $\mathcal{H}_{\mathcal{H}}(t) = U^\dagger(t, t_0) \mathcal{H}(t) U(t, t_0)$. If $\mathcal{H}$ is time-independent, we have $\mathcal{H}_{\mathcal{H}}(t) = \mathcal{H}$, as then $U(t, t_0) = e^{\frac{i}{\hbar} \mathcal{H}(t-t_0)}$ commutes with $\mathcal{H}$. However, if $\mathcal{H}(t)$ is explicitly time-dependent, $U(t, t_0)$ does not necessarily commute with $\mathcal{H}(t)$, as $\mathcal{H}(t)$ itself does not necessarily commute with $\mathcal{H}(t')$ at some other time $t' \neq t$.

² In taking this limit, we neglect initial correlations imposed by $\mathcal{H}_{\text{int}}$. It is in principle possible to include these correlations when developing a perturbation theory for the Green's function, as is, for example, done in Refs. [61, 63, 64, 66]. These accounts take the limit $t_0 \rightarrow -\infty$ after the perturbation theory has been developed - but this is equivalent to already taking the limit $t_0 \rightarrow -\infty$ in the very beginning, as we do here. Taking the limit early simplifies the subsequent discussion considerably, and nevertheless yields identical results.

Usually it is argued that initial correlations are washed out by the interactions, as eventually steady state is reached [61]. It is possible to develop a perturbation theory including the initial correlations [67, 68], but such a treatment becomes quite involved. In any case, as we are not considering interacting systems in this work anyway, we need not worry about this point.

where $\Theta(t)$ is the Heaviside step function. The anti-commutator in the definitions of advanced and retarded Green's function involves field operators at different times and can thus not be evaluated with the anti-commutation relations at equal times, Eq. (2.18a).

These various Green's functions carry different information. For example, the lesser Green's function directly allows the evaluation of physical observables, such as the local charge carrier density,

$$n(\mathbf{x}, t) = \langle \psi_{\mathcal{H}}^{\dagger}(\mathbf{x}, t) \psi_{\mathcal{H}}(\mathbf{x}, t) \rangle = -i\hbar G^{<}(\mathbf{x}, t, \mathbf{x}, t) \quad (2.20)$$

As a further example, the retarded Green's function can be used to extract the transmission probability of charge carriers through a system (see Section 2.5.2).

From the definitions (2.19a)-(2.19d) it is obvious that these various Green's functions are not independent. For example, we find the useful identities

$$G^r(\mathbf{x}, t, \mathbf{x}', t') = (G^a(\mathbf{x}', t', \mathbf{x}, t))^* \quad (2.21)$$

$$G^{<}(\mathbf{x}, t, \mathbf{x}', t') = -(G^{<}(\mathbf{x}', t', \mathbf{x}, t))^* \quad (2.22)$$

$$G^r - G^a = G^{>} - G^{<} \quad (2.23)$$

In equilibrium [58] and in non-equilibrium steady state [61] the Green's functions only depend on the time difference, $G(\mathbf{x}, t, \mathbf{x}', t') = G(\mathbf{x}, \mathbf{x}'; t - t')$. In this situation, it is useful to introduce the energy-dependent Green's function through Fourier transformation,

$$G(\mathbf{x}, \mathbf{x}'; E) = \int_{-\infty}^{\infty} dt e^{\frac{i}{\hbar} E t} G(\mathbf{x}, \mathbf{x}'; t) \quad (2.24)$$

The inverse Fourier transformation is then given by

$$G(\mathbf{x}, \mathbf{x}'; t - t') = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dE e^{-\frac{i}{\hbar} E(t-t')} G(\mathbf{x}, \mathbf{x}'; E) \quad (2.25)$$

For the purpose of developing a perturbation theory for the ground state, the causal (or time-ordered) Green's function

$$G^c(\mathbf{x}, t, \mathbf{x}', t') = -\frac{i}{\hbar} \langle \mathcal{T} \psi_{\mathcal{H}}(\mathbf{x}, t) \psi_{\mathcal{H}}^{\dagger}(\mathbf{x}', t') \rangle \quad (2.26)$$

plays a crucial role (see, e.g. [58]). Here, $\mathcal{T}$ denotes the time ordering operator for fermionic operators,

$$\mathcal{T} A(t_1) B(t_2) = \begin{cases} +A(t_1) B(t_2) & \text{for } t_1 > t_2 \\ -B(t_2) A(t_1) & \text{for } t_2 > t_1 \end{cases} \quad (2.27)$$

and analogously for more than two operators, with a minus sign for every operator interchange³. A generalization of this causal Green's function will be the foundation of a non-equilibrium perturbation theory in Section 2.4. However, before rising to this challenge, we first examine the various Green's functions defined here in the limit of a non-interacting system.

¹²³ There was no minus sign in the definition of the time ordering operator in Eq. (2.5): The operator to be ordered there, $\mathcal{H}(t)$, always consists of an even number of fermionic operators, as does any observable.

19.1.3 2.3 Green's functions of non-interacting systems

We now consider the particular case of a non-interacting system, where $\mathcal{H} = \mathcal{H}_0$. The Hamiltonian can then be written as

$$\mathcal{H}_0 = \int d\mathbf{x} \psi^\dagger(\mathbf{x}) H_0(\mathbf{x}) \psi(\mathbf{x}) \quad (2.28)$$

where $H_0(\mathbf{x})$ is a single particle Hamiltonian⁴. Alternatively, introducing the set of eigenstates $\{\varphi_n\}$ of the single particle Hamiltonian, $H_0\varphi_n = E_n\varphi_n$, we can write the Hamiltonian as

$$\mathcal{H}_0 = \sum_n E_n c_n^\dagger c_n \quad (2.29)$$

where c_n ($c_n^\dagger$) removes (creates) a particle in state φ_n . Since the non-interacting Hamiltonian is time-independent, the time-dependence of the field operator in the Heisenberg picture is given by

$$\begin{aligned} \psi_{\mathcal{H}_0}(\mathbf{x}, t) &= e^{\frac{i}{\hbar} \mathcal{H}_0(t-t_0)} \psi_{\mathcal{H}_0}(\mathbf{x}, t_0) e^{-\frac{i}{\hbar} \mathcal{H}_0(t-t_0)} \\ &= e^{\frac{i}{\hbar} \mathcal{H}_0(t-t_0)} \psi(\mathbf{x}) e^{-\frac{i}{\hbar} \mathcal{H}_0(t-t_0)} \end{aligned} \quad (2.30)$$

Here, we use the subscript $\mathcal{H}_0$ to emphasize that the dynamics is governed by the non-interacting Hamiltonian. The equation of motion for the field operator is then

$$i\hbar \frac{\partial}{\partial t} \psi_{\mathcal{H}_0}(\mathbf{x}, t) = [\psi_{\mathcal{H}_0}(\mathbf{x}, t), \mathcal{H}_0] = H_0(\mathbf{x}) \psi_{\mathcal{H}_0}(\mathbf{x}, t) \quad (2.31)$$

where we made use of the anti-commutation relations (2.18). In the same manner, we find

$$i\hbar \frac{\partial}{\partial t} c_{n,\mathcal{H}_0}(t) = [c_{n,\mathcal{H}_0}(t), \mathcal{H}_0] = E_n c_{n,\mathcal{H}_0}(t) \quad (2.32)$$

that can be easily integrated as

$$c_{n,\mathcal{H}_0}(t) = e^{-\frac{i}{\hbar} E_n(t-t_0)} c_n \quad (2.33)$$

Thus, the time-dependence of the field operator can be explicitly written as

$$\psi_{\mathcal{H}_0}(\mathbf{x}, t) = \sum_n \varphi_n(\mathbf{x}) e^{-\frac{i}{\hbar} E_n(t-t_0)} c_n \quad (2.34)$$

Equipped with these expressions for the time dependence of the field operators, we now turn to the Green's functions. Note that we subsequently mark these with a

subscript as G_0 , in order to indicate the dependence on the non-interacting Hamiltonian $\mathcal{H}_0$. We begin with deriving an equation of motion for the retarded Green's function:

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} G_0^r(\mathbf{x}, t, \mathbf{x}', t') &= i\hbar \frac{\partial}{\partial t} \left[-\frac{i}{\hbar} \Theta(t-t') \left\langle \left\{ \psi_{\mathcal{H}_0}(\mathbf{x}, t), \psi_{\mathcal{H}_0}^\dagger(\mathbf{x}', t') \right\} \right\rangle \right] \\ &= \delta(t-t') \left\langle \left\{ \psi_{\mathcal{H}_0}(\mathbf{x}, t), \psi_{\mathcal{H}_0}^\dagger(\mathbf{x}', t) \right\} \right\rangle \\ &\quad - \frac{i}{\hbar} \Theta(t-t') \left\langle \left\{ i\hbar \frac{\partial}{\partial t} \psi_{\mathcal{H}_0}(\mathbf{x}, t), \psi_{\mathcal{H}_0}^\dagger(\mathbf{x}', t') \right\} \right\rangle \end{aligned} \quad (2.35)$$

¹²⁴ For example, the Hamiltonian of an electron in a potential V , $H_0(\mathbf{x}) = -\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{x})$

In the first term we have set $t = t'$ because of the delta function. Using the anticommutation relation (2.18a) to simplify the first term and the equation of motion for the field operator (2.31) for the second term, we find

$$\left(i\hbar\frac{\partial}{\partial t} - H_0(\mathbf{x})\right) G_0^r(\mathbf{x}, t, \mathbf{x}', t') = \delta(t - t') \delta(\mathbf{x} - \mathbf{x}') \quad (2.36)$$

The retarded Green's function thus obeys the single particle Schrödinger equation with a delta function source term ⁵. In contrast, we find for the lesser Green's function:

$$\left(i\hbar\frac{\partial}{\partial t} - H_0(\mathbf{x})\right) G_0^<(\mathbf{x}, t, \mathbf{x}', t') = 0 \quad (2.37)$$

Equivalent equations hold for the advanced and greater Green's functions. Using the notation $G_0^{-1} = (i\hbar\frac{\partial}{\partial t} - H_0(\mathbf{x}))$, we can summarize these results formally as

$$G_0^{-1} G_0^{r,a} = \mathbb{I} \quad \text{and} \quad G_0^{-1} G_0^{<, >} = 0 \quad (2.38)$$

After deriving the equation of motions for the Green's functions for non-interacting systems, we now turn to explicitly calculating them.

The retarded Green's function is given as

$$\begin{aligned} G_0^r(\mathbf{x}, t, \mathbf{x}', t') &= -\frac{i}{\hbar} \Theta(t - t') \left\langle \left\{ \psi_{\mathcal{H}_0}(\mathbf{x}, t), \psi_{\mathcal{H}_0}^\dagger(\mathbf{x}', t') \right\} \right\rangle \\ &= -\frac{i}{\hbar} \Theta(t - t') \sum_{n,m} \varphi_n(\mathbf{x}) \varphi_m^*(\mathbf{x}') e^{-\frac{i}{\hbar}(E_n(t-t_0) - E_m(t'-t_0))} \underbrace{\{c_n, c_m^\dagger\}}_{=\delta_{n,m}} \\ &= -\frac{i}{\hbar} \Theta(t - t') \sum_n \varphi_n(\mathbf{x}) \varphi_n^*(\mathbf{x}') e^{-\frac{i}{\hbar} E_n(t-t')} \end{aligned} \quad (2.39)$$

where we made use of Eq. (2.34) and the anti-commutation relations (2.16). Note that the dependence on the density matrix dropped out completely.

In order to calculate the energy dependent retarded Green's function, we use the representation

$$\Theta(t) = -\frac{1}{2\pi i} \int_{-\infty}^{\infty} dE' e^{-\frac{i}{\hbar} E' t} \frac{1}{E' + i\eta} \quad (2.40)$$

for the Heaviside step function, where η is an infinitesimally small positive number. We arrive at

$$\begin{aligned} G_0^r(\mathbf{x}, \mathbf{x}'; E) &= \int_{-\infty}^{\infty} dt e^{\frac{i}{\hbar} E t} G_0^r(\mathbf{x}, \mathbf{x}'; t) \\ &= \frac{1}{2\pi\hbar} \sum_n \varphi_n(\mathbf{x}) \varphi_n^*(\mathbf{x}') \int_{-\infty}^{\infty} dE' \int_{-\infty}^{\infty} dt e^{\frac{i}{\hbar}(E - E_n - E')t} \frac{1}{E' + i\eta} \\ &= \sum_n \varphi_n(\mathbf{x}) \varphi_n^*(\mathbf{x}') \int_{-\infty}^{\infty} dE' \delta(E - E_n - E') \frac{1}{E' + i\eta} \\ &= \sum_n \frac{\varphi_n(\mathbf{x}) \varphi_n^*(\mathbf{x}')}{E - E_n + i\eta} \end{aligned} \quad (2.41)$$

¹²₅ The Green's function thus obeys a equation of the form of $LG(x, x') = \delta(x - x')$, where L is a linear differential operator. This kind of equation is known in the context of inhomogeneous differential equations, e.g. in electrodynamics, where the concept of a Green's function was introduced first. This similarity in the the equation of motion is the reason for the name "Green's function" in the context discussed here.

The non-interacting many-body retarded Green's function is thus identical to the wellknown single-particle retarded Green's function [59, 60]. The retarded Green's function does not carry any information about temperature or chemical potential, it only depends on the eigenfunctions and -energies of the single-particle Hamiltonian H_0 . From Eq. (2.21) we immediately find an expression for the advanced Green's function,

$$G_0^a(\mathbf{x}, \mathbf{x}'; E) = \sum_n \frac{\varphi_n(\mathbf{x}) \varphi_n^*(\mathbf{x}')}{E - E_n - i\eta} \quad (2.42)$$

These expressions may be written in a more compact form in terms of an operator identity, using $H_0 |\varphi_n\rangle = E_n |\varphi_n\rangle$,

$$\begin{aligned} G_0^{r(a)}(\mathbf{x}, \mathbf{x}'; E) &= \langle \mathbf{x} | \sum_n \frac{|\varphi_n\rangle \langle \varphi_n|}{E - E_n \pm i\eta} | \mathbf{x}' \rangle \\ &= \langle \mathbf{x} | \left(\frac{1}{E - H_0 \pm i\eta} \sum_n \underbrace{|\varphi_n\rangle \langle \varphi_n|}_{= \mathbb{1}} \right) | \mathbf{x}' \rangle \\ &= \langle \mathbf{x} | (E - H_0 \pm i\eta)^{-1} | \mathbf{x}' \rangle \end{aligned} \quad (2.43)$$

In order to calculate $G_0^<$ and $G_0^>$, we derive relations connecting them with G_0^r and G_0^a , following an argument by Kadanoff and Baym [64]. We assume that the system is in equilibrium, governed by a single electrochemical potential μ . Then, the density matrix is of the form $e^{-\beta(\mathcal{H}_0 - \mu\mathcal{N})}$ and we find

$$\begin{aligned} G_0^<(\mathbf{x}, t, \mathbf{x}', 0) &= \frac{i}{\hbar} \frac{1}{Z} \text{Tr} \left(e^{-\beta(\mathcal{H}_0 - \mu\mathcal{N})} \psi_{\mathcal{H}_0}^\dagger(\mathbf{x}', 0) \psi_{\mathcal{H}_0}(\mathbf{x}, t) \right) \\ &= \frac{i}{\hbar} \frac{1}{Z} \text{Tr} \left(e^{-\beta(\mathcal{H}_0 - \mu\mathcal{N})} \psi_{\mathcal{H}_0}(\mathbf{x}, t) e^{-\beta(\mathcal{H}_0 - \mu\mathcal{N})} \psi_{\mathcal{H}_0}^\dagger(\mathbf{x}', 0) e^{+\beta(\mathcal{H}_0 - \mu\mathcal{N})} \right), \end{aligned} \quad (2.44)$$

where we made use of the cyclic invariance of the trace. Since $\psi_{\mathcal{F}_0}^\dagger(\mathbf{x}', 0)$ creates a particle, we have

$$\psi_{\mathcal{H}_0}^\dagger(\mathbf{x}', 0) f(\mathcal{N}) = f(\mathcal{N} - 1) \psi_{\mathcal{H}_0}^\dagger(\mathbf{x}', 0) \quad (2.45)$$

for any function $f(\mathcal{N})$ of the number operator, as can be seen by operating on states with a given number of particles. Therefore, we find

$$e^{\beta\mu\mathcal{N}} \psi_{\mathcal{H}_0}^\dagger(\mathbf{x}, 0) e^{-\beta\mu\mathcal{N}} = e^{\beta\mu} \psi_{\mathcal{H}_0}^\dagger(\mathbf{x}, 0), \quad (2.46)$$

and thus

$$\begin{aligned} e^{-\beta(\mathcal{H}_0 - \mu\mathcal{N})} \psi_{\mathcal{H}_0}^\dagger(\mathbf{x}, 0) e^{\beta(\mathcal{H}_0 - \mu\mathcal{N})} &= e^{-\beta\mathcal{H}_0} e^{\beta\mu\mathcal{N}} \psi_{\mathcal{H}_0}^\dagger(\mathbf{x}, 0) e^{-\beta\mu\mathcal{N}} e^{\beta\mathcal{H}_0} \\ &= e^{\beta\mu} e^{-\beta\mathcal{H}_0} \psi_{\mathcal{H}_0}^\dagger(\mathbf{x}, 0) e^{\beta\mathcal{H}_0} \\ &= e^{\beta\mu} \psi_{\mathcal{H}_0}^\dagger(\mathbf{x}, i\hbar\beta) \end{aligned} \quad (2.47)$$

Here we used the fact that the number operator $\mathcal{N}$ commutes with the Hamiltonian, and an analytic continuation of Eq. (2.30) into the complex plane. We finally arrive at

$$\begin{aligned} G_0^<(\mathbf{x}, t, \mathbf{x}', 0) &= \frac{i}{\hbar} e^{\beta\mu} \frac{1}{Z} \text{Tr} \left(e^{-\beta(\mathcal{H}_0 - \mu\mathcal{N})} \psi_{\mathcal{H}_0}(\mathbf{x}, t) \psi_{\mathcal{F}_0}^\dagger(\mathbf{x}, i\hbar\beta) \right) \\ &= -e^{\beta\mu} G_0^>(\mathbf{x}, t, \mathbf{x}', i\hbar\beta) \end{aligned} \quad (2.48)$$

Taking the Fourier transform, we find the energy-dependent lesser Green's function,

$$\begin{aligned}
 G_0^<(\mathbf{x}, \mathbf{x}'; E) &= -e^{\beta\mu} \int_{-\infty}^{\infty} dt e^{\frac{i}{\hbar}Et} G_0^>(\mathbf{x}, \mathbf{x}'; t - i\hbar\beta) \\
 &= -e^{-\beta(E-\mu)} \int_{-\infty-i\hbar\beta}^{\infty-i\hbar\beta} dt e^{\frac{i}{\hbar}Et} G_0^>(\mathbf{x}, \mathbf{x}'; t) \\
 &= -e^{-\beta(E-\mu)} G_0^>(\mathbf{x}, \mathbf{x}'; E)
 \end{aligned} \tag{2.49}$$

where we made use of the fact that $G_0^>(\mathbf{x}, \mathbf{x}'; t - t')$ is an analytic function for $-\hbar\beta < \Im(t - t') < 0$ [63, 64]. We now define the spectral density as

$$\begin{aligned}
 A(\mathbf{x}, \mathbf{x}'; E) &= i(G_0^r(\mathbf{x}, \mathbf{x}'; E) - G_0^a(\mathbf{x}, \mathbf{x}'; E)) \\
 &= i(G_0^>(\mathbf{x}, \mathbf{x}'; E) - G_0^<(\mathbf{x}, \mathbf{x}'; E)),
 \end{aligned} \tag{2.50}$$

where the second equality is due to Eq. (2.23). Inserting Eq. (2.49) we find

$$A(\mathbf{x}, \mathbf{x}'; E) = -i(G_0^<(\mathbf{x}, \mathbf{x}'; E) + e^{\beta(E-\mu)} G_0^<(\mathbf{x}, \mathbf{x}'; E)) \tag{2.51}$$

and hence

$$G_0^<(\mathbf{x}, \mathbf{x}'; E) = i f_0(E) A(\mathbf{x}, \mathbf{x}'; E) \tag{2.52a}$$

$$G_0^>(\mathbf{x}, \mathbf{x}'; E) = -i(1 - f_0(E)) A(\mathbf{x}, \mathbf{x}'; E) \tag{2.52b}$$

where $f_0(E) = 1/(e^{\beta(E-\mu)} + 1)$ is the well-known Fermi-Dirac distribution. Eqs. (2.52a) and (2.52b) are known as the fluctuation dissipation theorem and also hold for interacting systems in equilibrium ⁶.

Inserting the previous results for the retarded and advanced Green's function from Eqs. (2.41) and (2.42), Eq. (2.52a) yields an explicit expression for $G_0^<$. Making use of the Sokhotsky-Weierstrass theorem,

$$\frac{1}{E \pm i\eta} = \mathcal{P} \frac{1}{E} \mp i\pi\delta(E) \tag{2.53}$$

where $\mathcal{P}$ denotes the Cauchy principal value integral, we arrive at

$$A(\mathbf{x}, \mathbf{x}'; E) = 2\pi \sum_n \varphi_n(\mathbf{x}) \varphi_n^*(\mathbf{x}') \delta(E - E_n) \tag{2.54}$$

and corresponding equations for $G_0^<$ and $G_0^>$. From the diagonal elements of the spectral density we find the local density of states ⁷ (LDOS) [59]

$$\begin{aligned}
 d(\mathbf{x}, E) &= \sum_n |\varphi_n(\mathbf{x})|^2 \delta(E - E_n) = \frac{1}{2\pi} A(\mathbf{x}, \mathbf{x}; E) \\
 &= \frac{i}{2\pi} (G_0^r(\mathbf{x}, \mathbf{x}; E) - (G_0^r(\mathbf{x}, \mathbf{x}; E))^*) \\
 &= -\frac{1}{\pi} \Im(G_0^r(\mathbf{x}, \mathbf{x}; E))
 \end{aligned} \tag{2.55}$$

The electron density for a non-interacting system in equilibrium is then given as

$$\begin{aligned}
n(\mathbf{x}) &= -i\hbar G_0^<(\mathbf{x}, \mathbf{x}; 0) = -\frac{i}{2\pi} \int_{-\infty}^{\infty} dE G_0^<(\mathbf{x}, \mathbf{x}; E) \\
&= \int_{-\infty}^{\infty} dE f_0(E) d(\mathbf{x}, E)
\end{aligned} \tag{2.56}$$

Thus, in the case of an equilibrium system without interactions, the various Green's functions have a clear interpretation: The retarded Green's function contains the spectrum, i.e. the eigenfunctions and -energies, whereas the lesser Green's function describes how this spectrum is filled with particles. Even in a non-equilibrium situation, including interactions to some degree, this interpretation still holds, as is discussed below. However, the system is not described by a single Fermi-Dirac distribution anymore, and Eqs. (2.52) are replaced by a more complicated relation. The derivation of this relation for $G^<$ is the subject of the next section.

19.1.4 2.4 Non-equilibrium perturbation theory

19.1.5 2.4.1 The need for a non-equilibrium theory

The non-equilibrium perturbation theory has been originally developed by Kadanoff and Baym [64], and from a complementary point of view by Keldysh⁸ [65], and independently by Craig [71]. Since then, it has been the subject of many reviews and books [59, 61, 63, 66, 67, 72, 73].

The basic idea of perturbation theory, equilibrium and non-equilibrium, is to express operators with a time dependence governed by the full Hamiltonian $\mathcal{H}$, such as $\psi_{\mathcal{H}}(\mathbf{x}, t)$, in terms of operators governed only by the non-interacting Hamiltonian $\mathcal{H}_0$, such as $\psi_{\mathcal{H}_0}(\mathbf{x}, t)$. Whereas the time dependence of $\psi_{\mathcal{H}}(\mathbf{x}, t)$ is in general not known, the time dependence of $\psi_{\mathcal{H}_0}(\mathbf{x}, t)$ can be calculated, as seen in the previous paragraph.

When the time dependence of an operator is only governed by the non-interacting part $\mathcal{H}_0$,

$$A_{\mathcal{F}_0}(t) = e^{\frac{i}{\hbar}\mathcal{H}_0(t-t_0)} A_S e^{-\frac{i}{\hbar}\mathcal{H}_0(t-t_0)} \tag{2.57}$$

instead of the full Hamiltonian $\mathcal{H}$, it is said to be in the interaction picture. Formally, the Heisenberg and the interaction picture are connected by

$$A_{\mathcal{H}}(t) = u^\dagger(t, t_0) A_{\mathcal{H}_0}(t) u(t, t_0), \tag{2.58}$$

where

$$u(t, t_0) = e^{\frac{i}{\hbar}\mathcal{H}_0(t-t_0)} U(t, t_0) \tag{2.59}$$

is the time-evolution operator in the interaction picture. This operator obeys the equation of motion

$$\begin{aligned}
i\hbar \frac{\partial}{\partial t} u(t, t_0) &= -e^{\frac{i}{\hbar}\mathcal{H}_0(t-t_0)} \mathcal{H}_0 U(t, t_0) + e^{\frac{i}{\hbar}\mathcal{H}_0(t-t_0)} \mathcal{H}(t) U(t, t_0) \\
&= -e^{\frac{i}{\hbar}\mathcal{H}_0(t-t_0)} \mathcal{H}_0 e^{-\frac{i}{\hbar}\mathcal{H}_0(t-t_0)} u(t, t_0) \\
&\quad + e^{\frac{i}{\hbar}\mathcal{H}_0(t-t_0)} (\mathcal{H}_0 + \mathcal{H}'(t)) e^{-\frac{i}{\hbar}\mathcal{H}_0(t-t_0)} u(t, t_0) \\
&= \mathcal{H}'_{\mathcal{H}_0}(t) u(t, t_0)
\end{aligned} \tag{2.60}$$

¹²⁶ The proof, as presented here, carries over to the interacting case unchanged [64].

⁷ The local density of states, $d(\mathbf{x}, E) = \sum_n |\varphi_n(\mathbf{x})|^2 \delta(E - E_n)$ is the usual density of states, $\sum_n \delta(E - E_n)$, weighted by the spatial probability density $|\varphi_n(\mathbf{x})|^2$ of the respective electronic level.

where we made use of the Schrödinger equation for the time evolution operator $U(t, t_0)$, Eq. (2.4). As for $U(t, t_0)$, we can formally integrate this equation as

$$u(t, t_0) = \mathcal{T} e^{-\frac{i}{\hbar} \int_{t_0}^t dt' \mathcal{H}'_{\mathcal{H}_0}(t')} \quad \text{for } t > t_0 \quad (2.61)$$

Again, $\mathcal{T}$ denotes the time-ordering operator.

Note that the perturbation Hamiltonian in the interaction picture, $\mathcal{H}'_{\mathcal{F}_0}(t)$, itself only contains operators in the interaction picture, i.e. with a time-dependence governed by

$\mathcal{H}_0$. In principle, the complicated time dependence of the Heisenberg operators has thus been reduced to the simple time dependence of operators in the interaction picture - at the cost of a time-ordering operator and an infinite power series of operators, due to the exponential. In practice, this series is often only considered up to some order, in the form of a perturbation theory.

Before pursuing further into this direction, we first collect some useful identities for $u(t, t_0)$. Taking the Hermitian conjugate of Eq.(2.61), we obtain

$$u^\dagger(t, t_0) = \tilde{\mathcal{T}} e^{-\frac{i}{\hbar} \int_{t_0}^t dt' \mathcal{H}'_{\mathcal{F}_0}(t')} \quad \text{for } t > t_0, \quad (2.62)$$

where $\tilde{\mathcal{T}}$ denotes the anti-time ordering⁹. Whereas $u(t, t_0)$ corresponds to a movement forward in time from t_0 to t within the interaction picture, $u^\dagger(t, t_0)$ takes the time back from t to t_0 . From these, we can define the time-evolution operator in the interaction picture for arbitrary times $t_1 > t_2 > t_0$ as $u(t_1, t_2) = u(t_1, t_0) u^\dagger(t_2, t_0)$ and find:

$$u(t_1, t_2) u^\dagger(t_1, t_2) = u^\dagger(t_1, t_2) u(t_1, t_2) = \mathbb{I} \text{ for } t_1 > t_2, \quad (2.63a)$$

$$u(t_1, t_2) u(t_2, t_3) = u(t_1, t_3) \text{ for } t_1 > t_2 > t_3, \quad (2.63b)$$

$$u^\dagger(t_1, t_2) u(t_1, t_3) = u(t_2, t_3) \text{ for } t_1 > t_2 > t_3 \quad (2.63c)$$

Up to now, we have only collected ideas that are also used in the well-established ground state Green's function formalism, such as presented in [58]. Why, respectively when is a non-equilibrium theory needed?

In order to answer this question, we follow an argument by Craig [71]. Suppose $|\psi_n(t_0)\rangle$ denotes a non-interacting many-body quantum state for $t_0 \rightarrow -\infty$. The matrix element of an operator with respect to this state is given as

$$\begin{aligned} \langle \psi_n(t_0) | A_{\mathcal{H}}(t) | \psi_n(t_0) \rangle &= \langle \psi_n(t_0) | u^\dagger(t, t_0) A_{\mathcal{F}_0}(t) u(t, t_0) | \psi_n(t_0) \rangle \\ &= \langle \psi_n(t_0) | u^\dagger(t_\alpha, t_0) u(t_\alpha, t) A_{\mathcal{H}_0}(t) u(t, t_0) | \psi_n(t_0) \rangle \\ &= \sum_m \langle \psi_n(t_0) | u^\dagger(t_\alpha, t_0) | \psi_m(t_0) \rangle \times \\ &\quad \langle \psi_m(t_0) | u(t_\alpha, t) A_{\mathcal{F}_0}(t) u(t, t_0) | \psi_n(t_0) \rangle, \end{aligned} \quad (2.64)$$

where $t_\alpha > t$ is some time later than any other time in the problem. Here, we made use of Eqs. (2.63) and inserted a complete set of states. If $|\psi_n(t_0)\rangle$ is a non-degenerate state, and interactions are switched on adiabatically, the time propagator in the interaction picture (that contains the interaction Hamiltonian) cannot cause transitions between states, and

$$\langle \psi_n(t_0) | u^\dagger(t_\alpha, t_0) | \psi_m(t_0) \rangle = \langle \psi_n(t_0) | u^\dagger(t_\alpha, t_0) | \psi_n(t_0) \rangle \times \delta_{nm} \quad (2.65)$$

This is the case, when $|\psi_m(t_0)\rangle = |\Phi_0\rangle$, where $|\Phi_0\rangle$ is the non-degenerate ground state of the non-interacting system. Furthermore, from $u^\dagger(t_\alpha, t_0) u(t_\alpha, t_0) = \mathbb{I}$ and Eq. (2.65) we obtain

¹²⁸ Often, the non-equilibrium Green's function theory is thus referred to as Keldysh Green's function theory.

$$\langle \psi_n(t_0) | u^\dagger(t_\alpha, t_0) | \psi_n(t_0) \rangle = (\langle \psi_n(t_0) | u(t_\alpha, t_0) | \psi_n(t_0) \rangle)^{-1} \quad (2.66)$$

In the limit $t_\alpha \rightarrow \infty$ and $t_0 \rightarrow -\infty$, we then arrive at

$$\langle A_{\mathcal{H}}(t) \rangle = \frac{\langle \Phi_0 | u(\infty, t) A_{\mathcal{H}_0}(t) u(t, -\infty) | \Phi_0 \rangle}{\langle \Phi_0 | u(\infty, -\infty) | \Phi_0 \rangle} \quad (2.67)$$

which is the foundation of the ground state Green's function theory [58].

However, this is only true for a non-degenerate state and an adiabatic time dependence of the interactions. There may very well be transitions between states for finite temperatures, as excited states are commonly degenerate, and for non-equilibrium situations, where transitions between states happen by definition. In these cases, Eq. (2.65) does not hold, and it is not possible to eliminate $u^\dagger(t, t_0)$ from the equations.

In the general case, we have to take a non-interacting state in the far past, evolve it in time up to t , act on this evolved state and then bring it back into the far past again, before taking the overlap with the non-interacting state again. This is the essential content of Eq. (2.64),

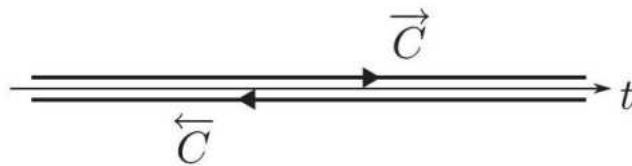
$$\langle \psi_n(t_0) | A_{\mathcal{H}}(t) | \psi_n(t_0) \rangle = \langle \psi_n(t_0) | u^\dagger(t, t_0) A_{\mathcal{H}_0}(t) u(t, t_0) | \psi_n(t_0) \rangle. \quad (2.68)$$

In contrast, in the ground state formalism, the ground state evolves uniquely in time, and as the interaction is switched off again in the far future, we arrive again at the ground state¹⁰. In the ground state Green's function theory it therefore suffices to only consider time-evolution in one direction, whereas the non-equilibrium theory necessarily involves going forward and backward in time, as no state in the future can be uniquely identified with a state in the past. The next section will render these arguments into a rigorous theory.

19.1.6 2.4.2 Contour-ordered Green's function theory

We now develop a non-equilibrium Green's function theory taking into account the arguments of the preceding section, and start by considering the causal Green's function G^c that is the foundation of a perturbation theory in the ground state formalism. Without loss of generality, we assume $t > t'$ and find

$$\begin{aligned} G^c(\mathbf{x}, t, \mathbf{x}', t') &= -\frac{i}{\hbar} \langle \mathcal{T} \psi_{\mathcal{H}}(\mathbf{x}, t) \psi_{\mathcal{H}}^\dagger(\mathbf{x}', t') \rangle \\ &= -\frac{i}{\hbar} \langle \psi_{\mathcal{H}}(\mathbf{x}, t) \psi_{\mathcal{H}}^\dagger(\mathbf{x}', t') \rangle \\ &= -\frac{i}{\hbar} \langle u^\dagger(t, -\infty) \psi_{\mathcal{F}_0}(\mathbf{x}, t) u(t, -\infty) u^\dagger(t', -\infty) \psi_{\mathcal{H}_0}^\dagger(\mathbf{x}', t') u(t', -\infty) \rangle \\ &= -\frac{i}{\hbar} \langle u^\dagger(t, -\infty) \underbrace{u^\dagger(\infty, t) u(\infty, t)}_{=1} \psi_{\mathcal{H}_0}(\mathbf{x}, t) u(t, t') \psi_{\mathcal{H}_0}^\dagger(\mathbf{x}', t') u(t', -\infty) \rangle \\ &= -\frac{i}{\hbar} \langle u^\dagger(\infty, -\infty) \times u(\infty, t) \psi_{\mathcal{F}_0}(\mathbf{x}, t) u(t, t') \psi_{\mathcal{H}_0}^\dagger(\mathbf{x}', t') u(t', -\infty) \rangle, \end{aligned} \quad (2.69)$$



¹²⁹ Taking the Hermitian conjugate of a product of operators reverses the order of the operators. Thus, the Hermitian conjugate of a time-ordered product of operators is anti-time-ordered.

¹²¹⁰ Up to some infinite phase. For details on the ground state formalism, see Ref. [58].

Figure 2.1 - The Keldysh contour

where we made extensive use of Eqs. (2.63). Note that all operators in the second term are time-ordered. We can therefore simplify it as

$$\begin{aligned}
 u(\infty, t) \psi_{\mathcal{H}_0}(\mathbf{x}, t) u(t, t') \psi_{\mathcal{F}_0}^\dagger(\mathbf{x}', t') u(t', -\infty) &= \\
 &= \mathcal{T} \left\{ u(\infty, t) \psi_{\mathcal{F}_0}(\mathbf{x}, t) u(t, t') \psi_{\mathcal{H}_0}^\dagger(\mathbf{x}', t') u(t', -\infty) \right\} \\
 &= \mathcal{T} \left\{ u(\infty, -\infty) \psi_{\mathcal{H}_0}(\mathbf{x}, t) \psi_{\mathcal{H}_0}^\dagger(\mathbf{x}', t') \right\} \\
 &= \mathcal{T} \left\{ e^{-\frac{i}{\hbar} \int_{-\infty}^{\infty} dt_1 \mathcal{H}'_{\mathcal{H}_0}(t_1)} \psi_{\mathcal{H}_0}(\mathbf{x}, t) \psi_{\mathcal{F}_0}^\dagger(\mathbf{x}', t') \right\} \quad (2.70)
 \end{aligned}$$

Here, we made use of the fact that the operators commute within the time-ordering operator, and that the time-evolution operator u always contains an even number of Fermion operators. Thus, no additional minus signs appear when interchanging the operators. Using Eq. (2.62), we finally arrive at

$$\begin{aligned}
 G^c(\mathbf{x}, t, \mathbf{x}', t') &= \\
 &= -\frac{i}{\hbar} \left\langle \tilde{\mathcal{T}} \left\{ e^{-\frac{i}{\hbar} \int_{-\infty}^{\infty} dt_1 \mathcal{H}'_{\mathcal{H}_0}(t_1)} \right\} \mathcal{T} \left\{ e^{-\frac{i}{\hbar} \int_{-\infty}^{\infty} dt_1 \mathcal{H}'_{\mathcal{F}_0}(t_1)} \psi_{\mathcal{F}_0}(\mathbf{x}, t) \psi_{\mathcal{F}_0}^\dagger(\mathbf{x}', t') \right\} \right\rangle \quad (2.71)
 \end{aligned}$$

In the ground state theory, we would only have the second, time-ordered term. However, we can formally bring the result (2.71) into a form equivalent to the ground state Green's function formalism, by introducing the ordering $\mathcal{T}_C$ along a contour C , with a part $\vec{C}$ from $-\infty$ to ∞ , and a part $\overleftarrow{C}$ from ∞ to $-\infty$, as depicted in Fig. 2.1. The contour C is often referred to as the Keldysh contour.

We will in general denote a time on the contour C as τ , and the corresponding real time as t . When it is necessary to distinguish whether a real time variable t belongs to the upper or lower branch, i.e. $t \in \vec{C}$ or $t \in \overleftarrow{C}$, we will use the notation $t^\rightarrow$ and $t^\leftarrow$ to indicate the respective branch. Note that ordering along the contour $\vec{C}$ corresponds to the normal time-ordering, whereas ordering along $\overleftarrow{C}$ corresponds to anti-time ordering. Therefore we find:

$$\begin{aligned}
 G^c(\mathbf{x}, t, \mathbf{x}', t') &= \\
 &= -\frac{i}{\hbar} \left\langle \mathcal{T}_{\overleftarrow{C}} \left\{ e^{-\frac{i}{\hbar} \int_{\overleftarrow{C}} dt_1 \mathcal{H}'_{\mathcal{H}_0}(t_1)} \right\} \mathcal{T}_{\vec{C}} \left\{ e^{-\frac{i}{\hbar} \int_{\vec{C}} dt_1 \mathcal{H}'_{\mathcal{H}_0}(t_1)} \psi_{\mathcal{H}_0}(\mathbf{x}, t^\rightarrow) \psi_{\mathcal{H}_0}^\dagger(\mathbf{x}', t'^\rightarrow) \right\} \right\rangle \\
 &= -\frac{i}{\hbar} \left\langle \mathcal{T}_C \left\{ e^{-\frac{i}{\hbar} \int_C d\tau_1 \mathcal{H}'_{\mathcal{H}_0}(\tau_1)} \psi_{\mathcal{H}_0}(\mathbf{x}, t^\rightarrow) \psi_{\mathcal{F}_0}^\dagger(\mathbf{x}', t'^\rightarrow) \right\} \right\rangle \quad \text{for } t, t' \in \vec{C}. \quad (2.72)
 \end{aligned}$$

Note that the definition of the contour-ordering operator $\mathcal{T}_C$ also includes a minus sign for every interchange of fermionic operators.

Up to now, the contour-ordering is only a neat trick to compactify the notation: Eq. (2.72) is only a formal paraphrase of Eq. (2.71). However, the power of the contour ordering lies in the possibility to describe several Green's functions at once in a unified way. In order to demonstrate this, we next consider the lesser Green's function:

$$\begin{aligned}
G^<(\mathbf{x}, t, \mathbf{x}', t') &= \frac{i}{\hbar} \langle \psi_{\mathcal{H}}^{\dagger}(\mathbf{x}', t') \psi_{\mathcal{H}}(\mathbf{x}, t) \rangle \\
&= \frac{i}{\hbar} \langle u^{\dagger}(t', -\infty) \psi_{\mathcal{F}_0}^{\dagger}(\mathbf{x}', t') u(t', -\infty) u^{\dagger}(t, -\infty) \psi_{\mathcal{H}_0}(\mathbf{x}, t) u(t, -\infty) \rangle \\
&= \frac{i}{\hbar} \langle u^{\dagger}(t', -\infty) \psi_{\mathcal{H}_0}^{\dagger}(\mathbf{x}', t') u(t', -\infty) u^{\dagger}(\infty, -\infty) u(\infty, -\infty) \\
&\quad u^{\dagger}(t, -\infty) \psi_{\mathcal{H}_0}(\mathbf{x}, t) u(t, -\infty) \rangle \\
&= \frac{i}{\hbar} \langle u^{\dagger}(t', -\infty) \psi_{\mathcal{H}_0}^{\dagger}(\mathbf{x}', t') u^{\dagger}(\infty, t') u(\infty, t) \psi_{\mathcal{F}_0}(\mathbf{x}, t) u(t, -\infty) \rangle \\
&= \frac{i}{\hbar} \langle \mathcal{T}_{\vec{C}} \left\{ e^{-\frac{i}{\hbar} \int_{\vec{C}} dt_1 \mathcal{H}'_{\mathcal{H}_0}(t_1)} \psi_{\mathcal{H}_0}^{\dagger}(\mathbf{x}', t'^{\leftarrow}) \right\} \mathcal{T}_{\vec{C}} \left\{ e^{-\frac{i}{\hbar} \int_{\vec{C}} dt_1 \mathcal{H}'_{\mathcal{H}_0}(t_1)} \psi_{\mathcal{H}_0}(\mathbf{x}, t^{\rightarrow}) \right\} \right\rangle \\
&= -\frac{i}{\hbar} \langle \mathcal{T}_C \left\{ e^{-\frac{i}{\hbar} \int_C d\tau_1 \mathcal{H}'_{\mathcal{H}_0}(\tau_1)} \psi_{\mathcal{F}_0}(\mathbf{x}, t^{\rightarrow}) \psi_{\mathcal{H}_0}^{\dagger}(\mathbf{x}', t'^{\leftarrow}) \right\} \rangle \quad \text{for } t \in \vec{C}, t' \in \overleftarrow{C}, \quad (2.73)
\end{aligned}$$

where the minus sign in the last line is due to the interchange of the field operators. Thus, the contour-ordered expression for $G^<$ is formally identical to G^c , only the time arguments belong to different branches of the contour.

The fact that the contour ordering provides a unified expression for several Green's functions which only contain field operators in the interaction picture motivates the definition of the contour-ordered Green's function,

$$G(\mathbf{x}, \tau, \mathbf{x}', \tau') = -\frac{i}{\hbar} \langle \mathcal{T}_C \psi_{\mathcal{H}}(\mathbf{x}, \tau) \psi_{\mathcal{H}}^{\dagger}(\mathbf{x}', \tau') \rangle \quad \text{for } \tau, \tau' \in C \quad (2.74)$$

This contour-ordered Green's function contains several of the already known Green's functions:

$$G(\mathbf{x}, \tau, \mathbf{x}', \tau') = \begin{cases} G^c(\mathbf{x}, t, \mathbf{x}', t') & \text{for } t, t' \in \vec{C} \\ G^<(\mathbf{x}, t, \mathbf{x}', t') & \text{for } t \in \vec{C}, t' \in \overleftarrow{C} \\ G^>(\mathbf{x}, t, \mathbf{x}', t') & \text{for } t \in \overleftarrow{C}, t' \in \vec{C} \\ G^{\text{ac}}(\mathbf{x}, t, \mathbf{x}', t') & \text{for } t, t' \in \overleftarrow{C} \end{cases} \quad (2.75)$$

Here, G^{ac} is the anti-causal (anti-time-ordered) Green's function.

Analogous to the derivations above, we may derive an expression for the contour-ordered Green's function only in terms of field operators in the interaction picture

$$G(\mathbf{x}, \tau, \mathbf{x}', \tau') = -\frac{i}{\hbar} \frac{\langle \mathcal{T}_C \left\{ e^{-\frac{i}{\hbar} \int_C d\tau_1 \mathcal{H}'_{\mathcal{H}_0}(\tau_1)} \psi_{\mathcal{H}_0}(\mathbf{x}, \tau) \psi_{\mathcal{H}_0}^{\dagger}(\mathbf{x}', \tau') \right\} \rangle}{\langle \mathcal{T}_C \left\{ e^{-\frac{i}{\hbar} \int_C d\tau_1 \mathcal{H}'_{\mathcal{H}_0}(\tau_1)} \right\} \rangle} \quad \text{for } \tau, \tau' \in C. \quad (2.76)$$

Note that the denominator in this expression is simply unity. This expression allows for the formulation of a perturbation theory analogous to the ground state formalism¹¹. In particular, the Green's function can be expressed in the form of a Dyson equation [58],

$$\begin{aligned}
G(\mathbf{x}, \tau, \mathbf{x}', \tau') &= G_0(\mathbf{x}, \tau, \mathbf{x}', \tau') + \\
&\quad \int_C d\tau_1 \int_C d\tau_2 \int d\mathbf{x}_1 d\mathbf{x}_2 G_0(\mathbf{x}, \tau, \mathbf{x}_1, \tau_1) \tilde{\Sigma}(\mathbf{x}_1, \tau_1, \mathbf{x}_2, \tau_2) G(\mathbf{x}_2, \tau_2, \mathbf{x}', \tau') \quad (2.78)
\end{aligned}$$

where $\tilde{\Sigma}$ is the proper self-energy consisting only of irreducible diagrams [58]. For the non-equilibrium theory, it is useful to divide the proper self-energy $\tilde{\Sigma}$ into two parts,

$$\tilde{\Sigma}(\mathbf{x}_1, \tau_1, \mathbf{x}_2, \tau_2) = V(\mathbf{x}_1, \mathbf{x}_2, \tau_1) \delta_C(\tau_1 - \tau_2) + \Sigma(\mathbf{x}_1, \tau_1, \mathbf{x}_2, \tau_2) \quad (2.79)$$

where δ_C is the delta function on the contour C . Here, V depends only on at most one time on the contour C , whereas Σ depends explicitly on two times on C . Such a V may

¹¹ Eq. (2.76) is formally equivalent to the corresponding expression of the ground state Green's function formalism, Eq. (2.67). Thus, we will not present a detailed derivation of a non-equilibrium perturbation theory here, but merely outline briefly how this perturbation arises from Eq. (2.76).

All the technical details are equivalent to the ground state formalism, and the reader is referred to the extensive literature reviewing this subject [58, 62, 74].

As for the ground state, a perturbation theory is developed by expanding the exponential

$$e^{-\frac{i}{\hbar} \int_C d\tau_1 \mathcal{H}'_{\mathcal{H}_0}(\tau_1)} = 1 - \frac{i}{\hbar} \int_C d\tau_1 \mathcal{H}'_{\mathcal{H}_0}(\tau_1) + \frac{1}{2} \frac{1}{\hbar^2} \int_C \int_C d\tau_1 d\tau_2 \mathcal{H}'_{\mathcal{H}_0}(\tau_1) \mathcal{H}'_{\mathcal{H}_0}(\tau_2) + \dots, \quad (2.77)$$

as a power series in terms of the perturbing Hamiltonian $\mathcal{H}_{\mathcal{H}_0}(\tau)$, and thus also in terms of field operators $\psi_{\mathcal{H}_0}, \psi_{\mathcal{H}_0}^\dagger$ in the interaction picture. The contour-ordered Green's function is then given as a sum of expectation values involving a contour-ordered product of an even numbers of field operators, $\langle \mathcal{I}_C \psi_{\mathcal{H}_0}^{(\dagger)}(\mathbf{x}_1, \tau_1) \dots \psi_{\mathcal{H}_0}^{(\dagger)}(\mathbf{x}_n, \tau_n) \rangle$, and a number of contour integrations $\int_C$. Invoking Wick's theorem, these expectation values are then reduced to a product of expectation values involving only two field operators, $\langle \mathcal{T}_C \psi_{\mathcal{H}_0}(\mathbf{x}, \tau) \psi_{\mathcal{H}_0}^\dagger(\mathbf{x}', \tau') \rangle = G_0(\mathbf{x}, \tau, \mathbf{x}', \tau')$. Thus, the Green's function of the interacting system, G , can be expressed solely in terms of Green's functions of the noninteracting system, G_0 . The systematics of the power series expansion of $e^{-\frac{i}{\hbar} \int_C d\tau_1 \mathcal{H}'_{\mathcal{H}_0}(\tau_1)}$ together with Wick's theorem is often expressed in terms of Feynman diagrams which are a convenient way of graphically representing all the different terms arising in this expansion. Because of the formal equivalence of the contour-ordered Green's function to the ground state formalism, the Feynman rules for the contour-ordered Green's function are simply obtained by replacing real time variables with a time variable on the Keldysh contour C . Thus, all results from the ground state formalism, such as the resummation of diagrams in the form of a Dyson equation, can also be applied to the non-equilibrium formalism.

Finally, a comment on the applicability of Wick's theorem to a contour-ordered product of operators. Note that there are two versions of Wick's theorem: an operator identity that is only applicable to the ground state formalism and an identity involving only expectation values, that also holds for example in the finite-temperature formalism. For constructing a perturbation theory, it is enough to consider expectation values of contour-ordered products. In order to prove Wick's theorem it is therefore more convenient to follow the proof of Wick's theorem for finite temperatures, as given for example in chapter 24 of Ref. [58]. Indeed, this proof only requires some time-like ordering and can be applied directly to the non-equilibrium case.

arise for example from a time-dependent or time-independent potential, or within the Hartree-Fock approximation to the Coulomb-interaction [58]. In contrast to Σ , which is an expression involving two-time Green's functions, V only contains potentials or observables such as the particle density, arising from Green's functions with equal time arguments in the perturbation expansion. The Dyson equation then takes the form

$$G = G_0 + G_0 V G + G_0 \Sigma G \quad (2.80)$$

where we compactified the notation by suppressing the arguments of the Green's function

and the integrations. We will continue to do so in the following, as long as this compactified notation is unambiguous. The Dyson equation can also be written in an alternative form as

$$G = G_0 + GVG_0 + G\Sigma G_0 \quad (2.81)$$

In the ground state formalism, the distinction between V and Σ is not important. However, it will be in the case of the non-equilibrium theory, when the perturbative terms in the contour-ordered Green's function are recast into a form involving only real-time Green's functions, such as $G^<$ and G^r ,¹² as presented in the next section.

19.1.7 2.4.3 Real-time formulation

The contour-ordered Green's function allows for the development of a perturbation theory, but does not have any clear physical meaning. It is thus advantageous to rewrite the Dyson equations (2.80) and (2.81) in terms of the real-time Green's functions G^r , G^a , $G^<$, and $G^>$ that can be directly used, e.g. for calculating an observable. In order to do so, we follow an approach by Langreth [72, 75].

The perturbation expansion contains terms that depend on two times on the contour: These can be single contour-ordered Green's functions, or products of contour-ordered Green's functions with internal times integrated out. In order to treat these terms on equal footing, it is useful to consider a general two-time function $A(\tau, \tau')$ on the contour, $\tau, \tau' \in C$. For such function, we define the corresponding real-time functions

$$A(\tau, \tau') = \begin{cases} A^c(t, t') & \text{for } t, t' \in \vec{C}, \\ A^<(t, t') & \text{for } t \in \overleftarrow{C}, t' \in \vec{C}, \\ A^>(t, t') & \text{for } t \in \vec{C}, t' \in \overleftarrow{C} \\ A^{ac}(t, t') & \text{for } t, t' \in \overleftarrow{C} \end{cases} \quad (2.82)$$

$$A^r(t, t') = \theta(t - t') [A^>(t, t') - A^<(t, t')], \text{ and} \quad (2.83)$$

$$A^a(t, t') = -\theta(t' - t) [A^>(t, t') - A^<(t, t')] \quad (2.84)$$

in analogy to the definitions (2.75), (2.19a), and (2.19b). This formal definition coincides with previous Green's function identities, if $A(\tau, \tau')$ is a single contour-ordered Green's function.

In the Dyson equation, we encounter terms of the form

$$C(\tau_1, \tau_2) = \int_C d\tau' A(\tau_1, \tau') B(\tau', \tau_2) \quad (2.85)$$

where A, B and C are two-time functions on the contour C . Specifying which branch the times τ_1 and τ_2 belong to, we can then obtain the desired real-time functions. For example, we obtain

$$\begin{aligned} C^<(t_1, t_2) &= C(t_1^{\rightarrow}, t_2^{\leftarrow}) = \int_C d\tau' A(t_1^{\rightarrow}, \tau') B(\tau', t_2^{\leftarrow}) \\ &= \int_{-\infty}^{\infty} dt' A(t_1^{\rightarrow}, t'^{\rightarrow}) B(t'^{\rightarrow}, t_2^{\leftarrow}) + \int_{\infty}^{-\infty} dt' A(t_1^{\rightarrow}, t'^{\leftarrow}) B(t'^{\leftarrow}, t_2^{\leftarrow}) \\ &= \int_{-\infty}^{\infty} dt' A^c(t_1, t') B^<(t', t_2) - A^<(t_1, t') B^{ac}(t', t_2). \end{aligned} \quad (2.86)$$

¹²¹² In essence, this distinction between V and Σ is necessary, because the definition of "lesser" or "retarded" involves two times on the contour C .

Using the identities $A^c = A^< + A^r$ and $B^{ac} = B^< - B^a$, we finally arrive at

$$C^<(t_1, t_2) = \int_{-\infty}^{\infty} dt' A^r(t_1, t') B^<(t', t_2) + A^<(t_1, t') B^a(t', t_2) \quad (2.87a)$$

an expression which only depends on real-time functions and integrations. In a similar fashion, we obtain:

$$C^>(t_1, t_2) = \int_{-\infty}^{\infty} dt' A^r(t_1, t') B^>(t', t_2) + A^>(t_1, t') B^a(t', t_2) \quad (2.87b)$$

$$C^r(t_1, t_2) = \int_{-\infty}^{\infty} dt' A^r(t_1, t') B^r(t', t_2) \quad (2.87c)$$

$$C^a(t_1, t_2) = \int_{-\infty}^{\infty} dt' A^a(t_1, t') B^a(t', t_2) \quad (2.87d)$$

These identities are commonly called analytic continuation or Langreth rules. Applying the Langreth rules repeatedly, any product of contour ordered Green's functions G can be reduced to an expression involving only $G^{<,>}$ and $G^{r,a}$. In particular, for a product of three two-time functions on the contour,

$$D = \int_C \int_C ABC, \quad (2.88)$$

we find

$$D^{<,>} = A^r B^r C^{<,>} + A^r B^{<,>} C^a + A^{<,>} B^a C^a, \quad (2.89a)$$

$$D^{r,a} = A^{r,a} B^{r,a} C^{r,a}. \quad (2.89b)$$

Here we again switched to a more compact notation, where we do not explicitly write all integration and Green's function arguments.

When applying the Langreth rules to the Dyson equations (2.80) and (2.81), it is essential to realize that the term

$$\int_C d\tau_1 G_0(\tau, \tau_1) V(\tau_1) G(\tau_1, \tau') \quad (2.90)$$

is (in terms of applying the Langreth rules) of the form $C = \int_C AB$. For the retarded Green's function, we then obtain the Dyson equations

$$G^r = G_0^r + G_0^r V G^r + G_0^r \Sigma^r G^r \quad (2.91)$$

and

$$G^r = G_0^r + G^r V G_0^r + G^r \Sigma^r G_0^r. \quad (2.92)$$

From the nature of the Langreth rules, i.e. that a retarded function can always be expressed only in terms of retarded functions, it might seem that this equation can only contain retarded functions, for any choice of the proper self-energy $\tilde{\Sigma}$. However, due to the nature of the perturbation expansion, in general there are terms containing $G^<$ or $G^>$ already in the contour-ordered expression¹³, and thus V and Σ^r may also contain $G^<$ and $G^>$.

For $G^<$, we obtain from the Langreth rules

$$G^< = G_0^< (1 + V G^a + \Sigma^a G^a) + G_0^r \Sigma^< G^a + (G_0^r V + G_0^r \Sigma^r) G^<. \quad (2.93)$$

Using Eq. (2.92) and $(1 + G^r V + G^r \Sigma^r)(1 + G_0^r V + G_0^r \Sigma^r) = 1$, we finally obtain

$$G^< = (1 + G^r V + G^r \Sigma^r) G_0^< (1 + V G^a + \Sigma^a G^a) + G^r \Sigma^< G^a. \quad (2.94)$$

Commonly, the first term of this expression is neglected: From the Dyson equation for the retarded function (2.92) and equivalently for the advanced function we can rewrite this term as $G^r \overleftarrow{G}_0^{-1} G_0^< \overrightarrow{G}_0^{-1} G^a$, where the arrows indicate on which function $G_0^<$ acts. Reversing this operator by partial integration then yields a term $\overrightarrow{G}_0^{-1} G_0^< = 0$, according to Eq. (2.38), and thus only a boundary term at times $\pm\infty$ is left. This boundary term can usually be neglected, as the initial correlations decay with time [76], except for certain cases, that we will comment on further in the next section. Hence, we obtain a compact expression for $G^<$,

$$G^< = G^r \Sigma^< G^a, \quad (2.95)$$

which is commonly referred to as the Keldysh equation.

Up to now, all expressions were for Green's functions in the time domain. In steadystate, i.e. when the properties of the (non-equilibrium) system are constant in time and the Green's functions only depend on time differences, it is more convenient to consider the real-time Green's functions in the energy domain. Note that this is not possible for the contour-ordered Green's function, as the contour C has two different branches.

Working in the energy domain is especially useful for steady state, as all expressions for real-time Green's functions involve convolutions that are reduced to simple multiplications: The Fourier transform of $A(t - t') = \int_{-\infty}^{\infty} dt_1 B(t - t_1) C(t_1 - t')$ yields $A(E) = B(E)C(E)$. This gives, for example, the Keldysh equation in the energy domain as

$$G^<(E) = G^r(E) \Sigma^<(E) G^a(E), \quad (2.96)$$

where again we suppressed the remaining spatial integrations. Our compact notation is particularly useful in this regard: All expressions can be understood to be valid both for the time and the energy domain. In the remainder of this work, we restrict ourselves to steady state and work exclusively in the energy domain.

During the course of the discussion about the Green's function theory, we have introduced the self-energy contributions V and Σ as general concepts. In the next section, we will derive explicit expressions for these objects.

19.1.8 2.5 Transport in tight-binding models

19.1.9 2.5.1 Green's functions in tight-binding

The non-equilibrium Green's function formalism of the previous chapter in principle provides a systematic prescription to perform a transport calculation for any system. However, even in the case of a non-interacting system, an exact analytic solution is usually not possible. Therefore, numerical computations are often the method of choice.

As presented above, the Green's functions obey differential or integral equations and depend on variables $\mathbf{x}, \mathbf{x}'$ with a continuous range of values. These continuous degrees of freedom are difficult to treat on a computer, and thus a numerical solution is in general only attempted

¹²¹³ The perturbation expansion contains terms of the form $\mathcal{T}_C H'_{\mathcal{H}_0}(\tau_1) \dots H'_{\mathcal{H}_0}(\tau_n) \psi_{\mathcal{H}_0}(\tau) \psi_{\mathcal{H}_0}^\dagger(\tau')$. The perturbation Hamiltonian $H'_{\mathcal{H}_0}(\tau_1)$ itself contains several field operators, at the same time τ_1 . These cannot be ordered by T_C , but are kept as they are ordered in the expression for $H'_{\mathcal{H}_0}(\tau_1)$. Then, when applying Wick's theorem, field operators at the same time are paired, and the result is $G_0^<$ or $G_0^>$. In Feynman diagrams, such terms appear, for example, in the form of Fermion bubbles, and are already present in the ground state formalism [58].

within a discrete basis set $\{|i\rangle\}$. In this basis set, the single-particle Hamiltonian can be written as

$$H = \sum_{ij} H_{ij} |i\rangle \langle j|, \quad (2.97)$$

where $H_{ij} = \langle i|H|j\rangle$. Very often, only few of the matrix elements H_{ij} are nonzero. Such tight-binding representations of the Hamiltonian are ubiquitous in quantum transport calculations and can arise from finite differences [77-79], from the finite element method [80], from atomic orbitals in empirical tight-binding [81-83] or Kohn-Sham orbitals within density functional theory [84-86]. In this work, we will employ both the finite difference approximation to the effective mass Schrödinger equation (Chapter

5 and Appendix F) and an empirical tight-binding model for graphene (Chapters 6-7 and Appendix G). In both cases, the states $|i\rangle$ form an orthonormal basis. In the following we therefore restrict the discussion to orthogonal tight-binding models.

The corresponding non-interacting many-particle Hamiltonian is then given as

$$\mathcal{H} = \sum_{ij} H_{ij} c_i^\dagger c_j \quad (2.98)$$

and is simply obtained by replacing the states $|i\rangle$ with the corresponding annihilation and creation operators [70]. Instead of the real-space Green's functions, that are given as expectation values of two field operators $\psi(\mathbf{x})$ and $\psi^\dagger(\mathbf{x}')$, we now consider the Green's functions in the representation of the discrete basis set $|i\rangle$. These are obtained by replacing $\psi(\mathbf{x}) \rightarrow c_i$ and $\psi^\dagger(\mathbf{x}') \rightarrow c_j^\dagger$. For example, the lesser Green's function in the basis set representation is then given as

$$G^<(i, t, j, t') = \frac{i}{\hbar} \langle c_j^\dagger(t') c_i(t) \rangle. \quad (2.99)$$

Note that in many cases, the basis states $|i\rangle$ are sufficiently localized at a point $\mathbf{x}_i$. In these cases, c_i plays the role of $\psi(\mathbf{x}_i)$ and the interpretation of the Green's functions remains unchanged. However, in any case, the Green's functions in tight-binding representation can be used to calculate any desired observable (see Appendix A).

Because of the discreteness of the basis set, the Green's functions can be interpreted as matrices, and we will use the notation $G_{ij}(t, t') = G(i, t, j, t')$ or $G_{ij}(E) = G(i, j; E)$ to emphasize this fact. All the results of the previous section also carry over to the Green's function in basis set representation, with integrations over $\mathbf{x}$ replaced by sums over basis states $i, \int d\mathbf{x}_1 A(\mathbf{x}, \mathbf{x}_1) B(\mathbf{x}_1, \mathbf{x}') \rightarrow \sum_k A_{ik} B_{kj}$. Therefore, in tight-binding representation, the results of the previous section can be directly interpreted as matrix equations. This is another advantage of the compact notation employed there.

19.1.10 2.5.2 Transport equations

We consider a non-interacting system consisting of a scattering region (also called central region) and N_L leads, as schematically depicted in Fig. 2.2. The leads and the scattering region are assumed to be disconnected in the far past, and the hopping is switched on adiabatically at later times [61]. The Hamiltonian of the system is then given as

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}' \quad (2.100)$$

Here, $\mathcal{H}_0$ is the Hamiltonian of the disconnected subsystems,

$$\mathcal{H}_0 = \sum_{l=1}^{N_L} \mathcal{H}_l + \mathcal{H}_S, \quad (2.101)$$

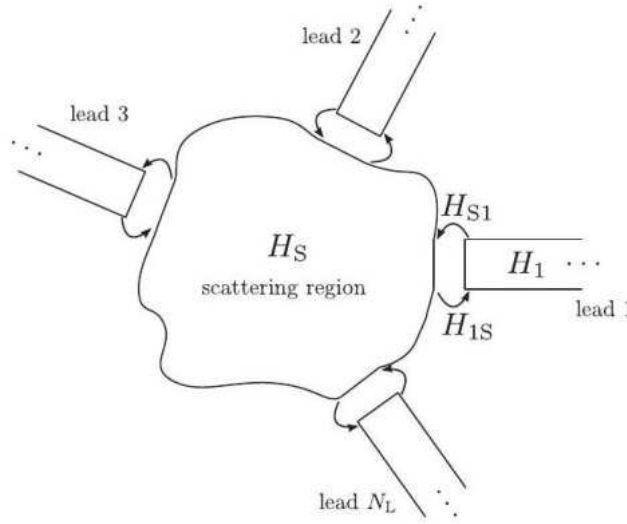


Figure 2.2 - In a transport calculation, the system under consideration consists in general of a scattering region attached to a number of leads.

where $\mathcal{H}_l$ is the Hamiltonian of lead l and $\mathcal{H}_S$ of the scattering region, and $\mathcal{H}'$ gives the coupling of the leads to the scattering region,

$$\mathcal{H}' = \sum_{l=1}^{N_L} \mathcal{H}_{lS} + \mathcal{H}_{Sl} \quad (2.102)$$

where $\mathcal{H}_{lS}$ and $\mathcal{H}_{Sl}$ describe the hopping from the scattering region to lead l and vice versa. Since $\mathcal{H}$ is a Hermitian Hamiltonian, we have $\mathcal{H}_{lS} = \mathcal{H}_{Sl}^\dagger$. The Hamiltonians of the isolated parts and of the hopping are given as

$$\mathcal{H}_S = \sum_{ij} H_{S,ij} c_{S,i}^\dagger c_{S,j}, \quad (2.103)$$

$$\mathcal{H}_l = \sum_{ij} H_{l,ij} c_{l,i}^\dagger c_{l,j} \quad (2.104)$$

$$\mathcal{H}_{lS} = \sum_{ij} H_{lS,ij} c_{l,i}^\dagger c_{S,j} \quad (2.105)$$

Here, the creation and annihilation operators are marked with a subscript, depending on whether they act in the scattering region ($c_{S,i}, c_{S,i}^\dagger$) or in lead l ($c_{l,i}, c_{l,i}^\dagger$). Note that the Hamiltonian is marked with a single subscript S or l only, if it contains operators within a single subsystem, and with two subscripts otherwise. We will employ the same notation for the Green's function later.

In the far past, the isolated subsystems are assumed to be in thermal equilibrium with electrochemical potential μ_l in lead l and μ_S in the scattering region¹⁴. The

electrochemical potentials μ_l may be different for every lead, and this difference will be the driving force for non-equilibrium. Such a difference may be for example due to different electron densities, temperatures or electrical potentials in the leads. Here, we consider in particular the

¹²¹⁴ The electrochemical potential of the leads μ_l is determined by the number of particles in lead l and

case of an applied bias, such that $\mu_l = \mu_0 + qV_l$, where V_l is the electrical potential in lead l , q the charge of the particles, and μ_0 the equilibrium chemical potential. In the following, we will specialize to electronic transport, where $q = -e$ is the charge of an electron.

We now apply the results of the non-equilibrium Green's function formalism of the previous section for the perturbation Hamiltonian $\mathcal{F}'$ and obtain the Dyson equation for the contour-ordered Green's function,

$$G = G_0 + G_0 V G = G_0 + G V G_0 \quad (2.106)$$

where $V = \sum_l H_{lS} + H_{Sl}$. Note that this equation is exact to all orders in the perturbation, as $\mathcal{H}'$ is simply a (off-diagonal) potential, and the proper self-energy given by a single term V [60].

In particular, we are interested in the Green's function within the scattering region, G_S . Since the leads and the scattering region are decoupled in the far past, there are no correlations between them, and as a consequence $G_{0,Sl} = G_{0,l}$ $s = 0$. Using this fact, we find

$$G_S = G_{0,S} + G_{0,S} \sum_l H_{Sl} G_{l,S} \quad (2.107)$$

and

$$G_{l,S} = G_{0,l} H_{l,S} G_S \quad (2.108a)$$

$$G_{Sl} = G_S H_{Sl} G_{0,l} \quad (2.108b)$$

Inserting Eq. (2.108a) into (2.107), we obtain

$$G_S = G_{0,S} + G_{0,S} \Sigma G_S, \quad (2.109)$$

where $\Sigma = \sum_l H_{Sl} G_{0,l} H_{lS}$ is the self-energy due to the leads.

We can now apply the analytic continuation rules of the previous section to obtain the real-time Green's functions. For the retarded Green's function, Eq. (2.91), we find $G_S^r = G_{0,S}^r + G_{0,S}^r \Sigma^r G_S^r$. The analytic continuation of the contour-ordered self-energy Σ is particularly simple in this case, as it consists only of a single Green's function, and we obtain $\Sigma^r = \sum_l \Sigma_l^r = \sum_l H_{Sl} G_{0,l}^r H_{lS}$, where Σ_l^r is the retarded self-energy due to a single lead l . In a tight-binding model, the hopping matrices H_{lS} only couple states at the boundary between the scattering region and the leads. Thus, Σ^r only depends on

the values of $G_{0,l}^r$ on the surface of lead l , decoupled from the scattering region. We denote the values of $G_{0,l}^r$ on the surface as the surface Green's function g_l^r , such that

$$\Sigma^r = \sum_l H_{Sl} g_l^r H_{l,S} \quad (2.110)$$

Using furthermore $G_{0,S} = (E - H_S + i\eta)^{-1}$, Eq. (2.43), we arrive at the final expression for the retarded Green's function,

$$G_S^r = (E - H_S - \Sigma^r + i\eta)^{-1} \quad (2.111)$$

Since the scattering region is finite, so are the matrices H_S and Σ^r , and the problem of calculating the retarded Green's function has been reduced to the problem of inverting a finite matrix. The infinitesimal $i\eta$ can be neglected in a numerical calculation, if the self-energy of

¹²the electrical potential arising from some applied voltage. In contrast, the electrochemical potential of the scattering region μ_S is somewhat ill-defined, as there is no connection to a reservoir with a well-defined electrical potential. Fortunately, the electrochemical potential of the scattering region does not enter the final result.

the leads has a nonzero imaginary part, $\Im(\Sigma^r) \neq 0$. This is the case, if there are propagating states in the leads at the energy E .

The lesser Green's function of the scattering region, Eq. (2.95), is given as

$$G_S^< = G_S^r \Sigma^< G_S^a \quad (2.112)$$

where $\Sigma^< = \sum_l H_{Sl} g_l^< H_{lS}$. Since the leads were assumed to be in equilibrium in the far past, we can employ Eq. (2.52a) to calculate $\Sigma^<$:

$$\begin{aligned} \Sigma^< &= \sum_l H_{Sl} g_l^< H_{lS} \\ &= \sum_l H_{Sl} (-f_0(E, \mu_l)) (g_l^r - g_l^a) H_{lS} \\ &= \sum_l i f_0(E, \mu_l) \Gamma_l \end{aligned} \quad (2.113)$$

where we defined $\Gamma_l = i(\Sigma_l^r - \Sigma_l^a) = i(\Sigma_l^r - (\Sigma_l^r)^\dagger)$.

These Green's functions can now be used to calculate the current through the system. For this, we define the current in lead l as $I_l = \frac{d}{dt} Q_l$, where Q_l is the charge in lead l . With this definition, a positive current $I_l > 0$ corresponds to an increase of positive charge, and thus to electrons flowing out of the lead. With $Q_l = -e \langle \mathcal{N}_l \rangle$, where $\mathcal{N}_l = \sum_i c_{l,i}^\dagger c_{l,i}$ is the number operator in lead l , we obtain through direct calculation

$$\begin{aligned} I_l &= -e \left\langle \frac{d}{dt} \mathcal{N}_l \right\rangle = \frac{ie}{\hbar} \langle [\mathcal{N}_l, \mathcal{H}] \rangle \\ &= \frac{ie}{\hbar} \sum_{ij} H_{lS,ij} \langle c_{l,i}^\dagger c_{S,j} \rangle - H_{Sl,ji} \langle c_{S,j}^\dagger c_{l,i} \rangle \\ &= e \sum_{ij} H_{lS,ij} G_{Sl,ji}^< - H_{Sl,ji} G_{lS,ij}^< \\ &= e \text{Tr} (H_{lS} G_{Sl}^< - H_{Sl} G_{lS}^<) \end{aligned} \quad (2.114)$$

Here, we used the definition of the lesser Green's function, Eq. (2.19c), and the definition of the trace.

Note that the expressions are up to now still in the time-domain. In steady-state¹⁵, the current is constant in time, $I_l(t) = I_l(0) = I_l$. Hence, it is possible to work in the energy domain and write

$$I_l = I_l(0) = \frac{e}{\hbar} \int dE \text{Tr} (H_{lS} G_{Sl}^<(E) - H_{Sl} G_{lS}^<(E)) \quad (2.115)$$

For brevity, in the following we will again omit the dependence of the Green's functions on energy E . Inserting the analytic continuation of Eq. (2.108) we obtain

$$\begin{aligned} I_l &= \frac{e}{\hbar} \int dE \text{Tr} (H_{lS} G_S^r H_{Sl} G_{0,l}^< + H_{lS} G_S^< H_{Sl} G_{0,l}^a \\ &\quad - H_{Sl} G_{0,l}^r H_{lS} G_S^< - H_{Sl} G_{0,l}^< H_{lS} G_S^a) \\ &= \frac{e}{\hbar} \int dE \text{Tr} ((\Sigma_l^a - \Sigma_l^r) G_S^< + \Sigma_l^< (G_S^r - G_S^a)) \\ &= \frac{e}{\hbar} \int dE \text{Tr} (i \Gamma_l (G_S^< + f_0(E, \mu_l) (G_S^r - G_S^a))), \end{aligned} \quad (2.116)$$

where we made use of the cyclic invariance of the trace and Eq. (2.113). The expression for the current (2.116) was first derived by Meir and Wingreen [87], and holds also for a system with interactions in the scattering region. For the case of non-interacting particles - or more precisely for systems where $\Sigma^<$ only has contributions from the leads as in Eq. (2.113) - we can use Eqs. (2.112) and (2.23) to further simplify this expression:

$$\begin{aligned}
I_l &= \frac{e}{h} \int dE \operatorname{Tr} (i\Gamma_l \sum_{l'} G_S^r i f_0(E, \mu_{l'}) \Gamma_{l'} G_S^a + i f_0(E, \mu_l) \Gamma_l G_S^r \underbrace{(\Sigma^> - \Sigma^<)}_{=-i \sum_{l'} \Gamma_{l'}} G_S^a) \\
&= \frac{e}{h} \int dE \sum_{l'} \operatorname{Tr} (\Gamma_l G_S^r \Gamma_{l'} G_S^a) (f_0(E, \mu_l) - f_0(E, \mu_{l'})) \\
&= \frac{e}{h} \int dE \sum_{l'} T_{ll'}(E) (f_0(E, \mu_l) - f_0(E, \mu_{l'})), \tag{2.117}
\end{aligned}$$

where we defined $T_{ll'}(E) = \operatorname{Tr} (\Gamma_l G_S^r \Gamma_{l'} G_S^a)$. In the next section, we will give a physical meaning to this object.

When the difference in the electrochemical potentials, i.e. the applied voltage, is small such that $\mu_l \approx \mu_{l'}$, it makes sense to define a Fermi energy $E_F \approx \mu_l$ for the whole system. The current (2.117) depends on the difference between the electrochemical potentials μ_l and $\mu_{l'}$ through the difference of the Fermi functions. Thus, only a small energy window between and around μ_l and $\mu_{l'}$ contributes to the energy integral. Provided that $T_{ll'}(E)$ does not vary significantly in this energy window, we can approximate $T_{ll'}(E) \approx T_{ll'}(E_F)$ and take it out of the energy integral. For small bias voltages and low temperatures, i.e. in the linear response regime [59, 60], we then find

$$\begin{aligned}
I_l &\approx \frac{e}{h} \sum_{l'} T_{ll'}(E_F) \int dE (f_0(E, \mu_l) - f_0(E, \mu_{l'})) \\
&= \frac{e}{h} \sum_{l'} T_{ll'}(E_F) (\mu_l - \mu_{l'}) = \frac{e}{h} \sum_{l'} T_{ll'}(E_F) (-eV_l + eV_{l'}) \\
&= \sum_{l'} G_{ll'} (V_{l'} - V_l) \tag{2.118}
\end{aligned}$$

where we defined the conductance

$$G_{ll'} = \frac{e^2}{h} T_{ll'}(E_F) = \frac{e^2}{h} \operatorname{Tr} (\Gamma_l G_S^r \Gamma_{l'} G_S^a) \tag{2.119}$$

between two leads l and l' .

The equations for G_S^r and $G_S^<$, (2.107) and (2.112) respectively, are the central results for calculating transport properties in a tight-binding model: The currents I_l can be calculated from G_S^r , while any local observable may be obtained from $G_S^<$ (see Appendix A). The non-equilibrium formalism thus provides a framework that can be applied to both interacting and non-interacting systems, with equations that are very suitable for a computer implementation, as they only involve matrix multiplications and inversions. For this reason the NEGF formalism has become

¹²¹⁵ There is a seeming contradiction between the assumption of steady state, i.e. time-independent observables, and the definition of current as $I_l = \frac{d}{dt} Q_l$. Per definition, $\frac{d}{dt} Q_l = 0$ in steady-state, and thus it seems as there could not be any current under steady-state conditions. However, this is a problem of the order of taking limits: The leads are assumed to be infinitely large. Therefore, we have to take both the limit of approaching steady state ($\lim_{t_0 \rightarrow -\infty}$) and infinitely large leads, $\lim_{Q_l \rightarrow \infty}$. Physically, this means that the electrochemical potential of the leads is not changed by a current flow.

rather popular for numerical computations, even for non-interacting systems. As a matter of fact, the nonequilibrium formalism can even be applied as a black box to obtain physical results. However, in such an approach it is difficult to develop a physical intuition of the underlying processes.

Unfortunately, the physical meaning of the Green's functions in the non-equilibrium formalism is clouded by the rather involved derivation of the basic equations. In fact, for a fully interacting system an intuitive picture may not be possible at all. This is however not true for non-interacting systems: In the next section we will give a physical meaning to the Green's functions by showing the equivalence of the NEGF and the scattering formalism for non-interacting systems, and further discuss to what extent this interpretation carries over to interacting systems.

19.1.11 2.5.3 Connection with the scattering formalism

A very intuitive approach to quantum transport in non-interacting systems has been pioneered by Landauer and Büttiker [88, 89]. As already observed in the case of

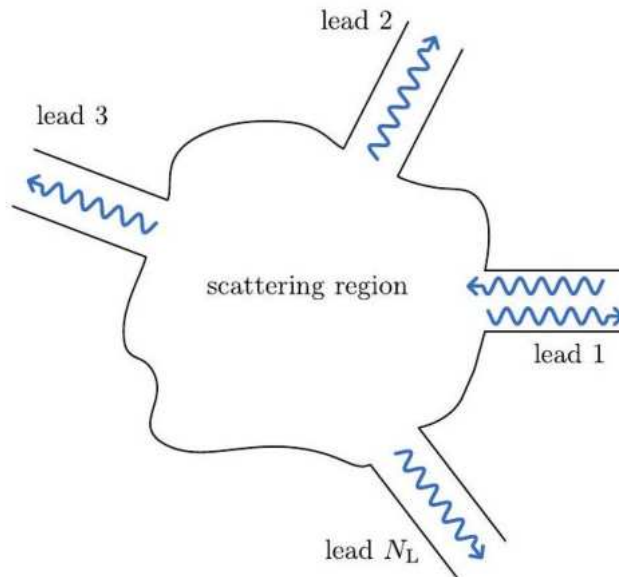


Figure 2.3 - An incoming wave from a certain lead is scattered inside the scattering region and may be either reflected back, or transmitted into another lead.

Green's functions, a non-interacting system can be characterized by the eigenstates of the single-particle Hamiltonian H . In the leads, these eigenstates are longitudinal plane waves with a transverse wave function with quantum index n (also called channel n ; here n includes both the orbital and spin degrees of freedom). An incoming wave is scattered within the scattering region, and can be either reflected back into the same lead or transmitted into a different lead (see Fig. 2.3). The quantum mechanical wave function that describes the scattering of an incoming wave in channel n from lead l is the scattering wave function $\Psi_n^{(l)}$. The central result of the Landauer-Büttiker formalism, also called scattering formalism, relates the conductance of a system to the properties of the scattering wave functions: the linear conductance between two leads is given by the sum of the transmission probabilities between the different channels in lead l and l' ,

$$G_{ll'} = \frac{e^2}{h} \sum_{mn} |t_{ll',nm}|^2 \quad (2.120)$$

where $t_{ll',nm}$ is the transmission amplitude from channel m in lead l' to channel n in lead l at the Fermi energy E_F . A detailed exhibition of the Landauer-Büttiker formalism in a

tight-binding model is given in Appendix C.

Whereas the transport properties in the Landauer-Büttiker formalism are uniquely defined by the scattering wave function, they are given in terms of Green's functions in the non-equilibrium Green's function formalism. In order to connect these complementary approaches, we need a relation between the scattering wave function and the

Green's function. Such a relation is commonly called Fisher-Lee relation ¹⁶. As shown in Appendix C for an arbitrary tight-binding model, Eq. (C.30), the scattering wave function at some (lattice) point $\mathbf{x}_i$ in the scattering region is given as

$$\Psi_n^{(l)}(\mathbf{x}_i) \propto G_{\mathbf{x}_i, 0_l}^r \Gamma_l \phi_{n,>}^{(l)} \quad (2.121)$$

where $\phi_{n,>}^{(l)}$ is the transverse wave function of an incoming channel n in lead l , and 0_l denotes the surface of lead l . This is the central result for showing the equivalence of the Green's function and the scattering formalism.

As shown in detail in Appendix C, we obtain

$$T_{ll'}(E_F) = \text{Tr}(\Gamma_l G_S^r \Gamma_{l'} G_S^a) = \sum_{mn} |t_{ll',nm}|^2 \quad (2.122)$$

proving the equivalence of the scattering and the non-equilibrium Green's function formalism. Thus, $T_{ll'}(E_F)$ can be interpreted as the total transmission probability of all channels in lead l' into all channels in lead l . In addition, we obtain an expression for $G_S^<$ as (see Eq. (C.38))

$$G_{S,\mathbf{x}_i\mathbf{x}_j}^<(E) \propto \sum_l f_0(E, \mu_l) \sum_n i \Psi_n^{(l)}(\mathbf{x}_i) \left(\Psi_n^{(l)}(\mathbf{x}_j) \right)^* D_n^{1\text{dd},(l)}(E), \quad (2.123)$$

where $D_n^{1\text{dd},(l)}(E)$ is the one-dimensional ¹⁷ density of states for a channel n in lead l . Thus, the lesser Green's function describes which scattering states are occupied at energy E . Whether a scattering wave function $\Psi_n^{(l)}$ is occupied or not is determined solely by the electrochemical potential μ_l of the respective leads.

Eq. (2.123) thus leads to a rather intuitive physical picture, depicted for the case of two leads, left (L) and right (R), in Fig. 2.4: The electronic states of the system are given by the scattering wave functions from left and right lead. For energies smaller than $\mu_R < \mu_L$, both scattering wave functions emanating from the left lead (rightmoving wave functions) and from the right lead (left-moving) are occupied ¹⁸. In the energy window $\mu_R < E < \mu_L$, only right-moving states are occupied, and there are no occupied states for energies above μ_L . This is a generalization of the interpretation of the lesser Green's function in Section 2.3, where we assumed a constant electrochemical potential throughout the whole system.

¹²¹⁶ It was Fisher and Lee that first published a relation between the retarded Green's function and the transmission amplitudes [90].

¹⁷ A channel n has an energy dispersion $E_{n,k}$ depending on a one-dimensional real wave vector k (see Section 3.1). Thus, a single channel behaves like a one-dimensional system.

¹⁸ For the sake of simplicity, we assume in our interpretation very low temperatures, $T \approx 0$, such that the Fermi function can be replaced by a step function $f_0(E, \mu) = \theta(\mu - E)$. For finite temperatures, there is some energy region around μ_R and μ_L where the smearing of the Fermi function also allows some additional occupancy.

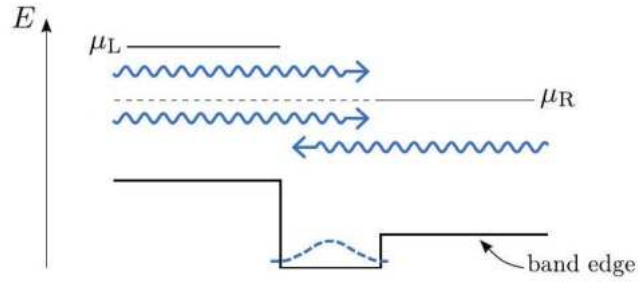


Figure 2.4-Schematic of a two-terminal system in non-equilibrium: scattering states emanating from the left (right) lead are occupied up to the electrochemical potential μ_L (μ_R), as depicted by wavy lines. In principle, there may also be localized states, as indicated by the dashed line.

From this picture, we can also understand the appearance of the factor $f_0(E, \mu_L) - f_0(E, \mu_R)$ in the current equation (2.117): For $E < \mu_R$, where $f_0(E, \mu_L) - f_0(E, \mu_R) = 0$ the currents carried by the left- and right-moving states cancel exactly. In contrast, for $\mu_R < E < \mu_L$ there are only right-moving states leading to a finite current from left to right.

In principle, there could also be localized states in the scattering region, that are not connected to any scattering state emanating from the leads, as indicated in Fig. 2.4. For example, such a state could lie below the band edges of the leads, and since there are no transitions between states of different energy in a non-interacting system, this localized state cannot be connected to any scattering state. These states are thus not included in the simplified form of the Keldysh equation (2.95). Instead, they are contained in the first term of Eq. (2.94) that is usually neglected [76]. Note that the first term of Eq. (2.94) involves the lesser Green's function $G_{0,S}^<$ of the disconnected scattering region, and thus μ_S . Essentially, this means that the electrochemical potential μ_S in the far past determines whether a localized state is occupied or not. This comes at no surprise, since it cannot be filled from the states of the leads that are at a different energy. Since μ_S is somewhat ill-defined as mentioned above, additional physical arguments are required to determine the occupancy of localized states. Fortunately, localized states do not contribute to transport in non-interacting systems, as they do not couple to the scattering states, and are thus of minor importance in this work. In interacting systems however, they may influence transport through the interaction term, and additional care must be taken to account for these states [85, 91].

Transport in a non-interacting system can thus be understood in a rather intuitive picture. To which extent does this interpretation carry over to the interacting case? Eqs. (2.122) and (2.123) are of central importance in showing the equivalence of the scattering and the non-equilibrium Green's function formalism. The main assumption behind the derivation of these equations is that the lesser self-energy $\Sigma^<$ only contains contributions from the leads. This is the case in non-interacting systems, but also for the example of Coulomb interaction within Hartree-Fock approximation and generally for any interaction in mean-field approximation (including density functional theory¹⁹). In all these cases, transport can be still understood in terms of scattering wave functions in an effective potential. Only if the interaction is treated beyond mean-field, leading to a contribution to $\Sigma^<$, this interpretation breaks down, as then (virtual) transitions between states of different energies become important.

Since the scattering formalism and the NEGF formalism for non-interacting systems are equivalent, often concepts of both approaches are used in a calculation at the same time. For example, since calculating the scattering wave function directly is in general numerically unstable [92], numerical calculations in the Landauer-Büttiker formalism usually first compute G_S^r from Eq. (2.111), and then the transmission amplitudes $t_{ll',nm}$ through the Fisher-Lee relation. Note that the $t_{ll',nm}$ contain more information than just the total transmission $T_{ll'}(E)$.

Because of that, and since in addition calculating the total transmission probability $T_{ll'}(E)$ from $t_{ll',nm}$ is computationally less expensive than computing $T_{ll'}(E)$ from the Green's functions, we also employ this approach in the numerical work. The main conceptual difference between the Landauer-Büttiker formalism and the non-equilibrium Green's function formalism is the calculation of $G_S^<$, and in this sense Eq. (2.112) contains the "non-equilibrium". As seen above, this quantity could also be calculated from the scattering wave functions, but this is usually not done. It should be emphasized that of course also the Landauer-Büttiker formalism is a non-equilibrium theory - there would be no current and hence no conductance in equilibrium. A non-equilibrium situation, i.e. different electrochemical potentials in the leads, enters explicitly in the derivation of the Landauer-Büttiker formula.

The scattering formalism has the advantage of providing a more intuitive picture of transport. In contrast, the Green's functions in the non-equilibrium Green's function formalism are very convenient to compute numerically, and the formalism can be easily extended to include interactions. Because of this the NEGF formalism has become rather popular for numerical computations.

19.1.12 2.6 Challenges for a numerical implementation

The transport formalism presented in the previous section is very general, as it is applicable to any tight-binding model. Central results were Eq. (2.111), $G_S^r = (E - H_S - \Sigma^r)^{-1}$, and Eq. (2.112), $G_S^< = G_S^r \Sigma^< G_S^a$. However, directly implementing these equations in a computer program still faces some difficulties. In this section, we will identify the remaining steps necessary to build a generic transport code for arbitrary

tight-binding models. In the following two chapters we will then present novel solutions to these problems.

The first step is obvious from Eq. (2.111): we need to compute Σ^r . In Chapter 3 we thus derive an expression for the self-energy Σ_l^r of a lead l , valid for arbitrary tight-binding models, paying close attention to the numerical stability of this method.

The second step is the calculation of G_S^r . At first glance, this does not appear problematic, as G_S^r can be in principle computed by direct inversion as given in Eq. (2.111). This inversion involves a matrix of size N_{grid} , where N_{grid} is the number of degrees of freedom in the tight-binding model, i.e. the number of lattice sites or orbitals in the scattering region. However, physical simulations involve in general large-scale systems, and a direct inversion is then computationally intractable.

Then again, it is often not necessary to calculate the full retarded Green's function G_S^r : For example, to obtain the conductance, it is enough to compute the Green's function between the leads, as can be seen from Eqs. (2.117) and (C.31) and the fact that the self energies from the leads Σ_l^r only couple to the surface of the scattering region. This and the fact that, in a tight-binding model, H_S is a sparse matrix, is the foundation of several quantum transport algorithms that avoid the direct inversion of Eq. (2.111). These transport algorithms however have the disadvantage of relying on some assumptions or geometrical restrictions on the underlying system, and are thus not applicable to arbitrary tight-binding models. In Chapter 4 we review several well-established transport algorithms and develop a preconditioning step that allows to apply these algorithms to an arbitrary tight-binding model. A detailed description of the quantum transport algorithm employed in this work, the recursive Green's function technique,

¹²¹⁹ In the case of density functional theory, the physical interpretation is complicated by the fact that the scattering wave functions are then solutions to the Kohn-Sham equations. However, the solutions of the Kohn-Sham equations do not have any physical meaning by themselves and neither do the corresponding scattering wave functions. But then, transport theory in density functional theory must assume anyway that the wave functions obtained like that are close to physical wave functions [84].

is given in Appendix B.

19.2 Chapter 3 Lead Green's functions

19.2.1 3.1 Introduction

Calculating transport properties in the non-equilibrium Green's function formalism requires the evaluation of the lead self-energies $\Sigma_l^r = H_{Sl} g_l^r H_l$, as seen in the previous chapter. Finding an expression for the surface Green's function of the lead g_l^r , and thus for Σ_l^r , is the intent of this chapter. Below, we will drop the subscript l since we are dealing here with a single lead only, and simply denote $g_l^r = g^r$. In addition, for the sake of brevity, we will often refer to the surface Green's function of the lead simply as the lead Green's function.

In general, a lead is a semi-infinite structure attached to the scattering region. In particular, we consider leads that in addition exhibit some periodicity, so that asymptotic scattering states can be defined. In tight-binding representation, such a lead is a semi-infinite repetition of unit cells, governed by the unit cell Hamiltonian H_0 , with neighboring unit cells connected by hopping Hamiltonians H_1 and $H_{-1} = H_1^\dagger$, as depicted in Fig. 3.1. If there are $N_{\text{u.c.}}$ degrees of freedom within a unit cell, these Hamiltonians can be represented by $N_{\text{u.c.}} \times N_{\text{u.c.}}$ matrices.

For transport, we have to calculate the surface Green's function of the lead detached from the scattering region. The Hamiltonian of such an isolated, semi-infinite lead is given by

$$H_{\text{semi-infinite}} = \begin{pmatrix} H_0 & H_1 & 0 & \dots \\ H_{-1} & H_0 & H_1 & \\ 0 & H_{-1} & H_0 & \\ \vdots & & & \ddots \end{pmatrix} \quad (3.1)$$

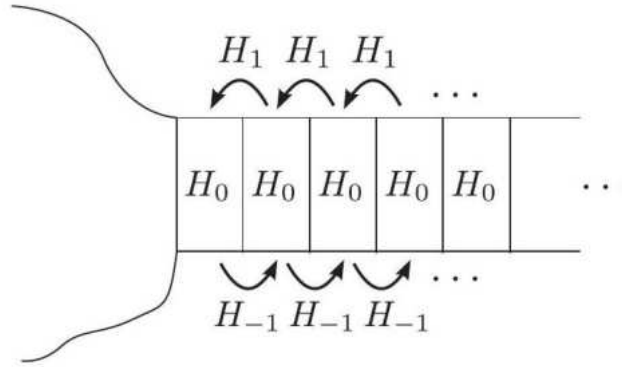


Figure 3.1 - Schematic view of a lead: A semi-infinite series of unit cells. and we can write this Hamiltonian in a form reminiscent of a one-dimensional wire,

$$H_{\text{semi-infinite}} = \sum_{j=0}^{\infty} H_0 |j\rangle \langle j| + H_1 |j\rangle \langle j+1| + H_{-1} |j\rangle \langle j-1|, \quad (3.2)$$

where j is the unit cell index. The surface Green's function g^r is defined only within the first unit cell of the semi-infinite wire, $g^r = \langle 0 | G_{\text{semi-infinite}} | 0 \rangle$, where $G_{\text{semi-infinite}}$ is the full Green's function of the isolated lead. Therefore, the quantity that we desire to calculate is a well-defined, finite matrix and its size is given by the size of the unit cell Hamiltonian H_0 .

A problem related to that of the semi-infinite lead is the infinite wire that extends to infinity in both directions. Its Hamiltonian is given by

$$H_{\text{infinite}} = \begin{pmatrix} \ddots & & & & \\ & H_0 & H_1 & & \\ & H_{-1} & H_0 & H_1 & \\ & & H_{-1} & H_0 & H_1 \\ & & & H_{-1} & H_0 \\ & & & & \ddots \end{pmatrix} \quad (3.3)$$

or in more compact form

$$H_{\text{infinite}} = \sum_{j=-\infty}^{\infty} H_0 |j\rangle \langle j| + H_1 |j\rangle \langle j+1| + H_{-1} |j\rangle \langle j-1| \quad (3.4)$$

The Hamiltonian H_{infinite} is periodic, and thus its eigenstates can be written in Bloch form [93]:

$$\varphi_{n,k}(j) = \phi_{n,k} e^{ikj} \quad (3.5)$$

where j again denotes the unit cell index, $k \in [-\pi, \pi]$, and $\phi_{n,k}$ is a vector representing the "transverse" part of the wave function within a unit cell. The wave function $\varphi_{n,k}$ is a solution of the Schrödinger equation

$$(E_{n,k} - H_{\text{infinite}}) \varphi_{n,k} = 0 \quad (3.6)$$

if the transverse wave function $\phi_{n,k}$ obeys

$$(E_{n,k} - H_0 - H_1 e^{ik} - H_{-1} e^{-ik}) \phi_{n,k} = 0 \quad (3.7)$$

The effective Hamiltonian $H(k) = H_0 + H_1 e^{ik} + H_{-1} e^{-ik}$ is Hermitian and thus, for a given k , the wave functions $\{\phi_{n,k}\}$ form a complete orthonormal set.

The infinite wire is related to the problem of the semi-infinite lead in several ways: First, away from the contact to the scattering region, the asymptotic eigenstates of the semi-infinite lead are given by the eigenstates $\varphi_{n,k}$ of the infinite wire. Furthermore, the Green's functions of the semi-infinite lead and the infinite wire are connected via Dyson's equation, as discussed below.

In principle, it would be possible to calculate g^r using the eigenfunction expansion of the Green's function, Eq. (2.41). For that, we would need to calculate the eigenstates of $H_{\text{semi-infinite}}$. Such an eigenstate can in principle be constructed as a linear combination of solutions of the infinite wire. The coefficients of this linear combination must be chosen such that the boundary conditions of the isolated semi-infinite lead are fulfilled. This can be done easily if the transverse wave functions $\phi_{n,k}$ do not explicitly depend on k [59]: By combining propagating and counter-propagating modes, $\varphi(j) \sim \phi_n e^{ikj} - \phi_n e^{-ikj} \sim \phi_n \sin(kj)$, the wave function can be made zero at the end of the semi-infinite lead and thus satisfies the boundary condition there.

However, in general the transverse wave function $\phi_{n,k}$ can depend on k . For example, in a wire including a strong magnetic field the wave function propagating in one direction is localized at one edge of the wire, whereas the mode propagating in the opposite direction is localized at the opposite edge. In this case the construction of an eigenstate of the semi-infinite lead becomes very difficult and it is essential to include not only the propagating eigenstates $\varphi_{n,k}$, but also evanescent modes. The construction is even further complicated by the fact that different transverse wave functions $\phi_{n,k}$ and $\phi_{n',k'}$ are in general not orthogonal anymore, if $k \neq k'$. This is due to the fact that they are solutions to different effective Hamiltonians $H(k)$ and $H(k')$. Only the full wave functions $\varphi_{n,k}$ form an orthonormal set. Because of these difficulties we will not attempt to compute the eigenstates of the semi-infinite lead. Instead, we

will first calculate the Green's function of the infinite wire for which the eigenstates are known and from that obtain the surface Green's function of the lead.

The problem of calculating the surface Green's function of a lead has been addressed frequently in the literature. First of all, the problem is strongly related to earlier accounts calculating the Green's function of an ideal crystal surface [94-97]. In contrast to a lead of finite transverse size as considered here, such a crystal surface is infinite and such are the respective Hamiltonian matrices. However, the problem can finally be reduced to a finite size invoking Bloch's theorem for the directions along the surface.

The numerical algorithms for computing the surface Green's function of a lead can be classified into two groups: Iterative methods and eigendecomposition based methods. The iterative algorithms calculate an approximate Green's function through some recursion relation [98-100] and require explicitly shifting the energy E by some small, but finite imaginary part. The eigendecomposition based methods compute the surface Green's function exactly-within the numerical precision-from the eigensolutions of an infinite wire. These methods have been derived independently several times [82, 86, 95-97, 101, 102], and equivalent expressions are also found in the mode matching method [103, 104], where instead of the surface Green's function the matching of lead eigenstates is considered¹. Compared to the iterative methods, the eigendecomposition based methods have been shown to be superior in accuracy and performance [101], and they provide additional information, such as the eigenmodes (channels) of the lead.

All of the accounts of the eigendecomposition based methods assume the hopping matrices H_1, H_{-1} to be non-singular. However, singular-i.e. non-invertible, rank-deficient-hopping matrices arise naturally in important examples and are not at all unusual: For instance, a unit cell size larger than the range of the hopping in tightbinding approximation immediately leads to singular H_1 and H_{-1} [82, 86]. Graphene nanoribbons, subject of Chapters 6 and 7, are such an example. Iterative methods would be applicable to the case of singular hopping matrices, but their performance for singular hopping matrices is even worse than for non-singular hopping matrices. Still, in principle any problem involving singular hopping matrices could be recast into a problem involving only invertible hopping matrices [86]. However, the algorithm to do so is much more complicated than the expressions we will derive in the course of this chapter.

Thus, we will derive an eigendecomposition based expression for the surface Green's function equally valid for singular and non-singular hopping matrices H_1, H_{-1} . To this end, the chapter is organized as follows: First, we will rigorously derive an expression for the surface Green's function of a lead analogous to previously derived expressions [82, 86, 101, 102], but without assuming the hopping matrices H_1, H_{-1} to be invertible. We then discuss cases for which a direct implementation of this expression is numerically unstable and develop an algorithm circumventing these stability problems.

19.2.2 3.2 A general expression for the lead Green's function

19.2.3 3.2.1 Green's function of the infinite wire

Before we begin to calculate the Green's function of the infinite wire, we collect some useful relations that will allow us to extract the surface Green's function from that of

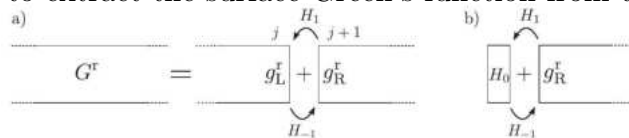


Figure 3.2 - (a) A infinite wire can be constructed by combining two semi-infinite wires. (b) A semi-infinite wire does not change upon adding another unit cell.

¹²1 The mode matching method is equivalent to the Green's function methods, as discussed in Ref. [105].

the infinite wire in the subsequent sections.

The surface Green's function of the lead and the Green's function of the infinite wire can be related by the Dyson equation for the retarded Green's function, Eq. (2.91), in a similar fashion as in the recursive Green's function algorithm: The infinite wire can be regarded as a composition of two semi-infinite leads with the hopping between left and right lead as perturbation, as depicted in Fig. 3.2(a). We denote the retarded surface Green's function of the right (left) lead as g_R^r (g_L^r) and the retarded Green's function of the infinite wire as G^r . g_R^r is the desired surface Green's function of the lead. Note that because of translational invariance the surface Green's functions $g_{R,L}^r$ are independent of the position j , where the infinite wire is cut. Furthermore, g_L^r differs from g_R^r only therein, that the hopping Hamiltonians H_1 and H_{-1} are switched.

From the Dyson equation $G^r = G_0^r + G_0^r V G^r = G_0^r + G^r V G_0^r$, Eqs. (2.91) and (2.92), we can obtain relations between the surface Green's functions and the Green's function of the infinite wire. In the situation as depicted in Fig. 3.2(a), the unperturbed Green's function G_0^r is given by the surface Green's functions $g_{R,L}^r$ of the isolated leads and the perturbation V by the hopping matrices H_1 and H_{-1} . From this, we find:

$$\begin{aligned} G_{j,j'+1}^r &= G_{j,j'}^r H_1 g_R^r & \text{for } j' \geq j \\ G_{j+1,j'}^r &= g_R^r H_{-1} G_{j,j'}^r & \text{for } j' \leq j \end{aligned} \quad (3.9)$$

Similar relations hold for g_L^r .

For the surface Green's function, another useful relation arises from the observation that the semi-infinite lead is invariant under addition of an extra unit cell (see Fig. 3.2(b)), and we obtain

$$g_R^r = (E - H_0 - H_1 g_R^r H_{-1})^{-1} \quad (3.10)$$

Using the eigenstates of the infinite wire, Eq. (3.6), and the eigenstate representation of the retarded Green's function, Eq. (2.41), the retarded Green's function of the infinite wire G^r is given as follows:

$$\begin{aligned} G_{j,j'}^r &= \frac{1}{2\pi} \int_{-\pi}^{\pi} dk \sum_n \frac{\phi_{n,k} e^{ikj} \phi_{n,k}^\dagger e^{-ikj'}}{E - E_{n,k} + i\eta} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} dk \sum_n \underbrace{\left(\frac{1}{E - E_{n,k} + i\eta} \phi_{n,k} \right)}_{= \frac{1}{E - H(k) + i\eta} \phi_{n,k}} \phi_{n,k}^\dagger e^{ik(j-j')} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} dk \frac{1}{E - H(k) + i\eta} \underbrace{\left(\sum_n \phi_{n,k} \phi_{n,k}^\dagger \right)}_{= \mathbb{1}} e^{ik(j-j')} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} dk \frac{e^{ik(j-j')}}{E - H(k) + i\eta}, \end{aligned} \quad (3.11)$$

where we have used that, for fixed k , the wave functions $\{\phi_{n,k}\}$ form a complete orthonormal set. By further substituting $z = e^{ik}$, the integration over k can be written as a contour integral on the unit circle and we arrive at

$$G_{j,j'}^r = \frac{1}{2\pi i} \oint dz \frac{z^{j-j'}}{(E + i\eta - H_0)z - H_1 z^2 - H_{-1}}. \quad (3.12)$$

It should be noted that the denominator of (3.12) is a matrix and the fraction should therefore be interpreted as a matrix inversion. Thus, the problem is more complicated than a

scalar valued contour integral. Still, in order to solve the matrix-valued contour integral, we first identify the zeroes of the denominator, just as in the scalar case.

19.2.4 3.2.2 Zeroes of $(\mathbf{E} + i\eta - \mathbf{H}_0) \mathbf{z} - \mathbf{H}_1 \mathbf{z}^2 - \mathbf{H}_{-1}$

The contour integral of (3.12) is nonzero only if the entries of the matrix $((E + i\eta - H_0)z - H_1z^2 - H_{-1})^{-1}$ have poles within the unit circle. The inverse of the matrix $\tilde{H}(E + i\eta, z) = (E + i\eta - H_0)z - H_1z^2 - H_{-1}$ has a pole at λ_n , if the matrix $\tilde{H}(E + i\eta, z)$ itself is singular at $z = \lambda_n$. That is, we seek solutions of the (quadratic) eigenvalue equation

$$((E + i\eta - H_0)\lambda_n - H_1\lambda_n^2 - H_{-1}) \mathbf{u}_n = 0 \quad (3.13)$$

In principle, the infinitesimal $i\eta$ could be taken care of in the numerical calculations by choosing some very small, but finite η . However, this introduces some unnecessary parameter dependence into the system, and η might need manual tuning in order to achieve convergence. We can avoid explicitly using η in the numerics by observing that η enters the calculation in a twofold way:

First, it determines which poles contribute to the contour integral in Eq. (3.12). In this respect, the infinitesimal η is crucial. Below, we will carefully distinguish different cases in order to account for this dependence on η .

Second, the values of the eigenvalues λ_n -i.e. the poles - and the eigenvectors $\mathbf{u}_n$ depend on η : $\lambda_n = \lambda_n(E + i\eta)$, $\mathbf{u}_n = \mathbf{u}_n(E + i\eta)$. However, λ_n and $\mathbf{u}_n$ are analytic functions of $E + i\eta$, and thus the dependence on η vanishes in the limit $\eta \rightarrow 0^+$. In particular, $\lim_{\eta \rightarrow 0^+} \lambda(E + i\eta) = \lambda_n(E)$, where $\lambda_n(E)$ is an eigenvalue of the quadratic eigenproblem without the infinitesimal shift $i\eta$. The same holds for the eigenvectors $\mathbf{u}_n$.²

Therefore, we seek solutions of the quadratic eigenvalue problem neglecting i (for a review on quadratic eigenproblems see [106])

$$((E - H_0)\lambda_n - H_1\lambda_n^2 - H_{-1}) \mathbf{u}_n = 0 \quad (3.14a)$$

or

$$\bar{\mathbf{u}}_n^\dagger ((E - H_0)\lambda_n - H_1\lambda_n^2 - H_{-1}) = 0 \quad (3.14b)$$

where $\mathbf{u}_n$ ($\bar{\mathbf{u}}_n$) is a right (left) eigenvector for the eigenvalue λ_n . Such a quadratic eigenproblem has $2N_{\text{u.c.}}$ eigenvalues, when $N_{\text{u.c.}}$ is the size of the matrix $\tilde{H}(E, z)$.³ However, these eigenvalues are not necessarily finite, if H_{-1} is rank-deficient [106] and thus must be discussed carefully.

The quadratic eigenproblem can be recast into a linear eigenproblem, albeit of twice the size. By introducing the vector $\mathbf{v} = (\mathbf{u}, \lambda\mathbf{u})^t$ Eq. (3.14a) can be rewritten as

$$\begin{pmatrix} 0 & 1 \\ -H_{-1} & E - H_0 \end{pmatrix} \begin{pmatrix} \mathbf{u}_i \\ \lambda_n \mathbf{u}_n \end{pmatrix} = \lambda_n \begin{pmatrix} 1 & 0 \\ 0 & H_1 \end{pmatrix} \begin{pmatrix} \mathbf{u}_n \\ \lambda_n \mathbf{u}_n \end{pmatrix} \quad (3.15)$$

which is a generalized eigenproblem of the form $A\mathbf{v} = \lambda B\mathbf{v}$. Such a generalized eigenproblem can be recast into an ordinary eigenproblem if the matrix B is invertible, i.e. if the matrix H_{-1} is invertible:

$$\begin{pmatrix} 0 & 1 \\ -H_1^{-1}H_{-1} & H_1^{-1}(E - H_0) \end{pmatrix} \begin{pmatrix} \mathbf{u}_n \\ \lambda_n \mathbf{u}_n \end{pmatrix} = \lambda_n \begin{pmatrix} \mathbf{u}_n \\ \lambda_n \mathbf{u}_n \end{pmatrix}. \quad (3.16)$$

In any case, there are standard algorithms for solving both the ordinary and the generalized linear eigenproblem [107].

The eigenvalues of Eqs. (3.14a) and (3.14b) fall into two groups with $|\lambda_n| \neq 1$ and $|\lambda_n| = 1$. We first consider the case $|\lambda_n| \neq 1$. Then, λ_n lies inside the unit circle if $|\lambda_n| < 1$ and outside if $|\lambda_n| > 1$, regardless of the infinitesimal imaginary part of the energy $i\eta$. In general, λ_n can be written as $\lambda_n = e^{ik} = e^{i\text{Re}(k)}e^{-\text{Im}(k)}$. $|\lambda_n| < 1$ corresponds to $\text{Im}(k) > 0$ and an evanescent mode decaying to the right, whereas

$|\lambda_n| > 1$ corresponds to $\text{Im}(k) < 0$ and an evanescent mode decaying to the left. Furthermore, there is an equal number of right- and left-decaying solutions: By taking the complex conjugate of Eq. (3.14a) and using $H_1^\dagger = H_{-1}$, we find that for every eigenvalue λ_n , $1/\lambda_n^*$ is an eigenvalue of the quadratic eigenproblem, too [102]:

$$\mathbf{u}_n^\dagger \left((E - H_0) 1/\lambda_n^* - H_1 (1/\lambda_n^*)^2 - H_{-1} \right) = 0. \quad (3.17)$$

Thus, for every eigenvalue within the unit circle there is a corresponding eigenvalue outside and vice versa.

As already mentioned above, the quadratic eigenvalue problem can have the eigenvalue ∞ , if the matrix H_{-1} (and thus also $H_1 = H_{-1}^\dagger$) is not invertible. In this case there is a vector $\mathbf{u}_n$ such that $H_{-1}\mathbf{u}_n = 0$ and $\lambda_n = 0$ solves Eq. (3.14a). Then also $1/\lambda_n^*$ is an eigenvalue and corresponds to ∞ . It is therefore quite natural to obtain diverging eigenvalues if the matrix H_{-1} is rank-deficient and not at all problematic as opposed to the statements of Ref. [86]. The eigenvectors for $\lambda_n = 0, \infty$ itself do not have any physical meaning, however their contribution in terms of the contour integral in Eq. (3.12) is well-defined. Therefore, we also call $\lambda_n = 0$ (which lies within the unit circle) a "right-decaying" mode. The importance of these solutions has been first pointed out in the context of complex band structure calculations [108].

If $|\lambda_n| = |e^{ik}| = 1$, the eigenvectors $\mathbf{u}_n$ are the propagating Bloch modes $\phi_{n,k}$ discussed above. Here we are considering a fixed energy E , and thus k is uniquely defined by $E_{n,k} = E$ and will be denoted by k_n . As a consequence, we will also drop the index k_n in the notation of the transverse wave function ϕ_n for the remainder of the chapter. Note that we will employ the notation ϕ_n if we want to emphasize that the eigenvector under consideration is a propagating mode, whereas $\mathbf{u}_n$ denotes all eigenvectors, propagating and evanescent.

For the propagating modes ϕ_n the corresponding eigenvalues λ_n lie on the unit circle. Thus, the infinitesimal imaginary shift $i\eta$ becomes crucial in identifying whether the eigenvalue lies inside or outside of the unit circle. As shown by Taylor expanding $\lambda_n(E + i\eta)$ in appendix D.1, this is determined by the group velocity v_n of the Bloch mode ϕ_n .

The group velocity v_n is given by

$$\begin{aligned} v_n &= \frac{1}{\hbar} \frac{dE_{n,k}}{dk} \Big|_{k=k_n} \\ &= \frac{1}{\hbar} \frac{d}{dk} \langle \phi_{n,k} | H_0 + H_1 e^{ik} + H_{-1} e^{-ik} | \phi_{n,k} \rangle \Big|_{k=k_n} \end{aligned} \quad (3.18)$$

Using the Feynman-Hellman theorem [109, 110] this expression can be simplified yielding

¹²² Note that the case of degenerate eigenvalues, $\lambda_n = \lambda_{n'}$ is a little more complicated in this respect and will be discussed in detail in the next section.

³ We can assume the quadratic eigenvalue problem to be regular. In an irregular eigenvalue problem $\det((E - H_0)\lambda - H_1\lambda - H_{-1}) = 0$ for all $\lambda \in \mathbb{C}$ which is unphysical. For example, in the infinite wire this would imply a flat band, $E(k) = \text{const.}$, with transverse wave function ϕ independent of k . This may only occur in ill-defined examples, such as a series of independent lattice points that are not coupled all.

$$\begin{aligned}
v_n &= \frac{1}{\hbar} \langle \phi_{n,k} | \frac{d}{dk} (H_0 + H_1 e^{ik} + H_{-1} e^{-ik}) | \phi_{n,k} \rangle \Big|_{k=k_n} \\
&= \frac{i}{\hbar} \phi_n^\dagger (H_1 e^{ik_n} - H_{-1} e^{-ik_n}) \phi_n \\
&= -\frac{1}{\hbar} \Im (2\phi_n^\dagger H_1 \lambda_n \phi_n).
\end{aligned} \tag{3.19}$$

As shown in appendix D.1, the eigenvalue λ_i lies inside the unit circle if $v_n > 0$, corresponding to a right-moving wave, and outside if $v_n < 0$, corresponding to a leftmoving wave.

Note that in this situation we cannot invoke Eq. (3.17) to make a statement about the number of right- and left-moving modes, as $1/\lambda_n^* = 1/(e^{ik})^* = e^{ik} = \lambda_n$. Eq. (3.17) thus only shows that for any right eigenvector ϕ_n , the Hermitian conjugate $\phi_n^\dagger$ is a left eigenvector to the same eigenvalue λ_n . Still, we can determine the number of right- and left-moving waves by considering the properties of the Bloch bands $E_{n,k}$. The Bloch bands $E_{n,k}$ are periodic in k with periodicity 2π . Therefore, at a given energy E , there is an equal number of right- and left-moving waves, as the Bloch bands $E_{n,k}$ intersect the energy E equally often with positive and negative slope [82].

Summarizing the arguments of this section, we find that the eigenvalues and -vectors of the quadratic eigenproblem, Eq. (3.14a), can be grouped into

- $N_{\text{u.c.}}$ right-going modes with $|\lambda_i| < 1$ or $|\lambda_i| = 1 \wedge v_n > 0$. These eigenvalues are denoted by $\lambda_{i,<}$, the corresponding eigenvectors by $\mathbf{u}_{i,<}$ or $\phi_{n,<}$. In addition, we define the matrix of right-going modes as $U_{<} = (\mathbf{u}_{1,<}, \dots, \mathbf{u}_{N_{\text{u.c.},<}})$.
- $N_{\text{u.c.}}$ left-going modes with $|\lambda_i| > 1$ or $|\lambda_i| = 1 \wedge v_n < 0$. These eigenvalues are denoted by $\lambda_{i,>}$, the corresponding eigenvectors by $\mathbf{u}_{i,>}$ or $\phi_{n,>}$. In addition, we define the matrix of left-going modes as $U_{>} = (\mathbf{u}_{1,>}, \dots, \mathbf{u}_{N_{\text{u.c.},>}})$.

It should be emphasized that these eigenvalues and -vectors can be computed and classified without having to resort to a finite imaginary shift $i\eta$.

19.2.5 3.2.3 The role of degenerate modes

The solutions of the quadratic eigenvalue problem (3.14a) are only unique, if all eigenvalues λ_n are different. If $\mathbf{u}_n$ and $\mathbf{u}_{n'}$ are solutions with the same eigenvalue $\lambda = \lambda_n = \lambda_{n'}$, then any linear combination solves the quadratic eigenproblem:

$$((E - H_0) \lambda + H_1 \lambda^2 + H_{-1}) (a\mathbf{u}_n + a'\mathbf{u}_{n'}) = 0 \quad \text{for } a, a' \in \mathbb{C} \tag{3.20}$$

Although this does not pose a problem for decaying solutions ⁴, it very well can for the propagating solutions, where $\lambda_n = e^{ik_n}$. As shown in Section 3.2 · 2, the group velocity

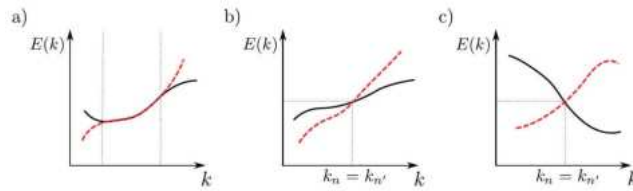


Figure 3.3 - Schematic band structure plots for different cases of degeneracies in Bloch modes: (a) Two bands are degenerate for a whole range of k -values. (b) Two bands cross at a

¹²⁴ Decaying solutions are classified only according to their eigenvalue λ_n , the eigenvector $\mathbf{u}_n$ does not enter. Furthermore, the final results, given in Eqs. (3.28) and (3.29) are invariant under a change of basis, i.e. choosing different linear combinations of degenerate solutions.

single k -point with slopes of the same sign. (c) Two bands cross with slopes of opposite signs. v_n of propagating modes determines whether a solution lies inside or outside the unit circle. If two or more propagating modes share the same eigenvalue, i.e. $k_n = k_{n'}$, the group velocity becomes ill-defined, as different linear combinations may result in different group velocities. Moreover, the eigensolutions $\phi_n(E)$ are not guaranteed to be smooth functions around $E = E_{n,k_n}$: Whereas at the degeneracy point E_{n,k_n} any linear combination of degenerate modes is a solution, moving away from this point in the complex plane will lift the degeneracy and lead to definite solutions.⁵

The different types of degeneracies that can occur are depicted schematically in Fig. 3.3. First, energy bands can be degenerate over a range of k -values, as shown in Fig. 3.3(a). Such a degeneracy can be due to symmetries of the underlying lattice, or occurs naturally if the system is degenerate with respect to an additional quantum number, such as spin. Second, two or more bands can intersect with a slope (and thus velocity) of the same sign (Fig. 3.3(b)). Both cases are not problematic, as then all degenerate eigenvalues are either inside or outside the unit circle simultaneously. If however, as in the third case (Fig. 3.3(c)), bands intersect with slopes of opposite sign, some of the degenerate eigenvalues are located inside and some outside the unit circle.

The troublesome degeneracy of eigenstates is an artifact of neglecting the infinitesimal shift $i\eta$. We can however find a remedy by observing that for non-degenerate λ_n the eigenvectors ϕ_n diagonalize the velocity operator (Eq. (3.19)), as shown in appendix C; this is not necessarily the case for degenerate eigenvalues. We can expect properties that depend on the eigenvectors ϕ_n , such as the diagonal form of the velocity operator, to change in a continuous way. Therefore, the condition that the velocity operator should be diagonal with respect to the eigenvectors ϕ_n uniquely defines a set of basis vectors with definite group velocity v_n that can be used to reliably classify these modes

as right- and left-moving.

Suppose that $\phi_{n_1}, \dots, \phi_{n_l}$ are eigenvectors for the eigenvalue $\lambda = e^{ik}$. Let $\Phi = (\phi_{n_1}, \dots, \phi_{n_l})$ be the matrix built from these eigenvectors. Then, the velocity operator in the subspace of these eigenvectors can be written as

$$V_\Phi = \Phi^\dagger i (H_1 e^{ik} - H_{-1} e^{-ik}) \Phi \quad (3.21)$$

V_Φ is Hermitian and thus can be diagonalized [82] as $V'_\Phi = S^\dagger V_\Phi S$ with some unitary matrix S . From the columns of $\Phi' = \Phi S$ we then find the new eigenvectors $\phi'_{n_1}, \dots, \phi'_{n_l}$ with well-defined velocities. Furthermore, the rotated ϕ'_n are smooth functions with respect to energy E .

19.2.6 3.2.4 Solving the contour integral

Before proceeding further in solving the contour integral in Eq. (3.12), we first study the eigenvalues and -vectors of $g_R^\dagger H_{-1}$, as this can give some indication on the final expression for the surface Green's function. Suppose $\mathbf{u}$ is an eigenvector of $g_R^\dagger H_{-1}$ with eigenvalue λ :

$$g_R^\dagger H_{-1} \mathbf{u} = \lambda \mathbf{u} \quad (3.22)$$

Then we find

¹²⁵ Indeed, the smoothness of $\phi_n(E)$ is a prerequisite of the derivation in appendix D. 1 that shows that the group velocity of propagating modes determines whether a eigenvalue lies inside or outside the unit circle.

$$\begin{aligned}
H_{-1}\mathbf{u} &= (g_{\text{R}}^{\text{r}})^{-1} \lambda \mathbf{u} \\
&= (E - H_0 - H_1 g_{\text{R}}^{\text{r}} H_{-1}) \lambda \mathbf{u} \\
&= ((E - H_0) \lambda - H_1 \lambda^2) \mathbf{u}
\end{aligned} \tag{3.23}$$

where we have used Eqs. (3.10) and (3.22). Every eigenvalue λ is thus also a solution of the quadratic eigenproblem

$$((E - H_0) \lambda - H_1 \lambda^2 - H_{-1}) \mathbf{u} = 0. \tag{3.24}$$

$g_{\text{R}}^{\text{r}} H_{-1}$ can be expressed in terms of its eigenvectors and -values and therefore takes the form

$$g_{\text{R}}^{\text{r}} H_{-1} = U \Lambda U^{-1} \tag{3.25}$$

where the matrices U and Λ are built up by choosing the appropriate eigenvectors and -values from the $2N$ solutions of the quadratic eigenproblem (3.14a). Unfortunately, these considerations do not give an indication on which solutions should be chosen.

The appropriate eigenvalues and -vectors can be selected by using the boundary condition of the retarded Green's function, namely that it should contain only outgoing waves [59, 82]. However, this intuitive physical arguments break down if systems with rank-deficient matrices H_1 and H_{-1} are considered and, as argued above, spurious solutions with vanishing or infinite eigenvalues appear. These solutions do not have any physical meaning, but are well-defined in the context of the contour integral of Eq. (3.12). Therefore we resort to this contour integration to extend the previous

expressions for the surface Green's function [82, 86, 101-103] also to the important case of rank-deficient hopping matrices.

We start from Eq. (3.12) and insert $U_{<} (U_{<})^{-1} = \mathbb{1}$: ⁶

$$\begin{aligned}
G_{j,j'}^{\text{r}} &= \frac{1}{2\pi i} \oint dz \frac{z^{j-j'}}{(E + i\eta - H_0) z - H_1 z^2 - H_{-1}} \\
&= \frac{1}{2\pi i} \oint dz U_{<} (U_{<})^{-1} \frac{z^{j-j'}}{\tilde{H}(E + i\eta, z)} \\
&= \frac{1}{2\pi i} \oint dz U_{<} \frac{z^{j-j'}}{\tilde{H}(E + i\eta, z) U_{<}}.
\end{aligned} \tag{3.26}$$

The denominator $\tilde{H}(E + i\eta, z) U_{<}$ now has columns $\tilde{H}(E + i\eta, z) \mathbf{u}_{i,<}$ with zeroes at $\lambda_{i,<}$. By carefully counting the multiplicities of the zeroes in the numerator and the denominator of the inverse $(\tilde{H}(E + i\eta, z) U_{<})^{-1}$, we can show that (see appendix D.2)

$$\begin{aligned}
G_{j,j'}^{\text{r}} &= \frac{1}{2\pi i} \oint dz U_{<} \Lambda_{<}^{j-j'} \frac{1}{\tilde{H}(E + i\eta, z) U_{<}} \\
&= U_{<} \Lambda_{<}^{j-j'} U_{<}^{-1} \frac{1}{2\pi i} \oint dz \frac{1}{(E + i\eta - H_0) z - H_1 z^2 - H_{-1}} \\
&= U_{<} \Lambda_{<}^{j-j'} U_{<}^{-1} G_{j',j'}^{\text{r}} \quad \text{for } j' \leq j.
\end{aligned} \tag{3.27}$$

Setting $j = j' + 1$ and using Eq. (3.9), $G_{j'+1,j'}^{\text{r}} = g_{\text{R}}^{\text{r}} H_{-1} G_{j',j'}^{\text{r}}$, we then obtain the expression:

$$g_{\text{R}}^{\text{r}} H_{-1} = U_{<} \Lambda_{<} U_{<}^{-1} \tag{3.28}$$

valid also for singular H_{-1} . This is the main finding of this chapter.

From Eq. (3.28) we immediately obtain an expression for the self-energy of the lead ⁸, $\Sigma_R = H_1 g_R^r H_{-1} = H_1 U_{<} \Lambda_{<} U_{<}^{-1}$. If the surface Green's function g_R^r itself is needed, it can be calculated either by multiplying Eq. (3.28) with $(H_{-1})^{-1}$ from the right, if H_{-1} is invertible, or through Eq. (3.10), which is also valid for rank-deficient hopping matrices.

In an analogous way we can show that every left eigenvector $\bar{\mathbf{u}}$ of $H_1 g_R^r$ with eigenvalue λ is a left eigenvector of the quadratic eigenvalue equation with eigenvalue $1/\lambda$. Consequently, we find that $H_1 g_R^r$ can be written in terms of left-going solutions:

$$H_1 g_R^r = (\bar{U}_{>}^\dagger)^{-1} \Lambda_{>}^{-1} \bar{U}_{>}^\dagger \quad (3.29)$$

If we were to calculate the advanced surface Green's function g_R^a , we would choose in Eq. (3.12) a negative infinitesimal shift, $E - i\eta$. The subsequent calculation proceeds in the same fashion as for the retarded Green's function, except that the propagating modes are handled differently: $-i\eta$ now selects left-moving instead of right-moving propagating modes. In contrast, the decaying modes are unchanged with respect to the retarded case.

We have found expressions for the surface Green's function in terms of eigenstates of the infinite wire. It should be emphasized that these expressions are fundamentally different from the normal eigenstate representation of the Green's function, Eq. (2.41). In the usual eigenstate representation, the Green's function is expressed in terms of all eigenstates of the system, and thus all energies $E_{n,k}$. In contrast, the expressions derived in this chapter express the surface Green's function in terms of eigenstates for a single energy E . The Green's function is in this case not only expressed in terms of physical eigenstates of the (infinite) system, but also unphysical ones: The decaying eigenstates grow exponentially in one direction and thus cannot be normalized. The fact that the eigenstates for a single energy E are enough to calculate the surface Green's function make these expressions very useful for numerical computations. A summary of the algorithm based on the results of this section can be found in Appendix D.3.1.

19.2.7 3.3 A more stable algorithm based on the Schur decomposition

19.2.8 3.3.1 Failure of the eigendecomposition based algorithm

The expressions for the surface Green's function of a lead are well-suited for a numerical implementation. However, it turns out that numerics based on these expressions can fail for certain, important cases. In order to illustrate this failure and to trace its origin, we will consider the example of an electronic wire in a magnetic field. The lessons learned from this example will then serve as guidelines for developing a more stable algorithm.

We will consider the case of a two-dimensional electron gas with a confinement potential of the form $V(x, y) = V(y)$, where

$$V(y) = \begin{cases} 0 & \text{for } 0 < y < W \\ \infty & \text{else} \end{cases} \quad (3.30)$$

¹²⁶ Here we have to assume without proof that $U_{<}$ is invertible. It is not obvious that the $N_{u.c.}$ rightgoing solutions $\mathbf{u}_{i,<}$ of the quadratic eigenproblem (3.14a) are linearly independent: For example, we have already seen that they are not necessarily orthogonal, as they are solutions to different Hamiltonians $H(\lambda_{i,<})$.

The eigenvectors of the linearized eigenproblem, $(\mathbf{u}_{i,<}, \lambda_{i,<} \mathbf{u}_{i,<})^t$ do form a linearly independent set, as they are solutions to a (generalized) linear eigenproblem. Unfortunately, this by itself does not prove the linear independence of the $\mathbf{u}_{i,<}$ alone.

⁷ This is uniquely defined, as $G_{j',j'}^r = (E - H_0 - H_1 g_R^r H_{-1} - H_{-1} g_L^r H_1)^{-1}$ is invertible.

⁸ Here, we assume that the hopping from lead l to the scattering region, H_{Sl} , is given by H_1 . Such a situation can always be achieved by formally assigning the first unit cell of the lead to the scattering region.

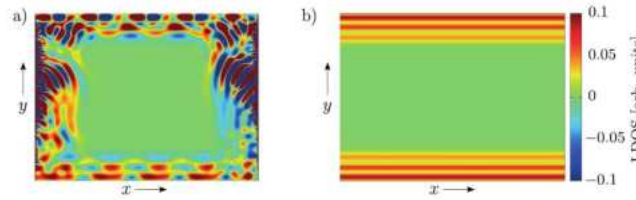


Figure 3.4 - Local density of states of wire in a perpendicular magnet field B , obtained numerically using the lead Green's functions calculated with (a) the eigendecomposition based algorithms of Section D.3.1 and (b) the Schur decomposition based algorithms of Section D.3.2. The parameters used were $W/l_{\text{cycl}} = 9$, $W = 150a$ and $E = 0.15 \frac{\hbar^2}{2ma_0^2}$, where a_0 is the lattice constant of the discretization and m the effective electron mass.

i.e. the electron gas is restricted to a wire of width W . Furthermore we assume a homogeneous magnetic field B perpendicular to the plane of the electron gas. Such a geometry is a simplified version of the systems considered e.g. for weak localization or quantum Hall effect, and as such reliable numerical solutions are of importance.

The Hamiltonian of the system is given by

$$H = \frac{1}{2m}(-i\hbar\nabla - e\mathbf{A}(\mathbf{x}))^2 + V(y) \quad (3.31)$$

where m is the effective mass of the system and e the electronic charge. We choose the vector potential $\mathbf{A}(\mathbf{x})$ in x -direction, $A_y = 0$, $A_x(y) = -By$. After discretizing the Hamiltonian (for details, see Appendix F), the Hamiltonian matrix of the unit cell H_0 is independent of the magnetic field. The magnetic field dependence is contained in the hopping matrices: The hopping matrix H_1 is a diagonal matrix with entries $(H_1)_{nn} = e^{i\alpha n}$, where α is some phase related to the magnetic field B and the lattice constant of the discretization. Thus, H_1 is invertible ($(H_1)^{-1} = H_1^\dagger = H_{-1}$) and we can use the expressions for invertible hopping matrices. We will measure the strength of the magnetic field in terms of the cyclotron radius l_{cycl} with respect to the wire width W , where $l_{\text{cycl}} = \frac{p_F}{eB}$ [111], with p_F the Fermi momentum.

Wires are a rather useful tool for testing numerical transport algorithms: Due to the translational symmetry in x direction, observables must be constant along the wire direction. Any spatial dependence along the wire is thus a clear indication of numerical errors. Furthermore, transport in large magnetic fields, i.e. in the quantum Hall regime, is very convenient for checking numerical calculations, as the wave functions of the edge states are relatively simple. Programming errors or failures of the numerics can then often be detected by visual inspection of observables, such as the local density of states.

In Fig. 3.4 we show the LDOS of a wire in magnetic field. The LDOS was calculated using the recursive Green's function technique first with the lead Green's function calculated from the expressions of the previous section (Fig. 3.4(a)), and second with the lead Green's function calculated with the stable algorithm developed in the remainder of the chapter (Fig. 3.4(b)). Fig. 3.4(b) shows nicely three edge modes forming a positive LDOS that is constant in x -direction, and we can consider this solution to be exact within the numerical limits. In contrast, Fig. 3.4(a) shows some complex spatial behavior of the LDOS with positive and even negative values, violating the simple physical considerations for wires mentioned above. When moving away from the leads, further into the region calculated via the RGF algorithm, some of the unphysical states introduced by the incorrect lead Green's function decay and the LDOS starts to resemble the exact solution. In fact, if we make the region calculated via the RGF algorithm longer, eventually all the unphysical solutions will decay so much that the LDOS will finally start to resemble the exact solution. This is a clear indication that indeed the problem lies in the (failed) calculation of the lead Green's function. Obviously this failure has catastrophic consequences here.

As we will show below, the failure of the algorithm of Section D.3.1 can be attributed to the inversion of $U_{<}$, that becomes ill-conditioned for the problem under discussion. To this end, we first define some mathematical measure in order to quantify how wellposed the problem of inversion is numerically. The condition number of a matrix A is defined as

$$\text{cond}(A) = \|A\|_2 \cdot \|A^{-1}\|_2 \quad (3.32)$$

where $\|\dots\|_2$ is the matrix 2-norm [112]. The numerical error in inverting a matrix A can then be estimated as $\approx \varepsilon \text{cond}(A)$, where ε is the machine precision. On a typical computer, the standard floating point data type is an eight byte double with $\varepsilon \approx 10^{-16}$. Numerically, the condition number can be estimated from the singular value decomposition of A as the ratio of the largest and the smallest singular value [112].

Another useful measure in this context is the rank of matrix A , $\text{rank}(A)$. If the columns (or rows) of a $m \times n$ -matrix A form a set of linearly independent vectors, the matrix has full rank, i.e. $\text{rank}(A) = \min(m, n)$. A square matrix A is invertible iff it has full rank. Numerically, we can calculate the approximate rank of a matrix A by counting the non-zero singular values of A [112]. This rank is only approximate as we need to specify some numerical tolerance in order to check whether a numerically computed singular value is zero, and the rank will depend on this tolerance. Still, the approximate rank can quantify whether the columns (or rows) of a matrix are linearly dependent or at least close to linearly dependent. The relation of the condition number and the approximate rank is obvious from the way they are calculated numerically: The condition number will become large, if at least a single singular value is close to zero, i.e. if a matrix has not full rank. In return, the approximate rank depends on the total number of singular values close to zero and thus gives a further estimate on the character of the ill-conditioned problem.

In Fig. 3.5(a) we show the condition number of $U_{<}$ as a function of $W/l_{\text{cycl}} \propto B$. For small magnetic field, the condition number of $U_{<}$ is small, indicating a wellconditioned problem. When the magnetic field increases and W/l_{cycl} approaches 1 -

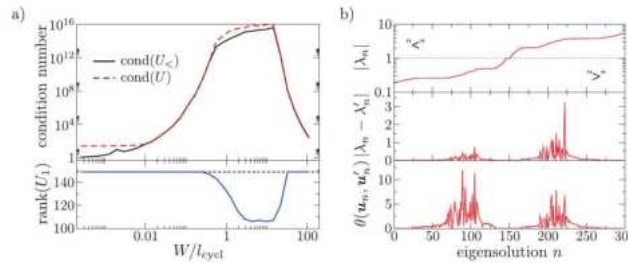


Figure 3.5 - (a) condition numbers (upper panel) and approximate rank (U_1) (lower panel) for the eigendecomposition based algorithm as a function of magnetic field strength W/l_{cycl} . The dashed line in the lower panel shows the value $\text{rank}(U_1)$ for linearly independent eigenvectors. (b) Detailed analysis of error bounds of individual eigenvalues λ_n for fixed magnetic field, $W/l_{\text{cycl}} = 9$: absolute value of λ_n (upper panel), error bound to the eigenvalue λ_n (middle panel), and error bound to the eigenvector $\mathbf{u}_n$ (lower panel) as a function of eigenvalue index n . The eigenvalues and -vectors have been sorted such that the right-going set (" $<$ ") precedes the left-going set (" $>$ "). The calculations were done using the same set of parameters as in Fig. 3.4.

i.e. the magnetic field begins to influence the wave functions of the wire beyond simple perturbations - $\text{cond}(U_{<})$ takes values in the range of $10^{10} - 10^{16}$. Numerically, the inversion of $U_{<}$ becomes meaningless in this case and so does the calculated surface Green's function. Only for very large magnetic field the matrix $U_{<}$ becomes wellconditioned again. The magnetic field range for which $U_{<}$ is ill-conditioned increases as the width W of the system becomes larger, i.e. the numerical error increases as the system size gets larger. This is an indication

that the inversion of $U_{<}$ becomes ill-defined due to numerical errors, rather than $U_{<}$ itself being non-invertible.

The $N_{\text{u.c.}} \times N_{\text{u.c.}}$ matrix $U_{<}$ is constructed from selected eigenvectors $(\mathbf{u}_n, \mathbf{u}_n \lambda_n)^\dagger$ of the $2N_{\text{u.c.}} \times 2N_{\text{u.c.}}$ linearization of the quadratic eigenvalue problem, Eq. (3.14a). The full matrix U of all eigenvectors $(\mathbf{u}_n, \mathbf{u}_n \lambda_n)^\dagger$ can be written as

$$U = \begin{pmatrix} U_{<} & U_{>} \\ U_{<} \Lambda_{<} & U_{>} \Lambda_{>} \end{pmatrix} = \begin{pmatrix} U_1 & U_2 \end{pmatrix}. \quad (3.33)$$

In Fig. 3.5(a) we also show that $\text{cond}(U) \approx \text{cond}(U_{<})$, in order to demonstrate that the failure of numerics is not due to choosing the submatrix $U_{<}$, but is inherent to the linearized quadratic eigenproblem itself. This can also be seen from the approximate rank(U_1) in the lower panel of Fig. 3.5(a). Whereas the rank of U_1 should be $N_{\text{u.c.}}$, as the $N_{\text{u.c.}}$ eigenvectors span a $N_{\text{u.c.}}$ -dimensional subspace, the numerically computed U_1 is rather rank-deficient in the magnetic field range where the numerics breaks down.

It should be noted that the linearized (generalized) eigenproblem of Eq. (3.15) involves computing the eigenvalues and -vectors of non-Hermitian matrices. In contrast to Hermitian matrices, the eigenvalues and -vectors of non-Hermitian matrices can be rather ill-conditioned [112]. For example, the eigenvectors of a non-Hermitian matrix need not be orthogonal, and thus the matrix U , while in principle invertible, could be easily near-singular.

We therefore analyze the error bounds of the numerically computed eigenvalues and -vectors. Suppose that λ'_n and $\mathbf{u}'_n$ is the exact eigensolution. Then the error in the numerically computed eigenvalue λ_n and -vector $\mathbf{u}_n$ can be quantified as $|\lambda_n - \lambda'_n|$ and the acute angle $\theta(\mathbf{u}_n, \mathbf{u}'_n)$, respectively [107]. In Fig. 3.5(b) we show the errors on the n -th eigenvalues and -vector together with the absolute value of the eigenvalue $|\lambda_n|$ for all $2N_{\text{u.c.}}$ eigensolutions in the regime where $U_{<}$ is ill-conditioned ($W/l_{\text{cycl}} = 9$, corresponding to the density plot in Fig. 3.4(a)). We find two clusters of eigensolutions in both the right- and left-going solutions with large error bounds: The error in the eigenvalues is of the order of the eigenvalue itself, and the angles between numerical and exact eigenvectors are of the order of π . Obviously, these solutions are not reliable. Note that these numerical problems are restricted to decaying solutions, whereas the propagating solutions (with $|\lambda_n| = 1$) have small error bounds. Presumably this is due to the fact that the propagating modes are solutions to a Hermitian Hamiltonian $H(k)$.

Summarizing the discussion, the failure of the eigendecomposition based algorithms for computing lead Green's functions is due to numerical errors in computing the eigenvalues and -vectors of a non-Hermitian matrix. This leads numerically to linearly dependent eigenvectors and thus to non-invertible eigenvector matrices.

Having established the cause of the failure of numerics as numerical errors due to finite precision, we are left with two options: First, we could resort to floating point arithmetics with higher precision. However, there are no readily available numerical routines beyond double precision, and extending the established linear algebra routines to higher precision is a major task and not trivial. Moreover, it is not clear whether higher precision would solve all problems. The condition numbers of Fig. 3.5(a) were calculated numerically. Thus they are also limited by the floating point precision and can only give a lower bound - the true condition number could be much larger. Due to these difficulties associated with higher precision we choose the second option: Develop an alternative algorithm that does not need to calculate all individual eigenvectors and -values.

19.2.9 3.3.2 Surface Green's function from the Schur decomposition

In order to develop a more stable algorithm for calculating the surface Green's function of a lead, we introduce the concept of an invariant subspace [112]. An invariant subspace of a

matrix A is a vector space V such that for every $v \in V$, also $Av \in V$. For example, a subspace spanned by eigenvectors of A is an invariant subspace.

Starting point for a more stable algorithm is the observation that it is indeed not necessary to calculate individual eigensolutions, but only the invariant subspace associated with a cluster of eigenvalues. The columns of the eigenvector matrix U_1 ,

$$U_1 = \begin{pmatrix} U_{11} \\ U_{21} \end{pmatrix} = \begin{pmatrix} U_{<} \\ U_{<} \Lambda_{<} \end{pmatrix}, \quad (3.34)$$

form an invariant subspace $\text{range}(U_1)$. Any invertible $N_{\text{u.c.}} \times N_{\text{u.c.}}$ matrix A generates a basis transformation of this subspace and vice versa, so that the columns of $U'_1 = U_1 A$ span the same subspace, $\text{range}(U'_1) = \text{range}(U_1)$. Then we find that

$$\begin{aligned} U'_{21} (U'_{11})^{-1} &= U_{21} A (U_{11} A)^{-1} \\ &= U_{<} \Lambda_{<} A A^{-1} U_{<}^{-1} \\ &= U_{<} \Lambda_{<} U_{<}^{-1}. \end{aligned} \quad (3.35)$$

Therefore, it is enough to compute the invariant subspace associated with the eigenvalues $\lambda_{n,<}$ in order to compute the surface Green's function of the lead.

In the example of the previous section, the failure of the numerics was linked to a rank-deficient matrix U_1 , i.e. a too small invariant subspace. Since the eigenvectors of a non-Hermitian matrix do in general not form an orthogonal basis, the eigenvectors can be nearly linearly dependent and numerical errors may lead to rank-deficient eigenvector matrices. Therefore, we seek an alternative method to compute the invariant subspace spanned by U_1 .

We start with discussing the case of invertible hopping matrices. The linearized quadratic eigenproblem can in this case be written as an ordinary eigenproblem of the form $A(\mathbf{u}_n, \lambda_n \mathbf{u}_n)^t = \lambda_n (\mathbf{u}_n, \lambda_n \mathbf{u}_n)^t$, as in Eq. (3.16). Then, there is a unitary transformation

$$Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} = (Q_1 \quad Q_2) \quad (3.36)$$

such that

$$Q^\dagger A Q = T = \begin{pmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{pmatrix} \quad (3.37)$$

where T is an upper triangular matrix with the eigenvalues λ_n of A on the diagonal. This decomposition is called Schur decomposition. Furthermore, the unitary transformation Q can be chosen such that the eigenvalues λ_n appear in any desired order along the diagonal. The first M columns of Q , $1 \leq M \leq 2N_{\text{u.c.}}$, then form an invariant subspace with respect to the first M eigenvalues, $\lambda_1 \dots \lambda_M$ [112].

Suppose now that the eigenvalues are ordered such that all eigenvalues $\lambda_{n,<}$ are located in the upper left $N \times N$ -block T_{11} . If T_{11} and T_{22} do not share common eigenvalues, the triangular matrix T can be block-diagonalized [112], i.e. there is a non-singular transformation

$$Y = \begin{pmatrix} \mathbb{I} & Y_{12} \\ 0 & \mathbb{I} \end{pmatrix} \quad (3.38)$$

such that

$$Y^{-1} T Y = Y^{-1} Q^\dagger A Q Y = \begin{pmatrix} T_{11} & 0 \\ 0 & T_{22} \end{pmatrix} \quad (3.39)$$

After this transformation, the $2N_{\text{u.c.}} \times 2N_{\text{u.c.}}$ eigenproblem is reduced to two separate $N_{\text{u.c.}} \times N_{\text{u.c.}}$ eigenproblems for the eigenvalues $\lambda_{n,<}$ and $\lambda_{n,>}$, respectively. The matrix A can thus be reduced to diagonal form - i.e. the full eigenproblem can be solved-by a block-diagonal transformation

$$X = \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \quad (3.40)$$

such that

$$X^{-1}Y^{-1}Q^\dagger A Q Y X = \begin{pmatrix} \Lambda_{<} & 0 \\ 0 & \Lambda_{>} \end{pmatrix} \quad (3.41)$$

The matrix of eigenvectors is then given by $U = QXY$. Hence, we find

$$\begin{aligned} g_{\text{R}}^\dagger H_{-1} &= U_{<} \Lambda_{<} U_{<}^{-1} = Q_{21} X_1 (Q_{11} X_1)^{-1} \\ &= Q_{21} Q_{11}^{-1}. \end{aligned} \quad (3.42)$$

The surface Green's function can thus be calculated from the orthonormal basis vectors of the invariant subspace associated with $\lambda_{n,<}$ obtained from the Schur decomposition, instead of the full eigendecomposition.

It should be noted that in the derivation of Eq. (3.42) it is necessary to assume that T_{11} and T_{22} do not share any eigenvalues. Hence, Eq. (3.42) is not valid in a situation as depicted in Fig. 3.3(c), when propagating modes share the same eigenvalue $\lambda_n = \lambda_{n'}$, but have group velocities v_n and $v_{n'}$ with opposite signs. In such a case it is necessary to explicitly shift the real energy either along the real or the imaginary axis (i.e. with a finite $i\eta$), lifting the eigenvalue degeneracy in order to apply Eq. (3.42). In any case, it is rather unlikely to hit such a situation numerically.

Finally, we turn to the case of a singular hopping matrix H_1 . The linearized quadratic eigenproblem forms in this case a generalized eigenvalue problem $A(\mathbf{u}_n, \lambda_n \mathbf{u}_n)^\dagger = \lambda_n B(\mathbf{u}_n, \lambda_n \mathbf{u}_n)^\dagger$, as in Eq. (3.15). Then, there are two unitary transformations Q and Z such that

$$Q^\dagger A Z = T \quad \text{and} \quad Q^\dagger B Z = S \quad (3.43)$$

where T and S are upper triangular [112] (generalized Schur decomposition). The eigenvalues of the generalized eigenproblem are thence given by $\lambda_n = t_{nn}/s_{nn}$, where t_{nn} (s_{nn}) are the diagonal elements of T (S). Again, we assume that the eigenvalues are ordered such that the $\lambda_{n,<}$ are contained in the leading diagonal positions of T and S . In a similar fashion as in the case of invertible hopping matrices, we find that the Green's function can be expressed in terms of the subspace spanned by the first $N_{\text{u.c.}}$ orthonormal columns of Z :

$$g_{\text{R}}^\dagger H_{-1} = Z_{21} Z_{11}^{-1} \quad (3.44)$$

A summary of the numerical algorithms based on the results of this section can be found in Appendix D.3.2.

19.2.10 3.3.3 Numerical Stability

The Schur decomposition based algorithm developed in the previous sections is superior to the standard, eigendecomposition based algorithms in terms of numerical stability. Although we will not proof this statement in a strict, mathematical sense, we will give some arguments why this is the case below.

In Section 3.3.1, the failure of the standard algorithm was linked to the rankdeficiency of the matrix U_1 . The numerically computed eigenvectors did not form a linearly independent set, and thus did not span the full invariant subspace range (U_1) associated with the eigenvalues $\lambda_{n,<}$.

In contrast to the eigendecomposition of a non-Hermitian matrix, the (generalized) Schur decomposition can be computed more stably, as it involves only unitary transformations [112]. In addition, reordering the (generalized) Schur form can be done in a numerically stable fashion, too. Since the transformation matrix Q/Z is unitary, its columns form an orthonormal basis set, as opposed to the basis spanned by the eigenvectors. Therefore, the Schur decomposition yields an orthonormal basis set for the invariant subspace range (U_1) = range(Q_1/Z_1). Due to the orthogonality of the N basis vectors, even when obtained numerically, the computed subspace will always have dimensionality N , in contrast to the standard algorithm.

Apart from the correct dimensionality, the computed invariant subspace must be associated with the correct eigenvalues. Note that we need not determine every single eigenvalue accurately, but we only need to cluster the eigenvalues into left- and rightgoing sets $\{\lambda_{n,>}\}$ and $\{\lambda_{n,<}\}$. The border of these sets is within the propagating modes where $|\lambda_n| = 1$. In the examples considered in this study, we have found that the error bounds on the eigenvalues and -vectors around $|\lambda_n| = 1$ are rather small (for an example, see Fig. 3.5(b)). Presumably this is due to the fact that the propagating modes (i.e. $|\lambda_n| = 1$) are governed by a Hermitian operator. Hence, the eigenvalues and -vectors separating left- and right-going sets can be calculated accurately and we can thus expect to obtain an accurate clustering. Indeed, the examples in the next section will show that we can calculate the invariant subspace range (U_1) = range(Q_1/Z_1) very accurately.

Finally, after calculating the invariant subspace we need to invert the block Q_{11}/Z_{11} . Unfortunately we cannot prove that this matrix is always well-conditioned-whereas the full matrix Q/Z is unitary and thus well-conditioned, the subblock Q_{11}/Z_{11} is in general not unitary. However, we found that it was always well-behaved in our examples, as shown in the next section.

In summary, while eigendecomposition and Schur decomposition based algorithms need to perform several computational steps alike, such as the classification of eigenvalues, the Schur decomposition performs significantly better in one crucial step, the computation of the invariant subspace.

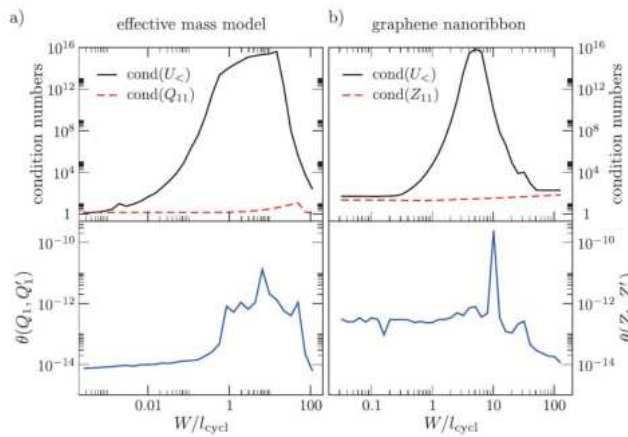


Figure 3.6 - Condition numbers (upper panels) and error in the computed invariant subspaces (lower panels) for the eigendecomposition and Schur based algorithms for the computation of the lead Green's function. The quantities are shown as a function of W/l_{cycl} for two benchmark systems involving a magnetic field: (a) a wire in effective mass approximation in a perpendicular magnetic field B , parameters as in Fig. 3.4. (b) a zigzag graphene nanoribbon in a perpendicular magnetic field B , with a width W corresponding to 302 atoms across the nanoribbon, and energy $E = 0.39t$.

19.2.11 3.3.4 Examples

In order to analyze their numerical performances, we apply both the eigendecomposition and Schur decomposition based algorithm to different systems. As a benchmark for the algorithms for invertible hopping matrices we take up again the system of Section 3.3.1, a wire in a magnetic field. For the case of singular hopping matrices, we consider a zigzag graphene ribbon described in nearest-neighbor tight-binding approximation⁹ in a perpendicular, homogeneous magnetic field B . In both cases we will again measure the magnetic field strength as W/l_{cycl} .

In the upper panels of Fig. 3.6 we show the condition numbers for the matrices that need to be inverted in the various algorithms: $U_{<}$ in the eigendecomposition

based algorithms, and Q_{11}, Z_{11} in the Schur decomposition based algorithms. In both examples, effective mass approximation and graphene, $\text{cond}(U_{<}) \approx 10^{10} - 10^{16}$ in a rather large range of magnetic field strength. In contrast, the matrices Q_{11} and Z_{11} are always well-behaved, with condition numbers smaller than 10^2 . As a consequence, the inversion of Q_{11} and Z_{11} is not problematic numerically, whereas the corresponding inversion of $U_{<}$ fails in the eigendecomposition based algorithms.

In the previous section we argued that the accuracy of the computed invariant subspace in the Schur decomposition based algorithms is expected to be superior to the eigendecomposition approaches. In order to demonstrate this, we define the angle between subspaces as a measure of the error in the computed subspace:

$$\theta(V, W) = \max_{\substack{v \in V \\ v \neq 0}} \max_{\substack{w \in W \\ w \neq 0}} \theta(v, w) \quad \text{for subspaces } V, W \text{ with } \dim V = \dim W \quad (3.45)$$

In the lower panels of Fig. 3.6 we show the angle of the subspaces spanned by the columns of the computed Q_{11} (Z_{11}) and the true Q'_{11} (Z'_{11}). As expected, the accuracy of the subspaces computed from the Schur decomposition is very high, in contrast to the eigendecomposition that even yielded subspaces with the wrong dimensionality. Consequently, the Schur decomposition based algorithms perform very well for the whole magnetic field range.

Finally, we discuss the run times of the different algorithms. Note that the first step of computing numerically the eigendecomposition also involves computing the Schur decompositions [107]. This computational step is most time-consuming in both algorithms. As a consequence, we find similar run times: sometimes the Schur decomposition based algorithm is faster, sometimes the eigendecomposition based algorithm, depending on the system under consideration. However, we found only differences up to 20%, and thus both algorithms can be considered equivalent with respect to run time.

19.2.12 3.4 Summary

In this chapter we have derived expressions for calculating the surface Green's function of a lead valid for arbitrary tight-binding models, including singular hopping matrices. A direct implementation of these eigendecomposition based expressions has been found numerically unstable for certain cases, and an alternative algorithm based on the Schur decomposition has been developed to remedy these problems. Since the Schur decomposition based algorithm is in general more stable than the eigendecomposition based algorithm and of comparable speed, it should be the method of choice.

In the literature, there is only one other algorithm that can compute the lead Green's function for arbitrary tight-binding models [86].¹⁰ There, the problem of computing

¹²⁹ The tight-binding Hamiltonian is given by $H = \sum_{ij} t| i \rangle \langle j |$, where the summation runs over nearest neighbors only, and $t \approx 2.7\text{eV}$. For details on the model for graphene, see Appendix G

¹²¹⁰ In fact, when the work on the contents of this chapter was started, there was no general algorithm

the surface Green's function for a lead with singular hopping matrices is mapped to an equivalent problem with invertible hopping matrices. The algorithms of this chapter are conceptually much simpler and require fewer computational steps than the algorithm of Ref. [86]. Still, it might be interesting to compare the accuracy of these algorithms on different models. In any case, the algorithm of Ref. [86] can also benefit from the stability of the Schur decomposition based algorithm developed in this chapter.

For transport properties, it is necessary to calculate Green's functions at real energies E , and thus the derivations of this chapter assumed E to be real. For computing equilibrium densities, it is however often convenient numerically to compute Green's functions energies E in the complex plane [84, 113]. When E is complex there are no propagating modes, and the eigenmodes $\mathbf{u}_n$ can be grouped into equally large sets of left- and right-decaying modes [97]. From these the lead Green's functions can be calculated as described in this chapter.

19.3 Chapter 4 Optimal block-tridiagonalization of matrices for quantum transport

19.3.1 4.1 Introduction

From the lead self-energies Σ_l^r derived in the previous chapter, we can now turn to calculating the retarded Green's function G_S^r of the scattering region. As discussed at the end of Chapter 2, the direct calculation of G_S^r by means of Eq. (2.111).

$$G_S^r = (E - H_S - \Sigma^r)^{-1} \quad (4.1)$$

is usually computationally intractable. There is however a number of quantum transport algorithms for two-terminal systems - i.e. systems with only two leads, that we

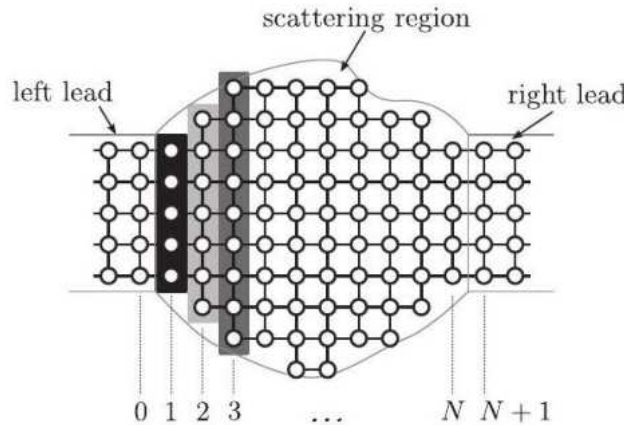


Figure 4.1 - Natural ordering of lattice points in a square finite difference grid leading to a block-tridiagonal matrix structure. The different matrix blocks are marked in alternating shades of grey.

denote as left (L) and right (L) - that circumvent this problem by making use of the sparsity of the Hamiltonian matrix H_S . In particular, these algorithm are based on the fact that the full Hamiltonian matrix H can be written in block-tridiagonal form:

available. Ref. [86] was published after the work on the algorithms of this chapter was finished.

dynamics [124]. Still, such algorithms are not as widely used as the large class of algorithms, that are directly based on the block-tridiagonal form of the Hamiltonian. In order to illustrate the typical computational tasks of this class of algorithms, we briefly state, as a representative example, the computational steps of the RGF algorithm. The details of this algorithm, including its extension to nonequilibrium, can be found in Appendix B.

The RGF technique is based on the Dyson's equation $G^r = G_0^r + G_0^r V G^r$, building up the system block by block, as depicted in Fig. 4.2. Let $G^{r,R(i)}$ denote the Green's function for the system containing all blocks $\geq i$. Then, the Green's function $G^{r,R(i-1)}$ is related to $G^{r,R(i)}$ by

$$G_{i-1,i-1}^{r,R(i-1)} = \left(E - H_{i-1,i-1} - H_{i-1,i} G_{i,i}^{r,R(i)} H_{i,i-1} \right)^{-1} \quad (4.3)$$

and

$$G_{N+1,i-1}^{r,R(i-1)} = G_{N+1,i}^{r,R(i)} H_{i,i-1} G_{i-1,i-1}^{r,R(i-1)}. \quad (4.4)$$

Starting from $G_{N+1,N+1}^{r,R(N+1)} = g_R^r$, the surface Green's function of the right lead, N slices are added recursively, until $G^{r,R(1)}$ has been calculated. The blocks of the Green's function of the full system necessary for transport are then given by

$$G_{0,0}^r = \left((g_L^r)^{-1} - H_{0,1} G_{1,1}^{r,R(1)} H_{1,0} \right)^{-1} \quad (4.5)$$

and

$$G_{N+1,0}^r = G_{N+1,1}^{r,R(1)} H_{1,0} G_{0,0}^r, \quad (4.6)$$

where g_L^r is the surface Green's function of the left lead. $G_{0,0}^r$ and $G_{N+1,0}^r$ are sufficient to calculate the conductance, as seen from Eqs. (2.117) and (C.31).

Each step of the algorithm performs inversions and matrix multiplications with matrices of size M_i . Since the computational complexity of matrix inversion and multiplications scales as M_i^3 , the complexity of the RGF algorithm is $\sum_{i=0}^{N+1} M_i^3$. Thus, it scales linearly with the "length" N , and cubically with the "width" M_i of the system. This scaling also applies to most of the other transport algorithms mentioned above.

While for particular cases general transport algorithms, such as the RGF algorithm, cannot compete with more specialized algorithms, such as the modular recursive Green's function technique [125, 126] that is optimized for special geometries, they are very versatile and easily adapted to many geometries - provided that the system has only two leads that are arranged collinearly.

The objective of this chapter is twofold: First, we intend to lift the discussed restrictions of the established quantum transport algorithms. We do this by bringing the Hamiltonian matrix H into a block-tridiagonal form suitable for quantum transport also for complex geometries, such as non-collinear leads or multi-terminal structures, that would otherwise need the development of specialized algorithms. Second, we improve the matrix H such that the block-tridiagonal form is optimal for transport. Although the block-tridiagonal structure of H , Eq. (4.2), that arises naturally in many problems appears to have a small "width" and thus seems to be quite suitable for transport algorithms, optimizing the matrix structure further may lead to significant speed-ups even in the two-terminal case, as we show below.

We achieve these goals by developing a matrix reordering algorithm based on graph partitioning techniques, that brings an arbitrary tight-binding matrix H into a blocktridiagonal form optimal for quantum transport. To this end, we identify a mathematical measure for this optimality and reformulate the matrix reordering problem in the language of graph theory. Based on graph-theoretical concepts we then develop the block-tridiagonalization algorithm. Finally, we apply this algorithm to various examples and investigate performance of the reordering and

of the RGF algorithm for the reordered Hamiltonian H , demonstrating the flexibility and power of this combined approach.

19.3.2 4.2 Definition of the problem

19.3.3 4.2.1 Definition of the matrix reordering problem

As shown in the previous section, the typical runtime of transport algorithms, proportional to $\sum_{i=0}^{N+1} M_i^3$, does depend on the particular block-tridiagonal structure of H . Therefore, the runtime of these algorithms can be improved in principle by conveniently reordering H with a permutation P ,

$$H' = PHP^{-1}. \quad (4.7)$$

In order to quantify the typical performance of a transport algorithm for a given matrix structure, we define a weight $w(H)$ associated with a matrix H as

$$w(H) = \sum_{i=0}^{N+1} M_i^3, \quad (4.8)$$

where M_i is the size of block $H_{i,i}$. Optimizing the matrix for transport algorithms is then equivalent to minimizing the weight $w(H)$. Since $\sum_{i=0}^{N+1} M_i = N_{\text{grid}}$, where N_{grid} is the total number of grid points, $w(H)$ is minimal, if all M_i are equal, $M_i = N_{\text{grid}} / (N + 2)$. Therefore, a matrix tends to have small weight, if the number N of blocks is large, and all blocks are equally sized. The reordering problem of the matrix H is thus summarized as follows:

Problem 4.1. Matrix reordering problem: Find a reordered matrix H' such that

1. $H'_{0,0}$ and $H'_{N+1,N+1}$ are blocks given by the left and right leads (as required by transport algorithms)
2. H' is block-tridiagonal ($H'_{i,j} \neq 0$, iff $j = i + 1, i, i - 1$),
3. the number N of blocks is as large as possible, and all blocks are equally sized.

In principle, this constrained optimization problem could be solved by generic optimization algorithms, such as Simulated Annealing. However, for larger problems the optimization could take much more time than the actual transport calculation, rendering the optimization process useless. It is therefore necessary to use heuristics especially designed for the problem at hand. To this end, we formulate the matrix reordering problem in the language of graph theory.

19.3.4 4.2.2 Mapping onto a graph partitioning problem

A graph $\mathcal{G}$ is an ordered pair $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is a set of vertices v and $\mathcal{E}$ a set of ordered pairs of vertices $(v_1, v_2) \in \mathcal{V} \times \mathcal{V}$. Such a pair is called an edge. A graph is called undirected, if for every edge $(v_1, v_2) \in \mathcal{E}$ also $(v_2, v_1) \in \mathcal{E}$. Two vertices v_1 and v_2 are called adjacent, if $(v_1, v_2) \in \mathcal{E}$. In order to simplify the notation, we will also consider a vertex v to be adjacent to itself.

There is a natural one-to-one correspondence between graphs and the structure of sparse matrices. For a given $n \times n$ matrix H , we define a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \{1, \dots, n\}$ and $(i, j) \in \mathcal{E}$ iff the entry $H_{ij} \neq 0$. A graph thus stores information about the structure of a matrix, i.e. which entries are nonzero. It does not contain any information about the values of the respective entries, although these may be stored easily along with the graph.

However, for the formulation of the quantum transport algorithms, only the block-tridiagonal form, i.e. the structure of the matrix, is relevant. Hermitian matrices, that are considered in quantum transport, have a symmetric structure of zero and nonzero entries, and therefore the corresponding graphs are undirected.

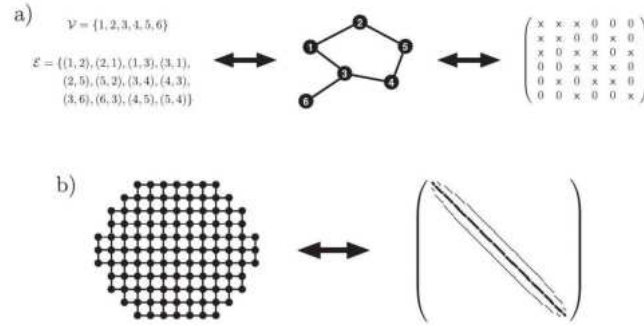


Figure 4.3 - (a) Simple example showing the connection between a graph, its graphical representations with dots and lines, and the zero-nonzero structure of a matrix. (b) Example of a finite difference grid, that can be interpreted as a graph, and the structure of the corresponding matrix. Nonzero entries are shown as black dots.

A graph can be depicted by drawing dots for each vertex v , and lines connecting these dots for every edge (v_1, v_2) , as shown in Fig. 4.3(a). It should be noted that a graphical representation of a tight-binding grid, such as shown in Fig. 4.3(b), can be directly interpreted as a representation of a graph and the corresponding matrix structure.

In terms of graph theory, matrix reordering corresponds to renumbering the vertices of a graph. Since we are only interested in reordering the matrix in terms of matrix blocks (the order within a block should not matter too much), we define a partitioning of $\mathcal{G}$ as a set $\{\mathcal{V}_i\}$ of disjoint subsets $\mathcal{V}_i \subset \mathcal{V}$ such that $\bigcup_i \mathcal{V}_i = \mathcal{V}$ and $\mathcal{V}_i \cap \mathcal{V}_j = \emptyset$ for $i \neq j$. Using these concepts, we can now reformulate the original matrix reordering problem into a graph partitioning problem:

Problem 4.2. Graph partitioning problem: Find a partitioning $\{\mathcal{V}_0, \dots, \mathcal{V}_{N+1}\}$ of $\mathcal{G}$ such that:

1. $\mathcal{V}_0$ and $\mathcal{V}_{N+1}$ contain the vertices belonging to left and right leads,
2. (a) vertices in $\mathcal{V}_0$ and $\mathcal{V}_{N+1}$ are only connected to vertices in $\mathcal{V}_1$ and $\mathcal{V}_N$, respectively,
 (b) for $0 < i < N + 1$, there are edges between $\mathcal{V}_i$ and $\mathcal{V}_j$ iff $j = i + 1, i - 1$,
3. the number $N + 2$ of sets $\mathcal{V}_i$ is as large as possible, and all sets $\mathcal{V}_i$ have the same cardinality $|\mathcal{V}_i|$. A partitioning with all $|\mathcal{V}_i|$ equally sized is called balanced.

A partitioning obeying requirement 4.2 .2 is called a level set with levels $\mathcal{V}_i$ [127]. Level sets appear commonly as an intermediate step in algorithms for bandwidth reduction of matrices [127-130]. These algorithms seek to find a level set of minimal width, i. e. $\max_{i=0 \dots N+1} |\mathcal{V}_i|$ as small as possible which is equivalent to requirement 4.2.3. The main difference between our graph partitioning problem and the bandwidth reduction problem is requirement 4.2.1: In the graph partitioning problem, $\mathcal{V}_0$ and $\mathcal{V}_N$ are determined by the problem at hand, while in the bandwidth reduction problem these can be chosen freely. Due to this difference, bandwidth reduction algorithms can be applied successfully to our graph partitioning problem only for special cases, as we show below.

The term graph partitioning usually refers to the general problem of finding a balanced partitioning $\{\mathcal{V}_i\}$ of a graph and has many applications in various fields such as very-large-scale integration (VLSI) design [131-133], sparse matrix reorderings for LU or Cholesky decompositions [134], or block ordering of sparse matrices for parallel computation [135-140]. In particular, the latter examples also include block-tridiagonal orderings [136, 137]. However, as these reorderings are geared towards parallel computation, they obtain a fixed number N of sets $\mathcal{V}_i$

given by the number of processors of a parallel computer, whereas in our block-tridiagonal reordering the number N should be as large as possible. In addition to that, the constraints on the blocks $\mathcal{V}_0$ and $\mathcal{V}_{N+1}$ (requirement 4.2.1) are again not present there.

As we cannot directly employ existing techniques to solve the graph partitioning problem, we will develop an algorithm combining ideas from both bandwidth reduction and graph partitioning techniques in the subsequent sections: Concepts from bandwidth reduction are employed to construct a level set which is then balanced using concepts from graph partitioning.

19.3.5 4.3 Optimal Matrix reordering by graph partitioning

19.3.6 4.3.1 A local approach—breadth first search

A breadth-first-search (BFS) [141] on a graph immediately yields a level set [127-130]. In our particular example, the level set is constructed as follows:

Algorithm 4.1. Level set construction by breadth-first-search.

A Start from $i = 0$. Then, $\mathcal{V}_i = \mathcal{V}_0$, as the first level is given by the constraints of requirement 4.2.1.

B If there is a vertex in $\mathcal{V}_i$ that is adjacent to a vertex in $\mathcal{V}_{N+1}$, assign all the remaining unassigned vertices into $\mathcal{V}_{N+1}$ and end the algorithm.

C All vertices adjacent to $\mathcal{V}_i$ not contained in the previous levels $\mathcal{V}_i, \mathcal{V}_{i-1}, \dots, \mathcal{V}_0$ are assigned to $\mathcal{V}_{i+1}$.

D Continue at step B with $i = i + 1$.

Note that the sets $\{\mathcal{V}_i\}$ form a level set by construction - a set $\mathcal{V}_i$ may only have vertices adjacent to $\mathcal{V}_{i-1}$ and $\mathcal{V}_{i+1}$. The construction by BFS not only obtains the

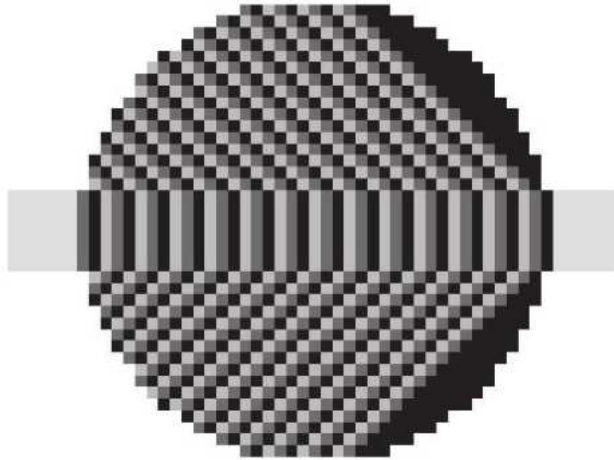


Figure 4.4 - Level set created by a BFS starting from $\mathcal{V}_0$. Different levels are shown in alternating shades of grey.

number of levels $N + 2$ for a particular realization, but yields a more general information:

Lemma 4.3. The number of levels $N + 2$ in the level set constructed by algorithm 4.1 is the maximum number of levels compatible with the constraints on the initial and final level $\mathcal{V}_0$ and $\mathcal{V}_{N+1}$ for a graph $\mathcal{G}$.

This can be seen from the fact that a BFS finds the shortest path in the graph between the initial sets $\mathcal{V}_0$ and $\mathcal{V}_{N+1}$, $(v_0, v_1, \dots, v_i, \dots, v_{N+1})$ where $v_0 \in \mathcal{V}_0$ and $v_{N+1} \in \mathcal{V}_{N+1}$. Any vertex on this shortest path can be uniquely assigned to a single level $\mathcal{V}_i$ and it would not be compatible with a larger number of levels than $N + 2$.

Algorithm 4.1 not only yields the maximum number of levels: All vertices contained in the first n levels of the BFS must be contained in the first n levels of any other level set:

Lemma 4.4. Let $\{\mathcal{V}_0, \mathcal{V}_1, \dots, \mathcal{V}_{N+1}\}$ be a level set constructed by algorithm 4.1, and $\{\mathcal{V}'_0, \mathcal{V}'_1, \dots, \mathcal{V}'_{N'+1}\}$ another level set consistent with the requirements of problem 4.2 with

$N' \leq N$. Then $\mathcal{V}_0 \cup \mathcal{V}_1 \cup \dots \cup \mathcal{V}_n \subset \mathcal{V}'_0 \cup \mathcal{V}'_1 \cup \dots \cup \mathcal{V}'_n$ for $0 \leq n \leq N' + 1$.

The statement is proved by induction. It is true trivially for $n = 0$ (because of requirement 1 in problem 4.2) and for $n = N' + 1$ (then the levels cover the whole graph). Suppose now that the statement holds for $n < N'$. Note that for the proof it suffices to show that $\mathcal{V}_{n+1} \subset \mathcal{V}'_0 \cup \mathcal{V}'_1 \cup \dots \cup \mathcal{V}'_{n+1}$. Consider now the set of all vertices adjacent to $\mathcal{V}_n$, adjacent ($\mathcal{V}_n$) = $\{v \in \mathcal{V} \mid v \text{ is adjacent to a } v' \in \mathcal{V}_n\}$. By construction, $\mathcal{V}_{n+1} \subset \text{adjacent}(\mathcal{V}_n)$. Since $\mathcal{V}_n \subset \mathcal{V}'_0 \cup \mathcal{V}'_1 \cup \dots \cup \mathcal{V}'_n$ and $\{\mathcal{V}'_i\}$ is a level set, all vertices adjacent to $\mathcal{V}_n$ must be contained in the set of vertices including the next level, i.e. $\text{adjacent}(\mathcal{V}_n) \subset \mathcal{V}'_0 \cup \mathcal{V}'_1 \cup \dots \cup \mathcal{V}'_n \cup \mathcal{V}'_{n+1}$. But then also $\mathcal{V}_{n+1} \subset \mathcal{V}'_0 \cup \mathcal{V}'_1 \cup \dots \cup \mathcal{V}'_{n+1}$, which concludes the proof.

Thus, the vertices contained in the first n levels of the BFS form a minimal set of vertices needed to construct n levels. However, this also implies that the last level which then covers the remaining vertices of the graph, may contain many more vertices than the average, leading to an unbalanced level set. This is not surprising, since the algorithm does not explicitly consider balancing and only local information is used, i.e. whether a vertex is adjacent to a level or not. An example for this imbalance is shown in Fig. 4.4, where the BFS construction yields a very large last level.

Note that throughout the manuscript we visualize the graph theoretical concepts using examples of graphs obtained from discretizing a two-dimensional structure. However, the ideas and algorithms presented here apply to any graph and are not limited to graphs with coordinate information. Two-dimensional graphs have the advantage of being visualized easily. In particular, the BFS search has an intuitive physical analog: Wave front propagation of elementary waves emanating from the vertices of the initial level $\mathcal{V}_0$.

The problem that a BFS does not yield a balanced partitioning was also noted in the theory of bandwidth reduction. The Gibbs-Poole-Stockmeyer (GPS) algorithm tries to overcome this deficiency by constructing a level set through the combination of two BFS searches starting from the initial and the final levels. However, there the initial and final levels are sought to be furthest apart, contrary to our problem. In general, the GPS construction only yields a balanced level set if the initial and final level are close to furthest apart, as we will show in Section 4.4.

19.3.7 4.3.2 A global approach—recursive bisection

In order to obtain a balanced partitioning, graph partitioning algorithms commonly perform a recursive bisection, i.e. successively bisect the graph and the resulting parts until the desired number of parts is obtained [131, 132, 136, 137, 142, 143]. This approach has the advantage of reducing the partitioning problem to a simpler one, namely bisection. Furthermore, if the resulting parts of every bisection are equally sized, the overall partitioning will be balanced. In addition, bisection is inherently a global approach, as the whole graph must be considered for splitting the system into two equally sized parts. Thus, it can be expected to yield better results than a local approach, such as BFS.

We intend to construct a level set with $N + 2$ levels, where $N + 2$ is the maximum number of levels as determined by algorithm 4.1. To this end we start from an initial partitioning $\{\mathcal{V}_0, \mathcal{V}_1, \mathcal{V}_{N+1}\}$, where $\mathcal{V}_0$ and $\mathcal{V}_{N+1}$ contain the vertices of the leads (requirement 4.2.1), and $\mathcal{V}_1$ all other vertices. The level set is then obtained by applying the bisection algorithm recursively to $\mathcal{V}_1$ and the resulting subsets, until N levels are obtained, as shown schematically in Fig. 4.5. Here bisection means splitting a set $\mathcal{V}_i$ into two sets, $\mathcal{V}_{i_1}$ and $\mathcal{V}_{i_2}$, such that $\mathcal{V}_{i_1} \cup \mathcal{V}_{i_2} = \mathcal{V}_i$ and $\mathcal{V}_{i_1} \cap \mathcal{V}_{i_2} = \emptyset$. In order to be applicable to the graph partitioning problem 4.2, the bisection must comply with certain requirements:

Problem 4.5. The bisection algorithm must be

1. compatible with a level set with $N + 2$ levels.

2. balanced.

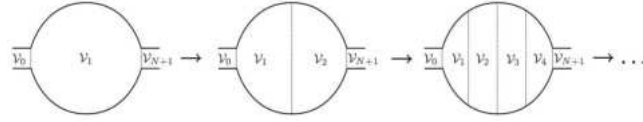


Figure 4.5 - Schematic depiction of recursive bisection.

3. performed such that subsequent bisections may lead to a balanced level set.

Requirement 4.5 .3 is formulated rather vaguely: Usually there are many different choices how to perform a bisection. A particular choice will influence the subsequent bisections (for a similar problem in graph partitioning see [143]), and thus the bisection algorithm must in principle take into account all following bisection steps. Since an exact solution to that problem seems computationally intractable, we will resort to heuristics there.

We start explaining how to comply with requirements 4.5.1 and 4.5.2. In the following we assume that $N > 0$, as $N = -1, 0$ are trivial cases. Then the initial partitioning $\{\mathcal{V}_0, \mathcal{V}_1, \mathcal{V}_{N+1}\}$ forms a level set, and so will the final result of the recursive bisection, if the result of every intermediate bisection yields a level set. For this, consider a set $\mathcal{V}_i$ with vertices adjacent to the sets $\mathcal{V}_{i_{\text{left}}}$ and $\mathcal{V}_{i_{\text{right}}}$, where "left" ("right") is defined as being closer to $\mathcal{V}_0$ ($\mathcal{V}_{N+1}$). Then the sets resulting from the bisection, $\mathcal{V}_{i_1}$ and $\mathcal{V}_{i_2}$ may only have vertices adjacent to $\mathcal{V}_{i_{\text{left}}}$, $\mathcal{V}_{i_2}$ and $\mathcal{V}_{i_1}$, $\mathcal{V}_{i_{\text{right}}}$, respectively.

Apart from the condition of forming a level set, requirement 4.5.1 also dictates the total number of levels. Due to the nature of the recursive bisection, the number of final levels contained in an intermediate step is always well-defined. If a set $\mathcal{V}_i$ contains N_i levels, then $\mathcal{V}_{i_1}$ and $\mathcal{V}_{i_2}$ must contain $N_{i_1} = \text{Int}(N_i/2)$ and $N_{i_2} = N_i - \text{Int}(N_i/2)$ levels, respectively. Here, $\text{Int}(\dots)$ denotes rounding off to the next smallest integer. The bisection is thus balanced, if

$$|\mathcal{V}_{i_1}| \approx \frac{N_{i_1}}{N_i} |\mathcal{V}_i| \quad \text{and} \quad |\mathcal{V}_{i_2}| \approx \frac{N_{i_2}}{N_i} |\mathcal{V}_i| \quad (4.9)$$

Note that N_i can take any value, and usually is not a power of two.

From Lemma 4.4 we know that the minimum set of vertices necessary to form n levels is given by a BFS up to level n . Let $\mathcal{V}_{i_1, \text{BFS}}$ ($\mathcal{V}_{i_2, \text{BFS}}$) denote the set of vertices found by a BFS starting from $\mathcal{V}_{i_{\text{left}}}$ ($\mathcal{V}_{i_{\text{right}}}$) up to level N_{i_1} (N_{i_2}). Then, for any bisection complying with requirement 4.5.1, $\mathcal{V}_{i_1, \text{BFS}} \subset \mathcal{V}_{i_1}$ and $\mathcal{V}_{i_2, \text{BFS}} \subset \mathcal{V}_{i_2}$. These vertices are uniquely assigned to $\mathcal{V}_{i_1}$ and $\mathcal{V}_{i_2}$ and are consequently marked as locked, i.e. later operations may not change this assignment. An example for the vertices found in a BFS is shown in Fig. 4.6(a). Note that in the initial bisection, $\mathcal{V}_i = \mathcal{V}_1$, $N_i = N$, $\mathcal{V}_{i_{\text{left}}} = \mathcal{V}_0$, and $\mathcal{V}_{i_{\text{right}}} = \mathcal{V}_{N+1}$.

The remaining unassigned vertices can be assigned to either set, and the bisection will still be compatible with a level set containing $N + 2$ vertices. Thus for complying with

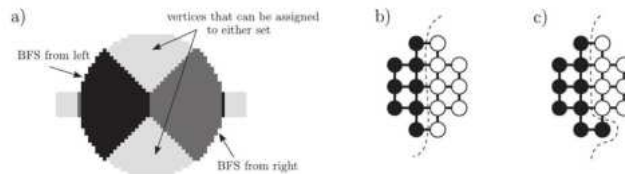


Figure 4.6 - (a) Example showing for a disk-type geometry the BFS from the left and right neighboring sets that construct the minimal set of vertices $\mathcal{V}_{i_1, \text{BFS}}$ (black) and $\mathcal{V}_{i_2, \text{BFS}}$ (dark grey) that must be contained in $\mathcal{V}_{i_1}$ and $\mathcal{V}_{i_2}$, respectively. The remaining vertices (light grey) can be assigned to either set. (b) and (c): Examples illustrating the difference between cut edges and cut nets. The number of cut edges is 5 in both (b) and (c), while the number of cut nets (surface vertices) is 10 in (b) and 9 in (c).

requirement 4.5.2, any prescription obeying the balancing criterion may be used. We choose to distribute the remaining vertices by continuing the BFS from $\mathcal{V}_{i_{\text{left}}}$ and $\mathcal{V}_{i_{\text{right}}}$ and assigning

vertices to $\mathcal{V}_{i_1}$ and $\mathcal{V}_{i_2}$ depending on their distance to the left or right neighboring set, while additionally obeying the balancing criterion. This approach assigning vertices to levels according to their distance from the initial and final set - is rather intuitive and probably the procedure that would be used if the level set were to be constructed "by hand". This procedure may lead to reasonable level sets, however in general, additional optimization on the sets $\mathcal{V}_{i_1}$ and $\mathcal{V}_{i_2}$ is needed, as discussed below. If this optimization is used, it can also be useful to distribute the unassigned vertices randomly, as this may help avoiding local minima.

As mentioned above, there is a lot of arbitrariness in distributing the unassigned vertices into $\mathcal{V}_{i_1}$ and $\mathcal{V}_{i_2}$. However, the particular choice of the bisection will influence whether a later bisection is balanced or not: If $\mathcal{V}_{i_1(i_2),\text{BFS}}$ contains more vertices than given by the balance criterion (4.9), the bisection cannot be balanced. Obviously, the BFS that constructs $\mathcal{V}_{i_1(i_2),\text{BFS}}$ depends on the details of the set $\mathcal{V}_i$ and thus on the details of the previous bisection step.

In order to formulate a criterion that may resolve the above-mentioned arbitrariness and help to find a balanced level set, it is useful to consider the matrix representation of the graph $\mathcal{G}$. Bisecting a graph means ordering the corresponding matrix into two blocks that are connected by an off-diagonal matrix H_{i_1,i_2} :

$$\left(\begin{array}{c|c} \ddots & \\ \hline & H_{i_1,i_2} \\ \hline H_{i_2,i_1} & \\ & \ddots \end{array} \right) \quad (4.10)$$

This off-diagonal matrix will be unchanged by further bisections and thus determines the minimum level width that can be achieved. Therefore, the size of the off-diagonal matrix H_{i_1,i_2} should be minimized.

In a bisection, an edge $(v_1, v_2) \in \mathcal{E}$ is said to be *cut*, if v_1 and v_2 belong to different sets, i.e. $v_1 \in \mathcal{V}_{i_1}$ and $v_2 \in \mathcal{V}_{i_2}$ or vice versa. The entries of H_{i_1,i_2} correspond to edges cut by the bisection, and minimizing the number of entries in H_{i_1,i_2} corresponds to minimizing the number of edges cut by the bisection (min-cut criterion). This criterion is often used in reordering matrices for parallel processing, where the offdiagonal matrix size determines the amount of communication between processors.

However, the number of entries in H_{i_1,i_2} is not directly related to the size of the matrix, as has been noted in the graph partitioning problem for parallel computing [138]. Instead, the size of the off-diagonal matrix is given by the number of surface vertices, i.e. the number of vertices that have cut edges. For this, we define a net of a vertex v in a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ as [136, 137]

$$\text{net}(v) = \{u \in \mathcal{V} \mid u \text{ is adjacent to } v\} \quad (4.11)$$

Note that $v \in \text{net}(v)$, as v is adjacent to itself. A net is said to be cut by a bisection, if any two vertices $v_1, v_2 \in \text{net}(v)$ are contained in different sets $\mathcal{V}_{i_1}$ and $\mathcal{V}_{i_2}$. Then, the number of surface vertices and thus the size of the off-diagonal matrix H_{i_1,i_2} is given by the number of cut nets. Thus, minimizing the number of cut nets (min-net-cut criterion) corresponds to minimizing the the number of surface vertices, and thus to minimizing the size of the off-diagonal matrix H_{i_1,i_2} . Furthermore, since the vertices in $\mathcal{V}_{i_{1/2},\text{BFS}}$ are determined by a BFS emanating from the surface vertices, minimizing the number of cut nets will usually also lead to a smaller number of vertices in $\mathcal{V}_{i_{1/2},\text{BFS}}$, leaving more freedom towards achieving a balanced bisection. Figs. 4.6(b) and (c) show a comparison of the min-cut and min-net-cut criterion for simple examples. In practice, when minimizing the number of cut nets, we also use the min-cut criterion to break ties between different bisections with the same number of cut nets

(min-net-cut-min-cut criterion) in order to avoid wide local minima, that occur frequently in the min-net-cut problem.

Both the min-cut and min-net-cut bisection problem have been shown to be $\mathcal{NP}$ hard [144]. Therefore, only heuristics are available to solve them. These heuristics start from an initial (balanced) bisection, such as constructed by the steps outlined above, and improve upon this initial bisection. Here, we choose to use the Fiduccia-Mattheyses (FM) algorithm [132], as it is readily available for min-cut and min-net-cut bisection¹. Furthermore, the FM algorithm can naturally deal with locked vertices that may not be moved between sets, is reasonably fast and its underlying concepts are easy to understand. The FM heuristic is a pass-based technique, i.e. it is applied repeatedly to the problem (several passes are performed), iteratively improving the bisection. More detailed information about the fundamentals of the Fiduccia-Mattheyses algorithm are given in appendix E.

We now summarize the steps outlined above and formulate an algorithm for bisection:

Algorithm 4.2. Bisection of set $\mathcal{V}_i$ containing N_i levels, with left (right) neighboring set $\mathcal{V}_{i_{\text{left}}}$ ($\mathcal{V}_{i_{\text{right}}}$).

A Stop, if $N_i = 1$.

B Do a BFS starting from $\mathcal{V}_{i_{\text{left}}}$ up to level $N_{i_1} = \text{Int}(N_i/2)$ and a BFS starting from $\mathcal{V}_{i_{\text{right}}}$ up to level $N_{i_2} = N - \text{Int}(N_i/2)$. The vertices found by the BFS are assigned to $\mathcal{V}_{i_1}$ and $\mathcal{V}_{i_2}$, respectively, and are marked as locked.

C Distribute the remaining unassigned vertices taking into account the balance criterion (4.9). The vertices may be assigned according to either one of the following prescriptions:

a) Continue the BFSs from step B and assign vertices to $\mathcal{V}_{i_1}$, if they are first reached by the BFS from $\mathcal{V}_{i_{\text{left}}}$, and to $\mathcal{V}_{i_2}$, if they are first reached by the BFS from $\mathcal{V}_{i_{\text{right}}}$. If a set has reached the size given by the balance criterion, assign all remaining vertices to the other set.

b) Distribute the unassigned vertices randomly to $\mathcal{V}_{i_1}$ and $\mathcal{V}_{i_2}$. If a set has reached the size given by the balance criterion, assign all remaining vertices to the other set.

D Optimize the sets $\mathcal{V}_{i_1}$ and $\mathcal{V}_{i_2}$ by changing the assignment of unlocked vertices according to some minimization criterion. In particular, the following optimizations may be performed:

a) No optimization.

b) Min-cut optimization using the FM algorithm.

c) Min-net-cut optimization using the FM algorithm.

d) Min-net-cut-min-cut optimization using the FM algorithm.

Recursive application of the bisection algorithm 4.2 then leads to an algorithm for constructing a level set complying with the requirements of the graph partitioning problem 4.2, and thus an algorithm for block-tridiagonalizing a matrix.

Algorithm 4.3. Block-tridiagonalization of matrix H

A Construct the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ corresponding to the matrix H , and the sets $\mathcal{V}_0$ and $\mathcal{V}_{N+1}$ corresponding to the leads.

B Use algorithm 4.1 to determine the maximum number of levels $N + 2$. If $N < 1$, stop.

C Construct $\mathcal{V}_1 = \mathcal{V} \setminus (\mathcal{V}_0 \cup \mathcal{V}_{N+1})$, containing N levels.

D Apply the bisection algorithm 4.2 to $\mathcal{V}_1$ and then recursively on the resulting subsets. Do not further apply if a set only contains one level.

It should be emphasized, that the block-tridiagonalization does not require any other input than the graph structure. In principle, the number of FM passes may affect the result. However, from experience, this number can be chosen as a fixed value, e.g. 10

¹²¹ In fact, min-net-cut bisection is a hypergraph partitioning problem, and the FM algorithm was originally designed for this task. For more information on hypergraphs, see Appendix E.

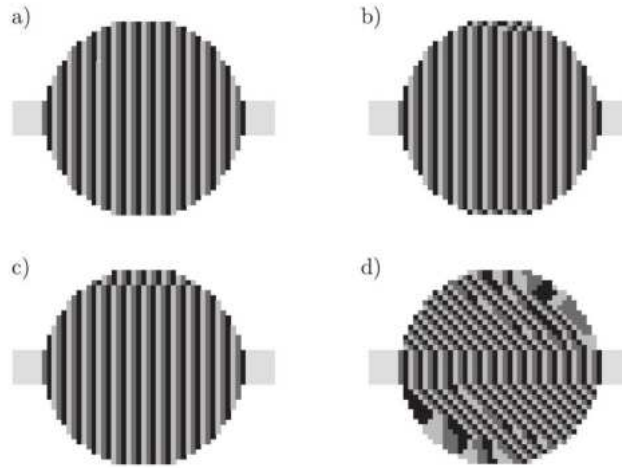


Figure 4.7 - Examples of level sets arising from (a) the natural ordering of grid points (as in Fig. 4.1), and application of the block-tridiagonalization algorithm 4.3 with distribution of vertices by BFS (algorithm 4.2, step C.(a)) (b) without further optimization, (c) with min-cut optimization, (d) with min-net-cut optimization.

FM passes, for all situations [132]. Thus, the block-tridiagonalization algorithm can serve as a black box.

In Fig. 4.7 we show examples of level sets arising from the natural ordering of grid points (Fig. 4.7(a), natural level set) and from the block-tridiagonalization algorithm developed in this work (Fig. 4.7(b)-(d)) for the case of a disk-type geometry. The level set in Fig. 4.7(b) arises from recursive bisection, where the vertices were distributed according to a BFS without any optimization. The resulting level set strongly resembles the natural level set. This is due to the highly symmetric structure and the fact that vertices are assigned to levels according to their distance from the leads—only small deviations are present due to the balance criterion. When the bisection is optimized according to the min-cut criterion, Fig. 4.7(c), the resulting level set changes little, as the min-cut criterion favors horizontal and vertical cuts for a square lattice, as presented in the example. In contrast, min-net-cut optimization (Fig. 4.7(d)) yields a new, nontrivial level set that has less symmetry than the underlying structure. Note that the minimization of surface vertices leads to levels in the form of "droplets", analogous to surface tension in liquids.

In fact, we will show in Section 4.4 that min-net-cut optimization usually leads to level sets and thus block-tridiagonal orderings that are superior to those arising from other methods. In particular, they are better than the natural level sets, leading to a significant speed-up of transport algorithms, as demonstrated in Section 4.4.1. In addition to that, the reordering algorithms allow one to use conventional two-terminal transport algorithms also for more complicated, even multi-terminal structures (see Sections 4.4.1 and 4.4.2).

19.3.8 4.3.3 Computational complexity

We conclude the theoretical considerations with an analysis of the computational complexity of algorithms 4.2 and 4.3.

The bisection algorithm involves a BFS search on $\mathcal{V}_i$, which scales linearly with the number of edges within $\mathcal{V}_i$, and thus has complexity $\mathcal{O}(|\mathcal{E}_i|)$, where $\mathcal{E}_i$ is the set of edges within $\mathcal{V}_i$. In addition to that, a single optimization pass of the FM algorithm scales also as $\mathcal{O}(|\mathcal{E}_i|)$ [132]. Usually, a constant number of passes independent of the size of the graph is enough to obtain converged results, and therefore the optimization process of several FM passes is also considered to scale as $\mathcal{O}(|\mathcal{E}_i|)$. Thus, the full bisection algorithm also has complexity $\mathcal{O}(|\mathcal{E}_i|)$.

Usually, the number of edges per vertex is approximately homogeneous throughout the

graph. Since the recursive bisection is a divide-and-conquer approach, the computational complexity of the full block-tridiagonalization algorithm is then $\mathcal{O}(|\mathcal{E}|\log|\mathcal{E}|)$ [141]. In typical graphs arising from physics problems, the number of edges per vertex is a constant, and the computational complexity can be written as $\mathcal{O}(N_{\text{grid}} \log N_{\text{grid}})$, where N_{grid} is the number of vertices in $\mathcal{V}$, or the size of the matrix H .

In contrast, many quantum transport algorithms, such as the recursive Green's function technique, scale as $\mathcal{O}(N(N_{\text{grid}}/N)^3) = \mathcal{O}(N_{\text{grid}}^3/N^2)$ in the optimal case of N equally sized matrix blocks (levels) of size N_{grid}/N . Often, the number of blocks (levels) $N \propto N_{\text{grid}}^\alpha$. Typically, to name a few examples, $\alpha = 1$ in one-dimensional chains, $\alpha = 1/2$ in two dimensions, and the transport calculation scales as $\mathcal{O}(N_{\text{grid}}^{3-2\alpha})$. Thus, except for the case of a linear chain, where $N = N_{\text{grid}}$ and matrix reordering is pointless anyway, the block-tridiagonalization algorithm always scales more favorably than the quantum transport algorithms. This scaling implies that the overhead of the matrix reordering in the transport calculation will become more negligible, the larger the system size.

19.3.9 4.4 Examples: Charge transport in two-dimensional systems

19.3.10 4.4.1 Ballistic transport in two-terminal devices

We now evaluate the performance of the block-tridiagonalization algorithm using representative examples from mesoscopic physics. The Schrödinger equation for the twodimensional electron gas (2DEG) is usually transformed into a tight-binding problem

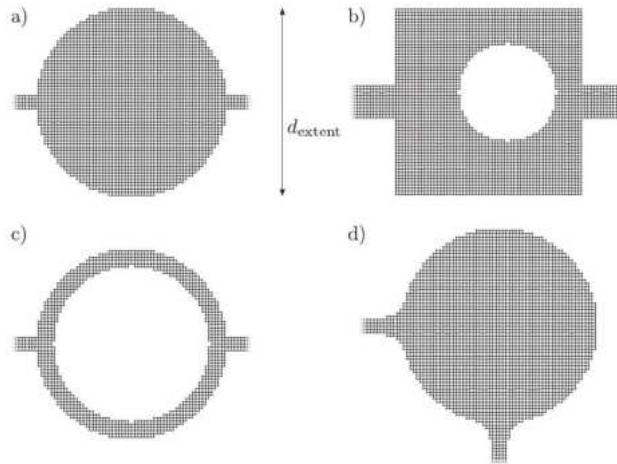


Figure 4.8 - Typical examples of structures considered in two-dimensional mesoscopic systems: (a) circle billiard, (b) asymmetric Sinai billiard, (c) ring, and (d) circular cavity with perpendicular leads. The tight-binding grid arises from the finite difference approximation to the Schrödinger equation. Note that the number of grid points used here was deliberately chosen very small for visualization purposes. In a real calculation, the number of grid points would be at least 2 orders of magnitude larger. d_{extents} denotes a length characterizing the extent for the different structures.

by the method of finite differences [77-79], where the continuous differential equation is replaced by a set of linear equations involving only the values of the wave function on discrete grid points, as discussed in Appendix F. Commonly, these points are arranged in a regular, square grid. This grid, together with the shape of the particular structure under consideration then defines the structure of the Hamilton matrix and the corresponding graph

The representative examples considered here are shown in Fig. 4.8: The circle (Fig. 4.8(a)) and the asymmetric Sinai billiard (Fig. 4.8(b)) that are examples of integrable and chaotic billiards in quantum chaos, respectively, the ring (Fig. 4.8(c)) that may exhibit various in-

terference physics, and the circular cavity with leads that are not parallel (Fig. 4.8(d)) as an example of a structure that does not have an intuitive, natural block-tridiagonal ordering. For all these structures, we introduce a length scale d_{extent} , given by the outer radius of the circular structures and the side length of the square structure, characterizing the maximum extent. The fineness of the grid, and thus the size of the corresponding graph will be measured in number of grid points per

Table 4.1 - Weights $w(H)$, Eq. (4.8), for the block-tridiagonal ordering constructed by different algorithms for the examples of Fig. 4.8. Optimization was done by 10 passes of the FM algorithm, when the initial bisection was constructed by BFS (algorithm 4.2, step C.(a)), and 20 passes, when the initial bisection was constructed by a random distribution of vertices (algorithm 4.2, step C.(b)). The minimal weights for each system are printed bold. In all examples, there were 400 grid points per length d_{extent} .

	Circular billiard	Asymmetric Sinai billiard	Ring	C
natural blocktridiagonal ordering	1.51×10^{10}	1.58×10^{10}	8.72×10^8	
Gibbs-Poole-Stockmeyer	1.15×10^{12}	7.84×10^{11}	2.14×10^8	
distribution by BFS, no optimization	1.51×10^{10}	9.29×10^9	2.1×10^8	
distribution by BFS, min-cut	1.51×10^{10}	9.67×10^9	2.1×10^8	
random distribution, min-cut	2.22×10^{10}	9.95×10^9	2.1×10^8	
distribution by BFS, min-net-cut	1.51×10^{10}	9.46×10^9	2.1×10^8	
random distribution, min-net-cut	1.46×10^{10}	9.0×10^9	2.09×10^8	
distribution by BFS, min-net-cut-min-cut	1.26×10^{10}	9.28×10^9	2.08×10^8	
random distribution, min-net-cut-min-cut	1.27×10^{10}	9.16×10^9	2.09×10^8	

length d_{extent} .

We now apply the block-tridiagonalization algorithm using the various optimization criteria discussed in the previous section, and compare the resulting orderings with the natural ordering and the ordering generated by the GPS algorithm. The weights $w(H)$, Eq. (4.8), of the different orderings are given in Table 4.1.

The initial distributions for the bisection algorithm are done in two different ways: The vertices are distributed both in an ordered way-by BFS-and randomly. The outcome after the optimization however is always similar for both types of initial distributions which indicates that the resulting weights are close to the global minimum and not stuck in a local minimum. Note that we use twice as many FM passes for a random initial distribution than for an initial distribution by BFS, as convergence is usually slower for a random initial distribution.

In all examples, the min-net-cut criterion yields orderings with the best weights, as expected from the considerations of the previous section. Based on the weight, orderings according to this criterion are expected to give the best performance in transport calculations such as the RGF algorithm. Note that the min-net-cut-min-cut ordering is on average closest to the best ordering. The min-net-cut ordering sometimes

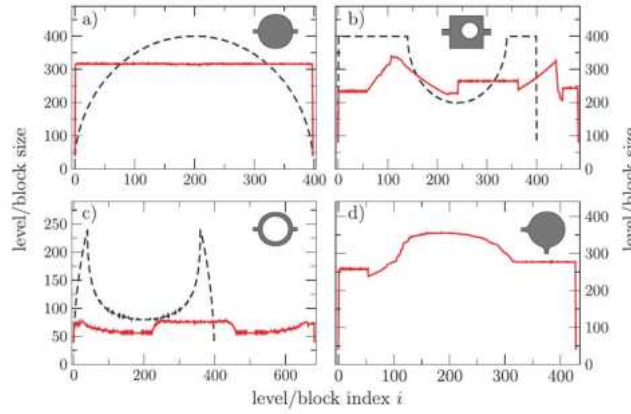


Figure 4.9 - Level (matrix block) size M_i as a function of the level (matrix block) index i for the natural level set (dashed line) and the min-net-cut-min-cut reordering (solid line), shown for (a) the circle billiard, (b) the asymmetric Sinai billiard, (c) the ring, and (d) the circular cavity with perpendicular leads. Note that for (d), there is no natural ordering. In all examples, there were 400 grid points per length d_{extent} .

suffers from slow convergence, when the algorithm must traverse a wide local minimum. The additional min-cut criterion helps to break ties and thus avoids these wide local minima.

Except for the ring, where all algorithms perform well, the GPS algorithm yields weights that are even larger than the weight of the natural ordering. As discussed above, the GPS algorithm performs well, if both leads are furthest apart in terms of the graph. In the case of the ring, this is approximately fulfilled. In the general case, when the leads are at arbitrary positions, the GPS algorithm usually produces some very large levels. As the level size enters cubically in the $w(H)$, this results in a prohibitively large weight. The GPS algorithm thus cannot be used as a generic reordering algorithm for quantum transport according to problem 4.2.

In summary, the block-tridiagonalization algorithm 4.3 in the combination of initial distribution by BFS and min-net-cut-min-cut optimization yields the best reorderings with respect to the weight $w(H)$. Experience shows that usually 10 FM passes are enough for optimizing a bisection. As a consequence, we will use this combination exclusively in the rest of this work.

The weight $w(H)$ of a matrix is a global measure of the quality of a ordering. Additional insight can be gained from the distribution of the sizes M_i of the matrix blocks/levels. In Fig. 4.9 we show this distribution before and after reordering. For the natural ordering of the finite difference grids, the number of matrix blocks is determined by the number of lattice points along the x -coordinate direction (see Fig. 4.1(b)). In contrast, the number of matrix blocks after reordering is given by the length of the shortest path between the two leads, in terms of the corresponding graph.

In the case of the circle billiard, Fig. 4.9(a), the number of matrix blocks is the same for the natural ordering and the reordered matrix, as the shortest path between the leads is simply a straight line along the x -coordinate direction. The improvements in the weight originate only from balancing the matrix block sizes: While the matrix block sizes vary for the natural ordering - the lateral size changes along the x -direction - the reordered matrix has equally sized matrix blocks. For this particular example, the result of the block-tridiagonalization algorithm is optimal, as it yields the best solution with respect to the requirements set forth in problems 4.1 and 4.2. Note that it is not always possible to find a perfectly balanced partitioning, but the circle billiard is such an example.

In contrast, in the case of the asymmetric Sinai billiard and the ring the number of matrix blocks generated by the block-tridiagonalization algorithm is larger than in the natural ordering (see Figs. 4.9(b) and (c), respectively). In both cases, the obstacle within the scattering region increases the length of the shortest path connecting the two leads. In both examples, this

increase in the number of matrix blocks leads to a significantly decreased weight $w(H)$ with respect to the natural ordering, although the partitioning is only approximately balanced. For instance, in the particular case of the ring, the number of matrix blocks after reordering is approximately given by the number of lattice points around half of the circumference. The reordered ring thus has a weight very similar to a straight wire with a width twice as large as the width of one arm of the ring, and a length given by half of the ring circumference.

For the cavity with perpendicular leads, there is no natural ordering, and a specialized transport algorithm would be required. The reordering creates a matrix with approximately balanced block sizes, and allows the direct application of conventional algorithms.

The weight $w(H)$ was introduced as a theoretical concept in order to simulate the computational complexity of a transport calculation. After discussing the influence of the reordering on this theoretical concept, we now demonstrate how the reordering increases the performance of an actual quantum transport calculation.

To this end we use a straight-forward implementation of the well-established recursive Green's function algorithm for two terminals, as described in Appendix B. It should be emphasized that the code that does the actual transport calculation - such as calculation of the Green's function and evaluation of the Fisher-Lee relation-is the same for all examples considered here, including the non-trivial cavity with perpendicular leads. The abstraction necessary for the reordering, i.e. the graph structure and the corresponding level set, allows for a generic computational code applicable to any tight-binding model.

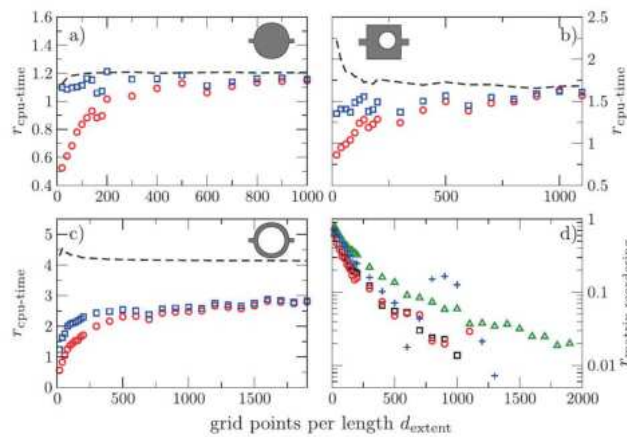


Figure 4.10 - (a)-(c): relative gain in computational time $r_{\text{cpu-time}}$, Eq. (4.12), through the reordering as a function of the grid size for the circular billiard, the asymmetric Sinai billiard, and the ring, respectively. $r_{\text{cpu-time}}$ is shown excluding ($\square$) and including ($\circ$) the overhead of matrix reordering. The estimate for $r_{\text{cpu-time}}$ from the weights $w(H)$ of the different orderings is shown as a dashed line. (d): fraction of time $r_{\text{matrix reordering}}$, Eq. (4.13), used for reordering the matrix as a function of the grid size. Data is shown for the circular billiard ($\square$), the asymmetric Sinai billiard ($\circ$), the ring ($\triangle$), and the circular cavity with perpendicular leads (+). The benchmarks were run on Pentium 4 processor with 2.8 GHz and 2 GBs of memory.

We measure the performance gain through matrix reordering as

$$r_{\text{cpu-time}} = \frac{\text{computing time for natural ordering}}{\text{computing time for reordered matrix}} \quad (4.12)$$

Note that during a real calculation, the conductance is usually not only calculated once, but repeatedly as a function of some parameters, such as Fermi energy or magnetic field. Thus, the respective quantum transport algorithm is executed repeatedly, too. In contrast, the block-tridiagonalization has to be carried out again only when the structure of the matrix and thus the corresponding graph changes. For the examples considered here this would correspond to changing the grid spacing or the shape of the structure. In such a case, the overhead of matrix

reordering must be taken into account for $r_{\text{cpu-time}}$. This overhead can be quantified as

$$r_{\text{matrix reordering}} = \frac{\text{overhead of matrix reordering}}{\text{computing time including reordering}} \quad (4.13)$$

In a typical calculation however, the matrix structure given by the underlying tightbinding grid does not change, and the matrix reordering must be carried out only once. In this common situation, the overhead of matrix reordering is negligible. For example, any change of physical parameters such as Fermi energy, magnetic field or disorder averages does not change the matrix structure.

In Fig. 4.10 we show the performance gain through matrix reordering, $r_{\text{cpu-time}}$, as a function of grid size for the circle billiard, the asymmetric Sinai billiard, and the ring (Figs. 4.10(a)-(c), respectively). We include both measurements excluding and including the overhead of matrix reordering, as discussed above. Remember that in the case of the cavity with perpendicular leads, Fig. 4.8(d), there is no natural ordering and thus a performance comparison is not possible. In fact for this system, only matrix reordering makes a transport calculation possible in the first place.

We find that block-tridiagonalization always increases the algorithmic performance in the typical situation, when the overhead of matrix reordering can be neglected. However, even if the reordering overhead is taken into account, we see a significant performance gain except for small systems-but there the total computing time is very short anyway. In fact, as the system sizes increases, the overhead of reordering becomes negligible, as predicted from the analysis of the computational complexity, and the performance gains including and excluding the reordering overhead converge. This can also be seen in Fig. 4.10(d), where we show the reordering overhead $r_{\text{matrix reordering}}$ as a function of system size.

Especially for large systems, the total computing time can become very long, and any performance gain is beneficial. Reordering leads to significant performance gains up to a factor of 3 in the case of the ring. The performance gain $r_{\text{cpu-time}}$ can also be estimated from the weights $w(H)$ of the original matrix (the natural ordering) and the reordered matrix, shown as the dashed line in Figs. 4.10(a)-(c). The actual, measured performance gain approaches this theoretical value, as the system size increases. Note that we do not fully reach the theoretically predicted performance gain in the case of the ring. On modern computer architectures, computing time does not only depend on the number of arithmetic operations [145], and thus the weight $w(H)$ overestimates the performance gain, though the performance still improves significantly.

Finally, we demonstrate the $\mathcal{O}(N_{\text{grid}} \log N_{\text{grid}})$ scaling of the reordering algorithm. Fig. 4.11 shows the computing times of the block-tridiagonalization algorithm as a function of matrix/graph size N for the geometries considered in this section. For all systems, the computing times scale according to the prediction from the complexity analysis in Section 4.3.3, as apparent from the fit $\propto N_{\text{grid}} \log N_{\text{grid}}$. Note that for large N_{grid} , $\mathcal{O}(N_{\text{grid}} \log N_{\text{grid}})$ scaling is practically indistinguishable from $\mathcal{O}(N_{\text{grid}})$ -scaling, as can also be seen in Fig. 4.11.

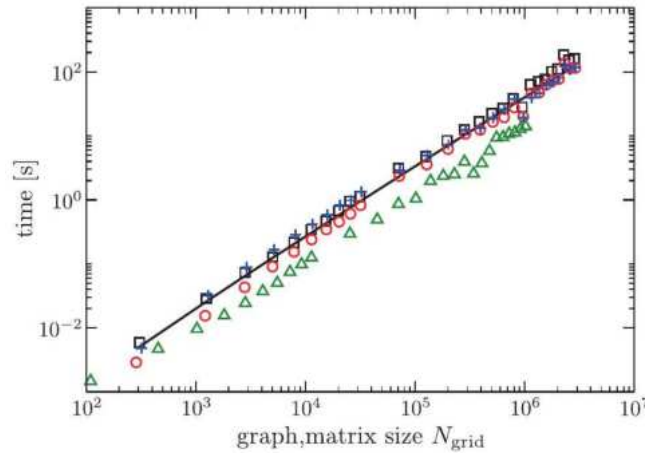


Figure 4.11 - Time spent for matrix reordering as a function of the total grid (matrix) size N , for the circular billiard ($\square$), the asymmetric Sinai billiard ($\circ$), the ring ($\triangle$), and the circular cavity with perpendicular leads ($+$). The solid line is a fit to the predicted scaling of the computational complexity, $N \log N$.

19.3.11 4.4.2 Multi-terminal structures

In the previous section, we demonstrated that matrix reordering increases the performance of quantum transport algorithms for two-terminal structures and additionally makes it possible to apply these conventional algorithms to non-trivial structures. Whereas there is a great variety of quantum transport algorithms for systems with two leads, there are only few algorithms that are suitable for multi-terminal structures, and most of them are restricted to rather specific geometries (e.g. Ref. [120]). Only recently algorithms have been developed that claim to be applicable to any multi-terminal structure. The knitting algorithm of Ref. [146] is a variant of the RGF algorithm where the system is built up adding every lattice point individually, instead of adding whole blocks of lattice points at a time. Therefore, instead of a matrix multiplication, the central computational step is an exterior product of vectors. Unfortunately, this implies that the knitting algorithm cannot use highly optimized matrix multiplication routines (Level 3 BLAS operations), that are usually much more efficient than their vector counterparts (Level 2 BLAS operations), as discussed in Ref. [145]. Another multi-terminal transport algorithm presented recently [147], is based on the transfer matrix approach. However, it requires the Hamiltonian to be in a specific block-tridiagonal form, and the corresponding level set is set up manually.

Here we show how to employ the block-tridiagonalization algorithm in order to apply the well-established two-terminal quantum transport algorithms to an arbitrary multi-

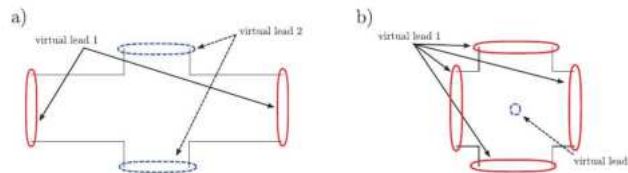


Figure 4.12 - A multi-terminal structure can be reduced to an equivalent two-terminal structure by collecting all leads in two virtual leads. (a) The leads are redistributed into two virtual leads. (b) All leads are combined in a single virtual lead, the second virtual lead is formed by a vertex furthest away.

terminal system. The basic idea is sketched in Fig. 4.12(a): Combining several (real) leads into only two virtual leads, the multi-terminal problem is reduced to an equivalent two-terminal problem. After reordering, the resulting problem can then be solved by conventional two-terminal algorithms. Note that in this approach the number of matrix blocks is given by the

shortest path between leads in two different virtual leads. If all leads are very close together, this may lead to only few, large blocks in the reordered matrix and respectively levels in the graph partitioning, leading to a very large weight $w(H)$. In such a case it is advisable to collect all leads into a single virtual lead. The second virtual lead is then formed by a vertex in the graph, that is furthest away from all leads as depicted in Fig. 4.12(b). Such a vertex can be found by a BFS search originating from all leads. Thereby the number of matrix blocks/levels is maximized. In fact, this approach yields a block-tridiagonal matrix structure as required by the algorithm of Ref. [147].

We now demonstrate these strategies on the example of the quantum Hall effect (QHE) in a 2DEG formed in a semiconductor heterostructure [148] and in graphene [32]. For this we use a four-terminal Hall bar geometry as sketched in Fig. 4.12(a), on top of a square lattice (finite difference approximation to 2DEG, see Appendix F) and a hexagonal lattice (graphene, see Appendix G). Again, it should be emphasized that the code of the actual transport calculation is the same as employed in the two-terminal examples of the previous section. The results of the calculation are shown in Fig. 4.13, where the integer QHE of the 2DEG and the odd-integer QHE of graphene are clearly visible.

The methods outlined above make it possible to calculate quantum transport in any system described by a tight-binding Hamiltonian. This generality is one of the main advantages gained by using the matrix reordering. However, generality also implies that it is difficult to make use of properties of specific systems, such as symmetries, in order to speed up calculations. Special algorithms developed specifically for a certain system however can, and will usually be faster than a generic approach-at the cost of additional development time.

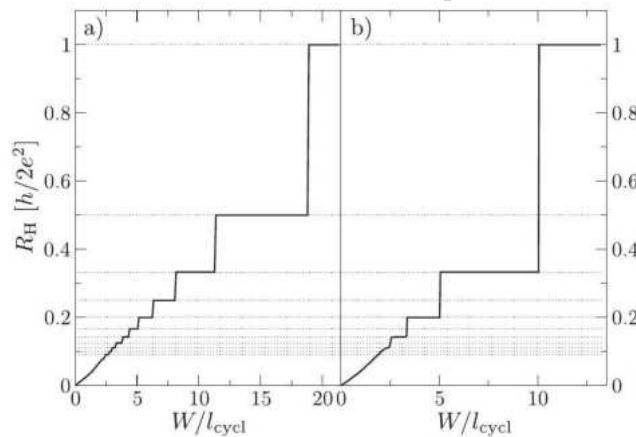


Figure 4.13 - Example of a four-terminal calculation: Quantum hall effect (a) in a two-dimensional electron gas and (b) in graphene. The Hall resistance R_H is shown as a function of W/l_{cycl} , where W is the width of the Hall bar and l_{cycl} the cyclotron radius in a magnetic field B . Note that $W/l_{cycl} \propto B$. The dotted lines indicate the quantized values of the Hall resistance, $h/2e^2 \times n^{-1}$, where n is a positive integer.

In the case of the Hall geometry in a 2DEG, such a special algorithm was presented by Baranger et al. [120], and we have implemented a variant of it. Comparing the computing times for the Hall bar geometry in a 2DEG, we find that the special algorithm is only a factor of 1.6 – 1.7 faster than our generic approach. Although such a performance comparison may depend crucially on the details of the system under consideration, experience shows that the use of the generic approach often does not come with a big performance penalty.

19.3.12 4.5 Summary

We have developed a block-tridiagonalization algorithm based on graph partitioning techniques that can serve as a preconditioning step for a wide class of quantum transport algorithms. The algorithm can be applied to any Hamiltonian matrix originating from an arbitrary

tight-binding formulation and brings this matrix into a form that is more suitable for many two-terminal quantum transport algorithms, such as the widely used recursive Green's function algorithm. The advantages of this reordering are twofold: First, the reordering can speed up the transport calculation significantly. Second, it allows for applying conventional two-terminal algorithms to non-trivial geometries including non-collinear leads and multi-terminal systems. The

block-tridiagonalization algorithm scales as $\mathcal{O}(N_{\text{grid}} \log N_{\text{grid}})$, where N_{grid} is the size of the Hamiltonian matrix, and thus induces only little additional overhead. We have demonstrated the performance of the matrix reordering on representative examples, including transport in 2DEGs and graphene.

The framework set by the non-equilibrium Green's function formalism in tight-binding representation, presented in Chapter 2, provides a clear pathway for numerical transport calculations. Together with the novel results of the last two chapters, namely the general expression for the surface Green's function of the leads and the block-tridiagonalization algorithm, it is then possible to develop a transport code applicable to arbitrary tight-binding systems. This is possible, since both the calculation of the surface Green's function and the block-tridiagonalization algorithm can operate as a black box, i.e. without the need for specifying external parameters other than the tight-binding Hamiltonian H . Such a generic transport code is desirable, as it minimizes development time and increases code quality, as only a few basic, thoroughly tested transport routines are necessary.

The algorithms developed in this work have been implemented in a computer program, as a set of transport routines requiring only the input of the tight-binding Hamiltonian H . Adapting this code to an arbitrary tight-binding models thus only consists of appropriately setting up the matrix H . This has been done for the cases of a two and three-dimensional effective mass model as well as for graphene, including the spin degree of freedom. In the remainder of this thesis we will employ this code to compute spin transport properties in magnetic tunnel junctions and graphene nanoribbons.

19.4 Chapter 5 Magnetic field effects on tunneling anisotropic magnetoresistance

19.4.1 5.1 Introduction

Amongst the various magnetoresistive phenomena, the giant magnetoresistance (GMR) [14, 15] and the tunneling magnetoresistance (TMR) effect [16-18] are probably the most prominent examples. In both cases, the resistance of two ferromagnetic layers separated by a spacer layer was found to depend on the relative angle of the magnetizations in the two ferromagnets. While this spacer is a metal in the GMR effect, it is an insulating material for the TMR effect. Although the basic experimental geometries are very similar in both cases, the underlying physics is quite different: diffusive transport in the case of GMR, and quantum mechanical tunneling in the case of TMR. Both the GMR and the TMR effect are employed in various electronic devices, such as read heads of hard disks or magnetic random access memory (MRAM) (for reviews see Refs. [12, 13, 19]).

In recent years, several experiments have found that the resistance of a magnetic tunnel junction (MTJ) may also depend on the absolute angle of the magnetization in a ferromagnet with respect to some crystallographic axis, in contrast to the conventional GMR and TMR effects, where the resistance depends only on the relative angle between the magnetizations in two ferromagnetic layers. Amongst those, the most striking are experiments on MTJs with only a single ferromagnetic layer that nevertheless show a magnetoresistance effect, depending on the orientation of the magnetization in the ferromagnet [47, 149]. This magnetoresistance

effect is called tunneling anisotropic magnetoresistance (TAMR) effect.

The TAMR effect was first observed in a (Ga,Mn)As/ Al_2O_3 /Au tunnel junction, featuring epitaxially grown magnetic semiconductor (Ga,Mn) As as the ferromagnetic elec-

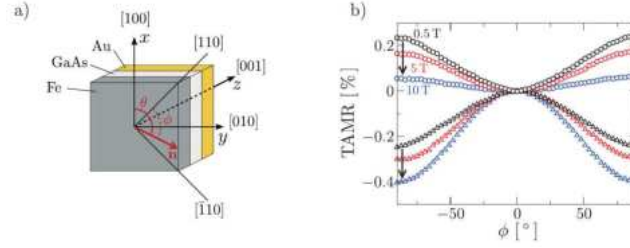


Figure 5.1 - (a) Schematic view of the magnetic tunnel junction in the experiments of Refs. [47, 150]. (b) TAMR effect for different bias voltages ($V_{\text{bias}} = +90\text{meV} : \bigcirc$; $V_{\text{bias}} = -90\text{meV} : \triangle$) and magnetic fields (black: $B = 0.5\text{ T}$; red: $B = 5\text{ T}$; blue: $B = 10\text{ T}$) in a Fe/GaAs/Au tunnel junction (from [150]).

trode, and an amorphous aluminum-oxide tunnel barrier [149]. The observed TAMR exhibited a four-fold anisotropy and was explained by an anisotropic density of states due to uniaxial strain in the magnetic semiconductor.

A qualitatively different TAMR effect was found in epitaxially grown Fe/GaAs/Au tunnel junctions with a single-crystalline tunnel barrier [47, 150], showing a uniaxial anisotropy in the magnetoresistance. The GaAs barrier was grown in the [001]direction, as indicated in the schematic of the experiment in Fig. 5.1(a). The angular dependence of the TAMR effect can be quantified as

$$\text{TAMR}(\phi) = \frac{R(\phi) - R(0)}{R(0)} \quad (5.1)$$

where $R(\phi)$ is the resistance of the tunnel junction, when the magnetization in Fe forms an angle ϕ with the [110]-direction, as indicated in Fig. 5.1(a). In the experiment, the TAMR was found to be periodic in ϕ , with a periodicity of π , as shown in Fig. 5.1(b). The TAMR effect shows both a peculiar bias as well as magnetic field dependence: For different bias voltages V_{bias} the TAMR changes in magnitude, and may also change sign. Moreover, the magnitude of the effect decreases with increasing magnetic field B when the TAMR effect is positive, and it increases when the TAMR effect is negative (magnetic field increases from the black to the blue symbols in Fig. 5.1(b)).

Theoretically, the TAMR effect for the case of small magnetic fields can be explained by the effects of spin-orbit interaction in the tunneling process [47, 151, 152]. This theory assumes that the momentum parallel to the layers (in-plane momentum) is conserved during the tunneling process, as appropriate for epitaxially grown (in contrast to amorphous) systems. The interference of Bychkov-Rashba spin-orbit coupling at the Fe/GaAs interface and Dresselhaus spin-orbit coupling within the GaAs-barrier then leads to the observed uniaxial anisotropy of the magnetoresistance. The bias dependence of the TAMR is explained by a bias-dependent value of the interfacial Bychkov-Rashba spin-orbit parameter. However, up to now, a theoretical explanation of the magnetic field dependence of the TAMR is still missing. Such an explanation is the objective of this chapter.

To this end, the chapter is organized as follows: First, we extend the theoretical model of Refs. [47, 151, 152] to include the orbital effects of a magnetic field B . The resulting model is then solved numerically and compared to experimental data. Finally, we develop a qualitative model for the B -field dependence of the TAMR, reproducing all characteristic features of the experiment.

19.4.2 5.2 Magnetic field dependence of TAMR

19.4.3 5.2.1 Model

Within the theoretical model of the TAMR effect, the Hamiltonian reads [47, 151, 152]

$$H = H_0 + H_D + H_{BR} \quad (5.2)$$

where

$$H_0 = \frac{1}{2} \mathbf{p} \frac{1}{m(z)} \mathbf{p} + V(z) - \frac{\Delta(z)}{2} \mathbf{n} \cdot \boldsymbol{\sigma} \quad (5.3)$$

Here, $\mathbf{p} = -i\hbar\nabla$, and the coordinate system was chosen such that the z -direction is in the growth direction of the MTJ, as indicated in Fig. 5.1(a). $m(z)$ is the effective mass, with $m(z) = 0.067m_e$ inside the GaAs barrier, and $m(z) \approx m_e$ within the Fe and Au layer, where m_e is the bare electron mass. $V(z)$ is the conduction band profile and $\Delta(z)$ the spin splitting. We neglect the Zeeman splitting in the barrier and the non-magnetic electrode, as these are much smaller than any other energy scale in the system, and only keep the spin splitting due to the magnetization in the ferromagnet. The direction of this magnetization is given by

$$\mathbf{n} = \begin{pmatrix} \cos \theta \\ \sin \theta \\ 0 \end{pmatrix} \quad (5.4)$$

where θ is the angle with respect to the x -axis. The experimentally used angle ϕ and θ are related through

$$\phi = \theta + \frac{\pi}{4} \quad (5.5)$$

as seen from the sketch in Fig. 5.1(a). $\boldsymbol{\sigma}$ is the vector of Pauli matrices. The shape of the conduction band offset resulting from this model is shown schematically in Fig. 5.2(a). The barrier height is approximately given by half the GaAs band gap, so that $V_{\text{bar}} \approx 0.75 \text{ eV}$. In addition, the Fermi wave vectors within *Fe* are $k_{\text{F,Fe}}^{\uparrow} = 1.05 \times 10^{10} \text{ m}^{-1}$ and

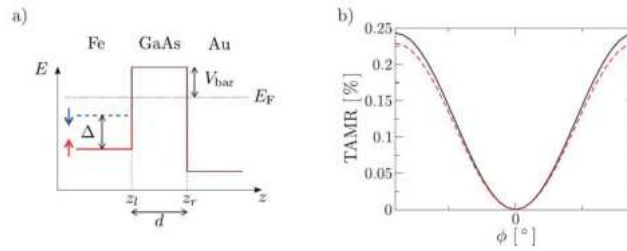


Figure 5.2 - (a) Schematic depiction of the conduction band profile for spin up (red solid line) and spin down (blue dashed line) for the theoretical model of the Fe/GaAs/Au tunnel junction. (b) Comparison between the TAMR calculated numerically from the finite difference approximation with $a_0 = 0.01 \text{ nm}$ (black solid line) and from an approximate analytical expression [152] (red dashed line). The Bychkov-Rashba parameter was chosen as $\alpha = 25 \text{ eV \AA}^2$. $k_{\text{F,Fe}}^{\downarrow} = 0.44 \times 10^{10} \text{ m}^{-1}$ for spin up and down, respectively. Inside the Au layer, the Fermi wave vector is given as $k_{\text{F,Au}} = 1.2 \times 10^{10} \text{ m}^{-1}$. H_D and H_{BR} describe the Dresselhaus and Bychkov-Rashba spin-orbit coupling, respectively. Here

$$H_D = \frac{1}{\hbar} (\sigma^x p_x - \sigma^y p_y) \frac{\partial}{\partial z} \left(\gamma(z) \frac{\partial}{\partial z} \right) \quad (5.6)$$

where the Dresselhaus parameter is $\gamma \approx 24\text{eV}\text{\AA}^3$ within the GaAs barrier and $\gamma = 0$ outside. The Bychkov-Rashba spin-orbit coupling is only relevant at the Fe/GaAs interface (which exhibits a broken inversion symmetry) and the Hamiltonian is given as

$$H_{\text{BR}} = \frac{1}{\hbar} \alpha (\sigma^x p_y - \sigma^y p_x) \delta(z - z_1), \quad (5.7)$$

where z_1 is the position of the left interface. In contrast to the Dresselhaus parameter γ which is a bulk property, the interface Bychkov-Rashba parameter is not known. Instead, α is used as a fitting parameter to describe the TAMR at a given bias voltage V_{bias} , hence resulting in a bias-dependent Bychkov-Rashba parameter $\alpha(V_{\text{bias}})$. A summary of the material parameters used in the model is given in Table 5.1. Furthermore, in the remainder of this chapter we fix the barrier width $d = 8$ nm, as in the experiments.

In order to include the orbital effects of a magnetic field $\mathbf{B}$ into this model, we first

m_{GaAs}	m_{Fe}	m_{Au}	V_{bar}	$k_{\text{F,Fe}}^\dagger$	$k_{\text{F,Fe}}$	$k_{\text{F,Au}}$	γ_{GaAs}
$0.067m_e$	m_e	m_e	0.75 eV	$1.05 \times 10^{10} \text{ m}^{-1}$	$0.44 \times 10^{10} \text{ m}^{-1}$	$1.2 \times 10^{10} \text{ m}^{-1}$	$24\text{eV}\text{\AA}^3$

Table 5.1 - Summary of the material parameters of the theoretical model for the Fe/GaAs/Au tunnel junction (from Refs. [13, 47, 151]).
note that the magnetic field is antiparallel to $\mathbf{n}$, so that ¹

$$\mathbf{B} = \begin{pmatrix} -B \cos \theta \\ -B \sin \theta \\ 0 \end{pmatrix} \quad (5.8)$$

We then choose the vector potential such that

$$\mathbf{A}(z) = \begin{pmatrix} -B \sin \theta z \\ B \cos \theta z \\ 0 \end{pmatrix} \quad (5.9)$$

where $\mathbf{B} = \nabla \times \mathbf{A}$. The magnetic field then enters the Hamiltonian via minimal coupling

$$\mathbf{p} \rightarrow \mathbf{p} + e\mathbf{A}(z) \quad (5.10)$$

This substitution must be done for any momentum operator in the Hamiltonian, including the spin-orbit terms.

In the chosen gauge, the vector potential in Eq. (5.9) only depends on the longitudinal coordinate z . Hence, the Hamiltonian is still translationally invariant in x and y direction, and Bloch's theorem applies. Therefore, the full wave function can be written as

$$\varphi(\mathbf{x}) = e^{i(k_x x + k_y y)} \varphi(z) \quad (5.11)$$

and the Schrödinger equation for the Hamiltonian H reduces to a one-dimensional differential equation for $\varphi(z)$. The transverse wave vector $\mathbf{k}_\parallel = (k_x, k_y)^\dagger$ is a good quantum number, as in the case without magnetic field.

At this point it should be emphasized that the conservation of $\mathbf{k}_\parallel$ during the tunneling process is not in contradiction to the cyclotron effects expected from a magnetic field. In fact, physically observable quantities, such as the electron velocity $\mathbf{v}_\parallel = \frac{1}{m} (\hbar \mathbf{k}_\parallel + e\mathbf{A}_\parallel(z))$ do show the transverse acceleration expected from the Lorentz force ².

¹²¹ Magnetic moments $\boldsymbol{\mu}$ align parallel to the magnetic field $\mathbf{B}$, but the magnetic moment of an electron is antiparallel to its spin: $\boldsymbol{\mu}_e = -\frac{e}{2m_e} g \frac{\hbar}{2} \boldsymbol{\sigma}$, where $g \approx 2$ is the gyromagnetic factor of the electron.

² For further discussion on the difference between canonical and mechanical momentum, see, for example, Ref. [153, App. III]

The choice of gauge in Eq. (5.9) also comes with a disadvantage: It is not possible to include a magnetic field in the contacts ³. In practice, we therefore switch off the magnetic field after some distance away from the barrier. This is a good approximation, as the effects of the magnetic field are expected to be strongest in the barrier and negligible in the metallic contacts: The cyclotron radius $l_{\text{cycl}} = p_F/eB$ [111] takes the values $l_{\text{cycl}} \approx 700$ nm in Fe and $l_{\text{cycl}} \approx 800$ nm in Au, much larger than any relevant length scale in the problem. Indeed, we find that the numerical results are insensitive to the exact position where the magnetic field is switched off.

In order to solve the Hamiltonian (5.2) including a magnetic field numerically, we employ the finite difference approximation that yields a tight-binding description of the system (for details, see Appendix F). The Delta function in Eq. (5.7) is approximated by a step function of width $2a_0$, centered at the Fe/GaAs interface. Here, a_0 is the lattice spacing of the discretization. In order to compare our numerical results with the previous analytical results of Refs. [47, 151, 152], we choose a very fine lattice spacing $a_0 = 0.01$ nm, to obtain a good approximation for the Delta function in Eq. (5.7). ⁴

For a given $\mathbf{k}_{\parallel}$, the transmission probability $T(E, \mathbf{k}_{\parallel})$ at energy E can then be calculated using the numerical techniques developed in Part I. Then, the total conductance is given as

$$G(\phi) = \frac{S}{(2\pi)^2} \int d\mathbf{k}_{\parallel} T(E, \mathbf{k}_{\parallel}) \quad (5.12)$$

where S is the cross-section area of the magnetic tunnel junction. Note that the finite difference approximation as presented in Appendix F is not restricted to situations where $\mathbf{k}_{\parallel}$ is conserved, but can be also applied to disordered systems [155]. In terms of conductances, the TAMR value is given as

$$\text{TAMR}(\phi) = \frac{R(\phi) - R(0)}{R(0)} = \frac{G(0) - G(\phi)}{G(\phi)} \quad (5.13)$$

In Fig. 5.2(b) we show a comparison of the TAMR effect without magnetic field calculated both numerically and from an approximate analytical expression [152]. Given the fact that the analytical expression is only approximate, we find a very good agreement between the numerical and analytical results ⁵ confirming the validity of the numerical procedure.

19.4.4 5.2.2 Numerical results

We now compare the results of numerical simulations on the magnetic field dependence of the TAMR effect with the corresponding experimental data. For this, we fit the parameter α at $B = 0.5$ T for every value of the bias voltage V_{bias} to the experimental data. In doing so, we obtain the bias voltage dependence of $\alpha = \alpha(V_{\text{bias}})$, as in Refs. [47, 151]. The magnetic field dependence can then be calculated without fitting any further parameter.

In Fig. 5.3(a) we show the angular dependence of the TAMR effect for different values of the bias voltage and magnetic field. The numerical simulations show the same trend as the experiment: The TAMR effect decreases with increasing B , when the effect is positive, and it

¹²³ First of all, scattering states can only be defined if the vector potential is parallel to the direction of the lead [154]. Second, also the method to calculate the surface Green's function presented in Chapter 3 is based on translational invariance along the lead direction.

⁴ For the description of the experiment it is not necessary to choose such a fine lattice spacing. In fact, no length scales smaller than the inter-atomic distance should play a role. However, the value of the Bychkov-Rashba parameter α depends on the width $2a_0$ of the step function mimicking the Delta function of Eq. (5.7); hence a fine lattice spacing is needed to compare with the previous analytical results.

⁵ When comparing results from this thesis and previous work, note that Refs. [47, 151] used a different sign convention for α .

increases, when the effect is negative. Furthermore, the numerical calculations reproduce the experimentally found change with magnetic field within a factor of $1.5 - 2$. This is an especially satisfying agreement, given the fact that the magnetic field dependence is calculated without fitting any parameter.

In order to quantify the magnetic field dependence of the magnitude of the TAMR effect, we show TAMR (90°) as a function of the magnetic field B in Fig. 5.3(b). In the experiment, we find that TAMR (90°) changes linearly with B . Moreover, the slope of this change is approximately independent of the bias voltage V_{bias} . Both of these characteristic features are also found in the numerical simulations. As before, the numerical results underestimate the slope of the magnetic field dependence only by a factor of $1.5 - 2$.

Up to now, detailed experimental data on the B -field dependence of TAMR (90°) is only available up to 5 T. From the magnetic field dependence of single TAMR curves as shown in Fig. 5.3(a), it seems that TAMR (90°) may change faster than linear in B for $B > 5$ T, although this is hard to decide from only a few points. In contrast, the numerical simulations result in a linear B -dependence also for fields up to 10 T. For large fields, other terms than those considered in Eq. (5.2) may become important [13]. In any case, the discussion of this point should be postponed until further experimental data is available.

The numerical results of this section show that the orbital effects of a magnetic field are enough to reproduce the characteristic features of the magnetic field dependence in the experiment. In order to understand the underlying physical mechanisms of this effect, we develop a qualitative model of the B -field dependence of the TAMR effect in the next section.

19.4.5 5.2.3 A qualitative picture

In order to develop a qualitative model for the magnetic field dependence of the TAMR, we extend the phenomenological model of Refs. [13, 151] to include finite magnetic fields.

The transmission probability can be written as a perturbative expansion in the spin-

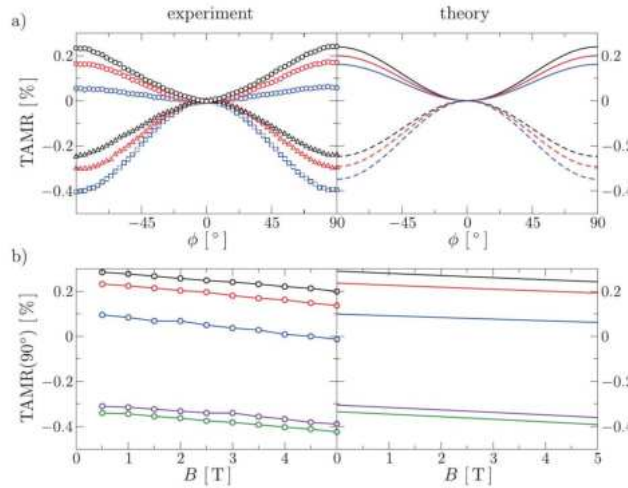


Figure 5.3 - (a) Angular dependence of the TAMR effect as found in the experiment (left panel, from [150]) and from numerical simulations (right panel). Data is shown for $V_{\text{bias}} = +90\text{meV}$ (experiment: $\circ$; theory: $\alpha = 25.1\text{eV}\text{\AA}^2$, solid lines) and $V_{\text{bias}} = -90\text{meV}$ (experiment: $\triangle$; theory: $\alpha = -25\text{eV}\text{\AA}^2$, dashed lines), as well as different magnetic fields: $B = 0.5$ T (black), $B = 5$ T (red), and $B = 10$ T (blue). (b) TAMR (90°) as a function of magnetic field in the experiment (left panel, from [48]) and from numerical simulations (right panel). Data is shown for $V_{\text{bias}} = +135\text{meV}$ ($\alpha = 29.3\text{eV}\text{\AA}^2$, black), $V_{\text{bias}} = +90\text{meV}$ ($\alpha = 24.3\text{eV}\text{\AA}^2$, red), $V_{\text{bias}} = +50\text{meV}$ ($\alpha = 10.5\text{eV}\text{\AA}^2$, blue), $V_{\text{bias}} = -50\text{meV}$ ($\alpha = -33.4\text{eV}\text{\AA}^2$, green), and $V_{\text{bias}} = -90\text{meV}$ ($\alpha = -30.8\text{eV}\text{\AA}^2$, violet).

orbit coupling [13, 151],

$$T(E, \mathbf{k}_{\parallel}) = T^{(0)}(E, k_{\parallel}) + T^{(1)}(E, k_{\parallel}) \mathbf{n} \cdot \mathbf{w}(\mathbf{k}_{\parallel}) + T^{(2)}(E, k_{\parallel}) (\mathbf{n} \cdot \mathbf{w}(\mathbf{k}_{\parallel}))^2 + \dots \quad (5.14)$$

where

$$\mathbf{w}(\mathbf{k}_{\parallel}) = \begin{pmatrix} \tilde{\alpha}k_y + \tilde{\gamma}k_x \\ -\tilde{\alpha}k_x - \tilde{\gamma}k_y \\ 0 \end{pmatrix} \quad (5.15)$$

is an effective spin-orbit field, obtained by averaging the spin-orbit field $\mathbf{B}_{\text{eff}}(z)$, $H_{\text{SO}} = H_D + H_{\text{BR}} = \mathbf{B}_{\text{eff}}(z) \cdot \boldsymbol{\sigma}$ over the unperturbed states of the system, i.e. in the absence of spin-orbit coupling. The effective Bychkov-Rashba and Dresselhaus parameters $\tilde{\alpha} = \alpha f_{\alpha}(k_{\parallel})$ and $\tilde{\gamma} = \gamma f_{\gamma}(k_{\parallel})$ are functions of $k_{\parallel} = |\mathbf{k}_{\parallel}|$ only, as are the expansion coefficients $T^{(n)}(E, k_{\parallel})$. Note that the $T^{(n)}(E, k_{\parallel})$ cannot be interpreted as transmission probabilities individually—for example, being expansion coefficients, $T^{(n)}(E, k_{\parallel})$ may very well be negative.

As before, the total transmission is obtained by integrating over all transverse wave vectors,

$$T = T^{(0)} + T^{(1)} + T^{(2)} + \dots \quad (5.16)$$

where $T^{(n)} = \langle T^{(n)}(E, k_{\parallel}) (\mathbf{n} \cdot \mathbf{w}(\mathbf{k}_{\parallel}))^n \rangle_{\mathbf{k}_{\parallel}}$ and $\langle \dots \rangle_{\mathbf{k}_{\parallel}} = \frac{1}{(2\pi)^2} \int d\mathbf{k}_{\parallel} \dots$. In principle, the integration over transverse wave vectors must be restricted to values within the Fermi sphere. However, in a tunnel junction the transmission probability is suppressed exponentially for increasing $k_{\parallel}$ (transmission is highest for perpendicular incidence, i.e. $\mathbf{k}_{\parallel} = 0$), and hence we can extend the $\mathbf{k}_{\parallel}$ -integration to infinity. To second order in the spin-orbit coupling, the conductance of the magnetic tunnel junction is then found as

$$G(\phi) = G^{(0)} + a_0^{(2)} + a_1^{(2)} \alpha \gamma \cos(2\phi), \quad (5.17)$$

where $G^{(0)}$, $a_0^{(2)}$, and $a_1^{(2)}$ are coefficients obtained by taking the respective $\mathbf{k}_{\parallel}$ -integrations for the zeroth- and second-order term in Eq. (5.14) (for details, see Ref. [13]). In particular, the coefficient $a_1^{(2)}$ does not depend on the spin-orbit parameters α and γ . The linear term in Eq. (5.14) vanishes exactly upon integration, as $\mathbf{w}(-\mathbf{k}_{\parallel}) = -\mathbf{w}(\mathbf{k}_{\parallel})$.

After reviewing the phenomenological model for $B = 0$, we now turn to the case of finite magnetic field B . In this situation, we can still expand the transmission probability with respect to the spin-orbit coupling, albeit with different expansion coefficients and a different effective spin-orbit field:

$$T_B(E, \mathbf{k}_{\parallel}) = T_B^{(0)}(E, \mathbf{k}_{\parallel}) + T_B^{(1)}(E, \mathbf{k}_{\parallel}) \mathbf{n} \cdot \mathbf{w}_B(\mathbf{k}_{\parallel}) + T_B^{(2)}(E, \mathbf{k}_{\parallel}) (\mathbf{n} \cdot \mathbf{w}_B(\mathbf{k}_{\parallel}))^2 + \dots \quad (5.18)$$

In the following we will derive approximate relations valid to linear order in B between the expansion coefficients $T_B^{(n)}$ and the effective spin-orbit field $\mathbf{w}_B$ to their counterparts for $B = 0$.

First, we consider the orbital effects of the magnetic field in the kinetic energy term of the Hamiltonian, and neglect the spin-orbit coupling. The kinetic part of the

Hamiltonian including a magnetic field reads

$$H_{\text{kin}} = \frac{\hbar^2}{2m} \left(\left(k_x - \frac{eB}{\hbar} \sin \theta z \right)^2 + \left(k_y + \frac{eB}{\hbar} \cos \theta z \right)^2 - \frac{\partial^2}{\partial z^2} \right) \quad (5.19)$$

where for the sake of argument we neglect the spatial dependence of the effective mass m . The effect of the magnetic field is two-fold: First, the maximum of the transmission probability is shifted with respect to $\mathbf{k}_{\parallel} = 0$. Instead, the maximum of $T_B(E, \mathbf{k}_{\parallel})$ is expected at $\mathbf{k}_{\parallel,0}$ with

$$\left\langle \left(k_{x,0} - \frac{eB}{\hbar} \sin \theta z \right)^2 \right\rangle = 0 \quad \text{and} \quad \left\langle \left(k_{y,0} + \frac{eB}{\hbar} \cos \theta z \right)^2 \right\rangle = 0 \quad (5.20)$$

where $\langle \dots \rangle$ denotes an average over a quantum mechanical state. Hence, $\mathbf{k}_{\parallel,0}$ can be written as

$$k_{x,0} = +b_1 B \sin \theta \quad (5.21a)$$

$$k_{y,0} = -b_1 B \cos \theta \quad (5.21b)$$

where b_1 is a coefficient that depends on both $\langle z \rangle$ and $\langle z^2 \rangle$. This shift can be regarded as the effect of the Lorentz force on the electrons. Second, the overall transmission probability decreases [156]. However, this effect is quadratic in B and will thus be neglected in the following. These considerations motivate the approximation

$$T_B^{(n)}(E, \mathbf{k}_{\parallel}) \approx T^{(n)}\left(E, \sqrt{(k_x - b_1 B \sin \theta)^2 + (k_y + b_1 B \cos \theta)^2}\right). \quad (5.22)$$

In addition to the transmission probability, also the effective spin-orbit field is shifted by the orbital terms from the magnetic field in Eqs. (5.6) and (5.7). The spin-orbit terms including a magnetic field read

$$H_{\text{SO}} = \alpha \left(\sigma_x \left(k_y + \frac{eB}{\hbar} \cos \theta z \right) - \sigma_y \left(k_x - \frac{eB}{\hbar} \sin \theta z \right) \right) \delta(z - z_l) + \\ \gamma \left(\sigma_x \left(k_x - \frac{eB}{\hbar} \sin \theta z \right) - \sigma_y \left(k_y + \frac{eB}{\hbar} \cos \theta z \right) \right) \frac{\partial}{\partial z} \left(\gamma(z) \frac{\partial}{\partial z} \right), \quad (5.23)$$

and therefore the effective spin-orbit field is given as

$$\mathbf{w}_B(\mathbf{k}_{\parallel}) \approx \mathbf{w}(k_x - b_2 B \sin \theta, k_y + b_2 B \cos \theta) \quad (5.24)$$

where b_2 is a coefficient that only depends on $\langle z \rangle$, as the spin-orbit terms only contain linear momenta in the transverse directions. Thus, in general we can expect $b_1 \neq b_2$.

With the approximations (5.22) and (5.24) we can now evaluate the total transmission. Since the zeroth order term is to first order only shifted by the magnetic field, we obtain the same result as in the case without magnetic field after integration:

$$\langle T_B^0(E, \mathbf{k}_{\parallel}) \rangle_{\mathbf{k}_{\parallel}} \approx T^{(0)} \quad (5.25)$$

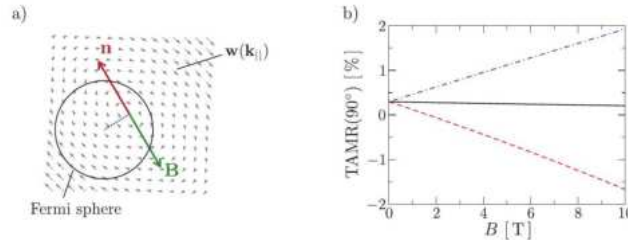


Figure 5.4 - (a) Schematic picture of the influence of the magnetic field: The Fermi sphere (or, more precisely, the maximum of the transmission probability) is shifted with respect to the effective spin-orbit field $\mathbf{w}(\mathbf{k}_{\parallel})$. (b) Magnetic field dependence of TAMR(90°) when the magnetic vector potential is only included in the kinetic term (red dashed line), only in the spin-orbit term (blue dash-dotted line), or in both terms (black solid line).

The term linear in the spin-orbit coupling evaluates to

$$\begin{aligned}
& \langle T_B^1(E, \mathbf{k}_{\parallel}) \mathbf{n} \cdot \mathbf{w}_B(\mathbf{k}_{\parallel}) \rangle_{\mathbf{k}_{\parallel}} \\
& \approx \frac{1}{(2\pi)^2} \int d\mathbf{k}_{\parallel} T^{(1)}(E, k_{\parallel}) \mathbf{n} \cdot \mathbf{w}(k_x + (b_1 - b_2)B \sin \theta, k_y - (b_1 - b_2)B \cos \theta) \\
& = \frac{1}{(2\pi)^2} \int d\mathbf{k}_{\parallel} T^{(1)}(E, k_{\parallel}) B (b_1 - b_2) \\
& \quad (-\tilde{\alpha} \cos^2 \theta + \tilde{\gamma} \sin \theta \cos \theta - \tilde{\alpha} \sin^2 \theta + \tilde{\gamma} \cos \theta \sin \theta) \\
& = (b_1 - b_2) B \left(\langle T^{(1)}(E, k_{\parallel}) f_{\gamma}(k_{\parallel}) \rangle \gamma \sin(2\theta) - \langle T^{(1)}(E, k_{\parallel}) f_{\alpha}(k_{\parallel}) \rangle \alpha \right). \quad (5.26)
\end{aligned}$$

In the second line we used the coordinate substitution $k_x \rightarrow k_x + b_1 B \sin \theta$ and $k_y \rightarrow k_y - b_1 B \cos \theta$, and in the third line we made use of the fact that terms linear in k_x or k_y vanish upon integration.

Hence, we find that the linear term, that vanished in the case without a magnetic field, gives a finite contribution for $B \neq 0$. This is due to the fact that the Fermi sphere is shifted by the Lorentz force. In particular, the effective spin-orbit field is shifted with respect to the center of the transmission maximum, as sketched in Fig. 5.4(a). In addition, we find a different angular dependence for the Bychkov-Rashba and the Dresselhaus spin-orbit coupling. This is due to different symmetries of the spin-orbit fields: The Bychkov-Rashba field exhibits a rotational symmetry [13] and thus leads to an angular-independent term in Eq. (5.26), whereas the Dresselhaus field does not, and hence leads to an angular dependence.

Finally, in addition to the original contribution for $B = 0$, the second-order term in the spin-orbit coupling also contains terms quadratic in B , due to the shift of the Fermi sphere. As before, we neglect these terms and approximate

$$\langle T_B^2(E, \mathbf{k}_{\parallel}) \rangle_{\mathbf{k}_{\parallel}} \approx T^{(2)}. \quad (5.27)$$

Hence, the conductance of the magnetic tunnel junction in a magnetic field B takes the form

$$G_B(\phi) = G^{(0)} + a_0^{(2)} + a_2^{(1)} \alpha B + (a_1^{(2)} \alpha \gamma - a_1^{(1)} \gamma B) \cos(2\phi), \quad (5.28)$$

where the coefficients

$$a_1^{(1)} = -\frac{Se^2}{(2\pi)^2 h} \langle T^{(1)}(E, k_{\parallel}) f_{\gamma}(k_{\parallel}) \rangle (b_1 - b_2) \quad (5.29a)$$

$$a_2^{(1)} = -\frac{Se^2}{(2\pi)^2 h} \langle T^{(1)}(E, k_{\parallel}) f_{\alpha}(k_{\parallel}) \rangle (b_1 - b_2) \quad (5.29b)$$

do not depend on the spin-orbit parameters. The TAMR then evaluates to

$$\text{TAMR}(\phi) \approx \frac{G_B(0) - G_B(\phi)}{G^{(0)}} \propto (a_1^{(2)} \alpha \gamma - a_1^{(1)} \gamma B) (1 - \cos(2\phi)). \quad (5.30)$$

This equation gives a qualitative explanation of all characteristic features observed in the experiment: First, it reproduces the linear change in the TAMR with magnetic field. Second, the slope of this change is independent of the bias voltage, as it only depends on the (bulk) Dresselhaus parameter. Thus, all the experimentally observed features can be explained with the orbital effects caused by the magnetic field.

From the numerical results we can deduce that $a_1^{(2)}, a_1^{(1)} > 0$. In addition, the numerical results

also confirm that $a_1^{(1)}$ contains a factor $(b_1 - b_2)$, i.e. opposing contributions from the orbital effects in the kinetic energy and in the spin-orbit coupling. In Fig. 5.4(b) we show the magnetic field dependence of TAMR (90°) when the magnetic vector potential is included only in the kinetic term (dashed line), only in the spin-orbit terms (dash-dotted line) and in both terms (solid line). When the magnetic field is only included in one term, we find a large change of the TAMR effect with magnetic field, however with opposite signs. When the vector potential is included in both the kinetic and the spin-orbit term, the contributions cancel to a large degree, leaving only a small effect. These numerical findings support the phenomenological model developed in this section.

Finally, it should be mentioned that although the magnetic field dependence of the TAMR (5.30) only depends on the Dresselhaus parameter γ , the conductance (5.28) also contains a B -dependent contribution from the Bychkov-Rashba parameter α . This term leads to a bias-voltage dependent linear slope in the magnetic field dependence of the conductance, in addition to the usual quadratic dependence [156]. Hence, it would be of great interest if such a dependence was indeed observed in the experiments.

19.4.6 5.3 Summary

In this chapter we have studied the magnetic field dependence of the TAMR effect in epitaxially grown Fe/GaAs/Au tunnel junctions. For this, we have extended the model used in previous works [47, 151, 152], which explained the TAMR effect through the spin-orbit coupling in the GaAs barrier, to include the orbital effects of a magnetic field. No additional fitting parameters were introduced in this step.

The observed magnetic field dependence - linear change in the TAMR with B and a bias-independent slope - could be explained both from the numerical simulations and a phenomenological model, where the orbital effects of the magnetic field were included as a shift of the Fermi sphere. In particular, the phenomenological model predicts that the slope of the TAMR change for a magnetic field B depends only on the (bulk) Dresselhaus spin-orbit parameter, and is hence bias-independent. This and other predictions from the phenomenological model are reproduced in numerical simulations, supporting its validity.

The TAMR effect in Fe/GaAs/Au tunnel junctions has been attributed to the effects of spin-orbit coupling in the GaAs barrier [47, 151, 152]. Including orbital effects into this model, we could explain the experimentally observed magnetic field dependence without any additional fitting parameters. This strongly supports the initial assumption that indeed spin-orbit coupling is the reason for the TAMR effect.

19.5 Chapter 6 The graphene edge state

19.5.1 6.1 Introduction

Graphene is a two-dimensional crystal of carbon atoms. Since its experimental discovery in 2004 [29], graphene has attracted a tremendous amount of interest (for reviews, see [31, 35, 157]). On the one hand, this is due to the fact that charge carriers in graphene obey an effective equation that mimics relativistic dynamics, leading to many unique physical properties, such as the odd-integer quantum Hall effect [32] that even prevails at room temperature [33], or Klein tunneling [34]. On the other hand, graphene is also a very promising material for electronics, because of its high mobility [36, 37], even at room-temperature [29], paving the road to ballistic transistors [31].

However, since (bulk) graphene does not feature a band gap, conventional transistor action is difficult to achieve. In contrast, graphene nanoribbons, strips of finite width cut from a graphene sheet, exhibit a finite band gap scaling with the nanoribbon width [158-160], and, as

a consequence, room-temperature transistor action has been demonstrated [161]. The origin of this band gap is still under discussion [162-166].

Whereas the edges of graphene nanoribbons were not well-controlled in the early experiments, recent progress in fabrication has shown that controlled manufacturing of nanoribbons with atomically defined edges seems within experimental reach in the near future [160, 167, 168]. From a theoretical point of view, the electronic properties of nanoribbons are expected to depend crucially on the structure of the edge [169]. In particular, a zigzag or zigzag-like edge is expected to support a state localized at the graphene edge [169-171]. This edge state has been observed experimentally using scanning tunneling (STM) microscopy for a graphene edge on the surface of graphite¹ [172-174]. Due to electron-electron interactions, the edge state is expected to exhibit

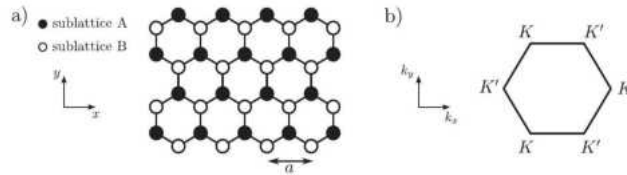


Figure 6.1 - (a) Lattice structure of graphene. (b) First Brillouin zone of graphene. magnetism [170, 175, 176]. This magnetism can be used in edge state based spintronics that we present in Chapter 7.

On the theoretical side, the transport in situations, where the edge state is important, is almost exclusively studied in nearest-neighbor tight-binding approximation (for example, see Refs. [162, 165, 166, 177-181]). However, it has already been shown that the effects of next-nearest neighbor hopping on the edge state are significant in terms of the band structure [182, 183]. In addition, edge magnetism also changes the edge state related band structure considerably [164, 176].

Therefore, it is the intent of this chapter to study the fundamental properties of edge state transport in graphene. Doing so, we show that the paradigm model of transport in graphene, the nearest-neighbor tight-binding model, is not suitable for studying edge state transport: The results of edge transport in nearest-neighbor tightbinding approximation change fundamentally upon exponentially small corrections to this model.

To this end, the chapter is organized as follows: We begin by summarizing the main results of the tight-binding model of graphene employed in this work (for a complete discussion, see Appendix G). In addition, we review symmetries of the graphene Hamiltonian, that impose certain restrictions on the graphene wave function. These results are then applied to characterize the properties of the edge state for different approximations to the graphene Hamiltonian. Finally, we discuss extensively the transport properties of the graphene edge state when corrections to the nearest-neighbor tightbinding approximation are included.

19.5.2 6.2 Fundamentals of graphene

19.5.3 6.2.1 Graphene Hamiltonian

The electronic structure of graphene is usually described in terms of a tight-binding model, that is discussed extensively in Appendix G. Here, we merely summarize the main results.

Graphene is a two-dimensional sheet of carbon atoms, arranged in a honeycomb network, as shown in Fig. 6.1(a). The honeycomb network is not a primitive Bravais lattice, but is generated by a triangular lattice with two basis atoms. These basis atoms are labeled A and B, respectively, and the triangular lattice constant is denoted as a .

¹²¹ Up to now, a direct experimental confirmation of the edge state for exfoliated or epitaxially grown graphene is still missing.

In single-orbital tight-binding approximation, the graphene Hamiltonian reads

$$H = \sum_{i,\alpha,j,\beta} t_{i\alpha j\beta} |i, \alpha\rangle \langle j, \beta|. \quad (6.1)$$

Here, $|i, \alpha\rangle$ denotes the orbital at lattice site $\mathbf{R}_i + \mathbf{d}_\alpha$, where $\mathbf{R}_i$ is a (triangular) lattice vector and $\mathbf{d}_\alpha$ the relative position of atom $\alpha = A, B$ within the unit cell. We consider matrix elements up to next-nearest neighbors, so that

$$t_{i\alpha j\beta} = \begin{cases} V_i + M_i & \text{if } i = j, \alpha = \beta = A \\ V_i - M_i & \text{if } i = j, \alpha = \beta = B \\ -t & \text{if } i, \alpha \text{ and } j, \beta \text{ are nearest neighbors,} \\ -t' & \text{if } i, \alpha \text{ and } j, \beta \text{ are next-nearest neighbors.} \end{cases} \quad (6.2)$$

Here V_i is a potential that has the same value on both sublattices, and M_i is a staggered potential, i.e. has opposite sign on atom A and B, respectively. The numerical values of the hopping parameter are chosen as $t = 2.7\text{eV}$ and $t' = 0.1t$ [183, 184].

In the case of an infinite graphene sheet and constant potentials $V_i = V_0$ and $M_i = M$, Bloch's theorem holds and in $\mathbf{k}$ -space the Hamiltonian reads (Eq. (G.23))

$$H(\mathbf{k}) = \left(V_0 - 2t' \sum_{l=1}^3 \cos(\mathbf{k}\delta'_l) \right) \mathbb{I} - \left(t \sum_{l=1}^3 \cos(\mathbf{k}\delta_l) \right) \sigma^x + \left(t \sum_{l=1}^3 \sin(\mathbf{k}\delta_l) \right) \sigma^y + M\sigma^z \quad (6.3)$$

where $\sigma^{x,y,z}$ are the Pauli matrices and δ_l (δ'_l) the distance vectors between nearest (next-nearest) neighbors. Note that the two basis atoms in the unit cell give rise to a wave function with two components, referring to sublattice A and B,

$$\varphi(j) = \begin{pmatrix} \varphi_A(j) \\ \varphi_B(j) \end{pmatrix} \quad (6.4)$$

such that the Hamiltonian $H(\mathbf{k})$ can be written as a 2×2 -matrix in Eq. (6.3). The wave vector $\mathbf{k}$ is defined within the first Brillouin zone of graphene that has the shape of a hexagon, as shown in Fig. 6.1(b). Only two of the six corners of the Brillouin zone are unique, as always three corners are connected by a reciprocal lattice vector. These two non-equivalent corners are called K and K' -point, respectively. Here we choose

$$\mathbf{K}_\tau = \begin{pmatrix} \tau \frac{4\pi}{3a} \\ 0 \end{pmatrix} \quad (6.5)$$

to represent the K ($\tau = +1, \mathbf{K} = \mathbf{K}_+$) and K' -point ($\tau = -1, \mathbf{K}' = \mathbf{K}_-$).

In a system close to charge neutrality, the states at the Fermi energy E_F have a wave vector close to the K and K' -point. States close to $\mathbf{K}$ and $\mathbf{K}'$ are said to be in the K and K' -valley. In the low energy regime and for smoothly varying potentials, the particle dynamics can be described by an effective Hamiltonian in each valley (Eq. (G.36)),

$$H_\tau = -i\hbar v_F (\tau \partial_x \sigma^x + \partial_y \sigma^y) + V(\mathbf{x}) + M(\mathbf{x}) \sigma^z \quad (6.6)$$

where $H_K = H_+$ is the effective Hamiltonian for the K -valley and $H_{K'} = H_-$ for the K' -valley. Here, $V(\mathbf{x})$ and $M(\mathbf{x})$ are a smoothly varying potential and staggered potential, respectively, and $v_F = \sqrt{3}ta/2 \approx 10^6$ m/s is the Fermi velocity. Again, the two sublattices give rise to a two-component envelope wave function

$$\varphi(\mathbf{x}) = \begin{pmatrix} \varphi_A(\mathbf{x}) \\ \varphi_B(\mathbf{x}) \end{pmatrix} \quad (6.7)$$

and the Pauli matrices $\sigma^{x,y,z}$ act within this sublattice space.

The effective Hamiltonian in Eq. (6.6) is equivalent to a Dirac Hamiltonian in two spatial dimensions, albeit with an "effective speed of light" v_F . As a consequence, the low-energy spectrum of graphene is linear close to the K and K' -point. This gives rise to a cone-like shape of the band structure in the K and K' -valley, that is commonly referred to as Dirac cone. Furthermore, the two-component wave function (6.7) behaves like a spinor under transformations, and the sublattice degree of freedom is thus referred to as pseudospin.

The envelope wave function (6.7) and the tight-binding wave function (6.4) are connected through (Eq. (G.38))

$$\varphi_\alpha(j) = e^{i\mathbf{K}(\mathbf{R}_j + \mathbf{d}_\alpha)} \varphi_\alpha(\mathbf{R}_j + \mathbf{d}_\alpha) + e^{i\mathbf{K}'(\mathbf{R}_j + \mathbf{d}_\alpha)} \varphi'_\alpha(\mathbf{R}_j + \mathbf{d}_\alpha) \quad (6.8)$$

where $\varphi_\alpha(\mathbf{x})$ and $\varphi'_\alpha(\mathbf{x})$ are the envelope wave functions in the K and K' -valley, respectively.

19.5.4 6.2.2 Symmetries of the graphene Hamiltonian

We will now review some important symmetries of the graphene Hamiltonian, in particular particle-hole symmetries. If the Hamiltonian H is particle-hole symmetric, there is a transformation $\mathcal{P}$ such that

$$\mathcal{P}^\dagger H \mathcal{P} = -H \quad (6.9)$$

Then, if φ is an eigenstate of H with energy E , $H\varphi = E\varphi$, $\mathcal{P}\varphi$ is an eigenstate with energy $-E$, $H\mathcal{P}\varphi = -E\mathcal{P}\varphi$. Thus, the energy spectrum is symmetric with respect to $E = 0$ and every "particle" wave function φ with positive energy has a "hole" partner $\mathcal{P}\varphi$ with negative energy.

Writing the tight-binding Hamiltonian in terms of Pauli matrices as in Eq. (6.3) has the advantage that we can treat it on equal footing with the effective Dirac Hamiltonian (6.6). The symmetries and their consequences considered below are thus valid for both the tight-binding and the effective Hamiltonian.

The Hamiltonian is particle-hole symmetric under the unitary transformation

$$\mathcal{P}_z = \sigma^z \quad (6.10)$$

if it contains only terms proportional to σ^x or σ^y . That is, $\mathcal{P}_z$ is a particle-hole transformation if the Hamiltonian does not include potentials $^2V(\mathbf{x})$ and $M(\mathbf{x})$ and, in the case of the tight-binding Hamiltonian, only includes nearest-neighbor hopping ³.

In order to exist also in a system of finite extent, the particle-hole symmetry must additionally be compatible with the boundary conditions [171]. In a finite graphene patch, the wave function must vanish on a set of missing atoms, as discussed in Section G.3.2. Since $\mathcal{P}_z$ does not mix sublattice A and B, it is compatible with any boundary.

The particle-hole transformation $\mathcal{P}_z$ is especially important, as it also has consequences for the occupation probabilities of the two sublattices [185]. Note that $\frac{1}{2}(1 \pm \sigma^z)$ is a projection operator onto the wave function of sublattice A and B, respectively. From this projection operator we obtain the following expression for an eigenstate φ with energy E , provided that $\{\mathcal{P}_z, H\} = 0$:

$$\begin{aligned}
\langle \varphi | \frac{1}{2} (1 + \sigma^z) H \frac{1}{2} (1 - \sigma^z) | \varphi \rangle &= \langle \varphi | \frac{1}{2} (1 + \sigma^z) \frac{1}{2} (1 + \sigma^z) H | \varphi \rangle \\
&= E \int d\mathbf{x} |\varphi_A(\mathbf{x})|^2
\end{aligned} \tag{6.11}$$

Commuting the Hamiltonian H to the left we also obtain

$$\langle \varphi | \frac{1}{2} (1 + \sigma^z) H \frac{1}{2} (1 - \sigma^z) | \varphi \rangle = E \int d\mathbf{x} |\varphi_B(\mathbf{x})|^2 \tag{6.12}$$

and thus, combining Eqs. (6.11) and (6.12),

$$\int d\mathbf{x} |\varphi_A(\mathbf{x})|^2 = \int d\mathbf{x} |\varphi_B(\mathbf{x})|^2 \quad \text{for } E \neq 0 \tag{6.13}$$

Hence, if the Hamiltonian is particle-hole symmetric under $\mathcal{P}_z$, every eigenstate with energy $E \neq 0$ occupies sublattice A and B with equal probability.

$\mathcal{P}_z$ is not the only particle-hole symmetry in the system. For example, the effective Hamiltonian is particle-hole symmetric under the antiunitary transformation [186]

$$\mathcal{P}_x = \mathcal{C} \sigma^x \tag{6.14}$$

	σ^x, σ^y	$M(\mathbf{x})\sigma^z$	$V(\mathbf{x})\mathbb{1}$	$t' \neq 0$
particle-hole symmetry	yes	yes ^a	no	no
$\int d\mathbf{x} \varphi_A(\mathbf{x}) ^2 = \int d\mathbf{x} \varphi_B(\mathbf{x}) ^2$	yes ^b	no	no	no

Table 6.1 - Summary of which terms in the graphene Hamiltonian break particle-hole symmetry, and which terms imply equal occupation probability on sublattice A and B.

^a In bulk; for armchair boundaries; for zigzag boundaries provided that there is an additional symmetry connecting zigzag edges of the opposite kind (such as in nanoribbons).

^b For $E \neq 0$.

where $\mathcal{C}$ denotes complex conjugation. The tight-binding Hamiltonian, in turn, is particle-hole symmetric under

$$\mathcal{P}_y = \mathcal{C} \sigma^y. \tag{6.15}$$

Both $\mathcal{P}_x$ and $\mathcal{P}_y$ are particle-hole transformations if the respective Hamiltonian only contains terms proportional to σ^x, σ^y and σ^z , i.e. including a staggered potential.

However, unlike $\mathcal{P}_z$, these particle-hole symmetries interchange sublattice A and B and are thus only compatible with boundary conditions that involve both sublattices on equal footing, such as armchair boundary conditions. In particular, a zigzag edge breaks these particle hole symmetries-unless there is an additional symmetry connecting different zigzag edges as in a zigzag nanoribbon. Still, even if $\mathcal{P}_{x,y}$ is a particle-hole transformation, it does not give rise to a relation for the occupation probability of the two sublattices.

The findings of this section are summarized in Table 6.1.

¹²² Note that a constant potential $V(\mathbf{x}) = V_0$ only leads to an overall energy shift; as such, it does not break particle-hole symmetry.

³ In the low-energy regime, the next-nearest neighbor hopping does not play a role (see Section G.2.3). Thus, close to the Dirac point, the spectrum is still approximately particle-hole symmetric, even if next-nearest neighbor hopping is included. However, this is not true anymore, if edge states are present (see Section 6.3.4).

19.5.5 6.3 Characterization of the edge state

19.5.6 6.3.1 Edge state at a single zigzag edge

Over the course of the next sections, we will examine the graphene edge state under different conditions. Here, we start with the simplest situation, a half-infinite graphene sheet with a single zigzag edge in nearest-neighbor tight-binding approximation.

It has been known for some time that a zigzag edge in a honeycomb network supports a state localized at the edge [96, 187], but it was Fujita that first noticed the physical relevance of this state [170]. As Fujita, we consider a half-infinite graphene sheet with a single zigzag edge in x -direction, as sketched in Fig. 6.2(a). In this situation, the outermost atoms are all of sublattice B, and the wave function vanishes on the first row of missing A atoms.

The half-infinite graphene sheet can be understood as a repetition of unit cells of length a , where the unit cell is given as a chain of carbon atoms, as indicated in black in Fig. 6.2(a). We label atoms within this unit cell by (j, α) , where $j = 0, 1, 2, \dots$ increases when moving away from the edge and $\alpha = A, B$ denotes the sublattice. Note that in

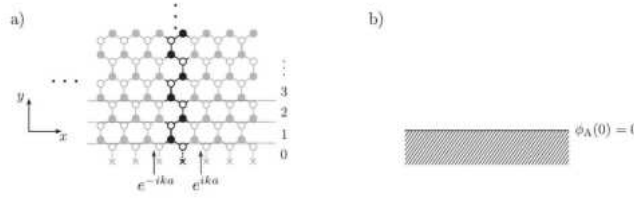


Figure 6.2 - (a) A half-infinite graphene sheet with a single zigzag edge in nearest-neighbor tight-binding approximation. The unit cell of the system is indicated in black, and the first row of missing atoms is indicated by crosses. (b) The same situation within the effective Dirac Hamiltonian.

our convention, the first missing atom has index $(0, A)$. The unit cells themselves are labeled with an index n , that increases with increasing x .

Because of translational symmetry in x -direction, we can write the wave function of the system in Bloch form,

$$\varphi_\alpha(n, j) = \tilde{\phi}_\alpha(j) e^{ikan} \quad (6.16)$$

where $k \in [-\frac{\pi}{a}, \frac{\pi}{a}]$. Within a unit cell, the atomic positions change by $\pm \frac{a}{2}$ in x -direction when going from j to $j+1$ or $j-1$. In order to eliminate the phase associated with this lateral displacement, it is convenient to define

$$\begin{aligned} \tilde{\phi}_\alpha(j) &= \phi_\alpha(j) & \text{for } j \text{ even,} \\ \tilde{\phi}_\alpha(j) &= \phi_\alpha(j) e^{-ika/2} & \text{for } j \text{ odd} \end{aligned} \quad (6.18)$$

Making use of the Bloch form (6.16) we can then write down the stationary Schrödinger equation $H\varphi = E\varphi$ as a set of equations only involving $\phi_\alpha(j)$:

$$E\phi_A(j) = -t\phi_B(j) + te^{-\gamma}\phi_B(j-1) \quad (6.19)$$

$$E\phi_B(j) = -t\phi_A(j) + te^{-\gamma}\phi_A(j+1), \quad (6.20)$$

where we defined $e^{-\gamma} = -2\cos(ka/2)$. This set of equations is closed by the boundary condition $\phi_A(0) = 0$, so that

$$E\phi_B(0) = -t\phi_A(1) \quad (6.21)$$

It is easy to see that these equations have a solution for $E = 0$:

$$\phi_A(j) = 0, \quad \phi_B(j) = e^{-\gamma j}. \quad (6.22)$$

This solution is normalizable - and thus represents a physical solution-if $|e^{-\gamma}| = |2 \cos(ka/2)| < 1$, i.e. $-\frac{\pi}{a} \leq k < -\frac{2\pi}{3a}$ and $\frac{2\pi}{3a} < k < \frac{\pi}{a}$. From

$$\sum_{j=0}^{\infty} e^{-2\gamma j} = \frac{1}{1 - e^{-2\gamma}} \quad (6.23)$$

we find the normalized solution

$$\phi_A(j) = 0 \quad \phi_B(j) = \sqrt{1 - e^{-2\gamma}} e^{-\gamma j}. \quad (6.24)$$

This is the celebrated graphene edge state. Note that this state resides on a single sublattice - since it is a zero-energy solution, Eq. (6.13) does not hold, although the nearest-neighbor tight-binding Hamiltonian is particle-hole symmetric with respect to $\mathcal{P}_z$.

The edge state is a solution that decays exponentially in the direction perpendicular to the edge. The perpendicular decay constant γ , and thus the perpendicular extension of the edge state, depends uniquely on the lateral wave vector k , $\gamma = \gamma(k)$. This is in contrast to an edge in an ordinary, semiconductor based two-dimensional electron gas, where the perpendicular and the lateral direction are decoupled. The edge state is maximally localized for $k = \pm \frac{\pi}{a}$, where it resides on the outermost atom only,

$$|\phi_B(j)|^2 = \begin{cases} 1 & \text{for } j = 0 \\ 0 & \text{else} \end{cases} \quad (6.25)$$

and extends through the whole system for $k = \mp \frac{2\pi}{3a}$ (that is, at the K and K' -point, respectively),

$$|\phi_B(j)|^2 = \text{const.} \quad (6.26)$$

Close to the K -points, we find

$$\gamma = i\pi - \frac{\sqrt{3}a}{2} \left(k \mp \frac{2\pi}{3a} \right) + \mathcal{O} \left(\left(k \mp \frac{2\pi}{3a} \right)^2 \right), \quad (6.27)$$

i.e. the perpendicular decay constant depends linearly on the lateral wave vector k .

In contrast to the decay length, the energy of the edge state $E(k) = 0$ is a constant, independent of k . Thus, the edge state is dispersionless, i.e. has zero group velocity. In Fig. 6.3 we show a schematic picture of the band structure of a half-infinite graphene sheet including the edge state (for a discussion on how to project the two-dimensional band structure of graphene into one dimension, see Section G.3.1). The edge state band is perfectly flat and connects the two K -points.

After discussing the edge state within the tight-binding model, we now turn to the effective Dirac Hamiltonian. Again, because of translational symmetry in x -direction, the wave function of the system can be written in Bloch form

$$\varphi(\mathbf{x}) = \phi(y) e^{iqx} \quad (6.28)$$

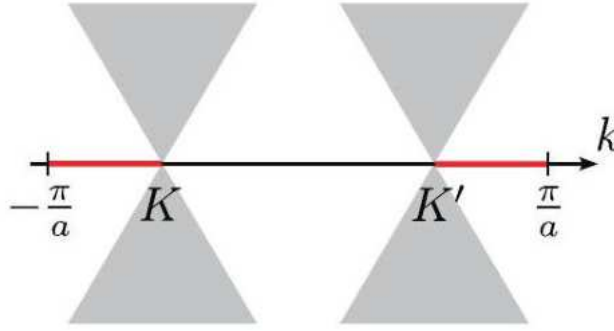


Figure 6.3-Schematic depiction of the edge state dispersion $E(k)$, shown in red. The Dirac cones are indicated in light grey.

Note that a wave vector in the effective Hamiltonian is always given with respect to the K or K' -point, as discussed in Section G.2.3. We will always denote such a relative wave vector by q , in order to distinguish it from a wave vector k in the tight-binding model.

Using Eq. (6.28) in the effective Dirac Hamiltonian (6.6) we obtain a one-dimensional differential equation for the transverse wave function

$$\hbar v_F (\tau q - \partial_y) \phi_B(y) = E \phi_A(y) \quad (6.29)$$

$$\hbar v_F (\tau q + \partial_y) \phi_A(y) = E \phi_B(y) \quad (6.30)$$

where $\tau = +1$ for the K -valley and $\tau = -1$ for the K' -valley. The boundary condition for a half-infinite graphene sheet as sketched in Fig. 6.2(b) is then

$$\phi_A(0) = 0 \quad (6.31)$$

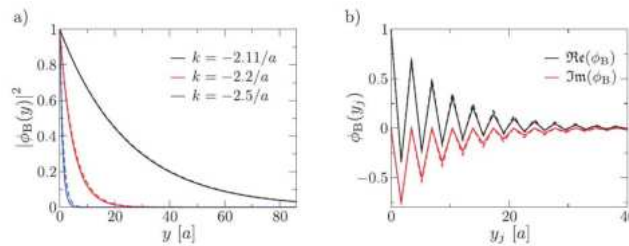
For $E = 0$, these equations have the solution

$$\phi_A(y) = 0 \quad \phi_B = e^{\tau q y} \quad (6.32)$$

This solution is normalizable, if $\tau q < 0$, i.e. for negative (positive) wave vectors q with respect to the K (K')-point. Note that this is consistent with the tight-binding result, where $k < -\frac{2\pi}{3a}$ (K -point) or $k > \frac{2\pi}{3a}$ (K' -point). The normalized edge state solution then reads

$$\phi(y) = \begin{pmatrix} 0 \\ \sqrt{-2\tau q} e^{\tau q y} \end{pmatrix}. \quad (6.33)$$

Thus, also within the effective Hamiltonian we find an edge state solution that decays exponentially in the perpendicular direction. Again, the decay length depends on the lateral wave vector q . Furthermore, the perpendicular decay constant τq is in agreement with the linear expansion of γ close to the K -points, Eq. (6.27).⁴ Thus, close to the K -points, the edge state is well-described by the effective Dirac Hamiltonian.



¹²⁴ For this, note that the y -coordinate of the j -th atom is given by $y_j = j\sqrt{3}a/2$.

Figure 6.4 - Comparison between the edge state solutions of the tight-binding (solid lines) and the effective Hamiltonian (dashed lines). (a) shows the probability density $|\phi_B(y)|^2$ for different values of k , and (b) the wave functions $\phi_B(y)$ for $k = -2.3/a$ as a function of the distance y from the edge. In (b), we denote this distance as y_j in order to emphasize that the values of the wave function are only given for fixed atom positions. In order to conveniently compare the wave functions for different values of k , they have been normalized such that $\phi_B(0) = 1$.

In Fig. 6.4(a) we show a comparison between the edge state from the tight-binding and the effective Hamiltonian for different values of k . As expected, close to the K point, both solutions agree very well. Moving away from the K -point, the edge state from the effective Hamiltonian decays somewhat slower than the edge state within the tight-binding model. In fact, the tight-binding edge state solution is maximally localized for $k = \pm \frac{\pi}{a}$, whereas the edge state solution in the effective Hamiltonian is localized strongest for $\tau q \rightarrow -\infty$. The point of maximal localization, $k = \pm \frac{\pi}{a}$, is located halfway between the K and K' -points. Around this point, the validity of the effective Hamiltonian, that is an expansion around a single K -point, breaks down. Thus, it is not surprising that the effective Hamiltonian underestimates the strength of the edge state decay, but as seen in Fig. 6.4, this deviation is small.

In addition to the exponential decay, the edge state also features a change in sign from atom to atom, since $e^{-\gamma} < 0$. In order to obtain such an oscillation on the scale of the lattice from the smooth envelope wave functions of the effective Hamiltonian, we have to use Eq. (6.8) to relate the envelope wave functions to the tight-binding solution. For the case of the edge state, this is shown in Fig. 6.4(b), where we find a good agreement between both solutions.

Hence, the effective Dirac Hamiltonian describes the main features of the graphene edge state rather well, but somewhat underestimates the strength of the decay. Therefore, when keeping in mind this limitation, the effective Hamiltonian can be applied successfully to study the edge state in more complex situations, when an analytical solution of the tight-binding model is not possible anymore.

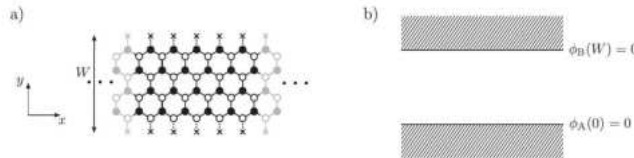


Figure 6.5 - (a) A zigzag graphene nanoribbon and (b) its equivalent within the effective Dirac Hamiltonian.

19.5.7 6.3.2 Edge state in a zigzag nanoribbon

We now turn to a system containing two parallel zigzag edges, i.e. a zigzag graphene nanoribbon. It is still possible to solve the tight-binding Hamiltonian in nearestneighbor approximation analytically in this case [183]. However, the resulting expressions are rather unwieldy. Thus, here we restrict the discussion to the effective Dirac Hamiltonian in order to obtain the characteristic features of the edge state in a nanoribbon.

Since the zigzag nanoribbon again has translational symmetry in x -direction, the eigenstates will have the Bloch form as given in Eq. (6.28). The boundary conditions then read

$$\phi_A(0) = 0, \quad \phi_B(W) = 0, \quad (6.34)$$

where W is the width of the nanoribbon (see Fig. 6.5). The problem of solving the effective Dirac Hamiltonian (6.6) with these boundary conditions has been addressed in Refs. [53, 185]. The edge state solution is given as

$$\phi(y) \cong \begin{pmatrix} \sinh(zy) \\ \text{sgn}(E) \sinh(z(W-y)) \end{pmatrix} \quad (6.35)$$

where the decay constant z is real and obeys

$$q = -\tau \frac{z}{\tanh(zW)}. \quad (6.36)$$

Again, $\tau = +1$ for the K and $\tau = -1$ for the K' -valley. The edge state has the eigenenergy

$$E = \pm \frac{z}{\sinh zW} \quad (6.37)$$

The solution (6.35) is decaying exponentially in the perpendicular direction and can thus be called an edge state, if the transcendental equation (6.36) has a real solution z . This is the case for $\tau q < -1/W$. Note that the critical value of q for obtaining an edge state solution has been shifted slightly with respect to the single edge, where

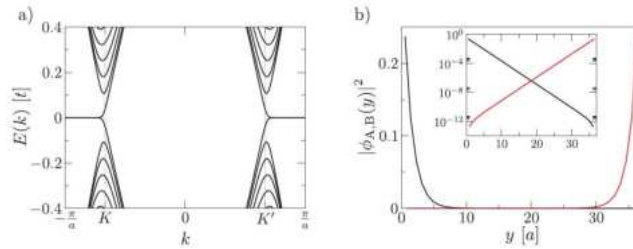


Figure 6.6 - A zigzag nanoribbon in nearest-neighbor tight-binding approximation: (a) The band structure of the zigzag nanoribbon. (b) Probability density profile of the edge state across the zigzag nanoribbon, for $k = -2.4/a$. The inset shows the same density profile on a logarithmic scale. The probability density for sublattice A is shown in red, for sublattice B in black. In both (a) and (b) the nanoribbon had a width $W = 64a/\sqrt{3}$.

$\tau q < 0$ was required for the existence of an edge state. For wide enough ribbons, this difference becomes negligible.

The state at $q = -\tau/W$ has $z = 0$ and connects edge state and bulk-like solutions, as does the $q = 0$ state for the single edge. However, unlike for the case of the single edge, where $\phi_B(y) = \text{const.}$, this connecting state decreases linearly from one edge to the other [53], as it must satisfy the boundary conditions at both edges.

For $zW \gg 1$, i.e. a very wide ribbon, or a strongly decaying state, from Eq. (6.36) we obtain $z \approx -\tau q$ and the eigenstate simplifies to

$$\phi(y) \approx \begin{pmatrix} e^{\tau q(W-y)} \\ \pm e^{\tau q y} \end{pmatrix}. \quad (6.38)$$

In this limit, the edge state in the zigzag nanoribbon is given as two decaying exponentials at each edge, residing on opposite sublattices (the boundary conditions at both edges are opposite to each other), reminiscent of the edge state at a single edge as discussed in the previous section. From Eq. (6.37) we find the energy of this state as

$$E \approx \mp 2\tau q e^{\tau q W}. \quad (6.39)$$

This energy is exponentially close to zero, since $-\tau q W = zW \gg 1$. Thus, we find two states with an exponentially small, but finite energy splitting. There are two states, since the nanoribbon features two edges.

These states appear as two almost dispersionless bands in the band structure of a zigzag nanoribbon as shown in Fig. 6.6(a), where we calculated numerically the band structure of a zigzag nanoribbon in nearest-neighbor tight-binding approximation.

Since the energy of these states is exponentially small, these bands appear to have the same energy when plotted in a linear scale. However, there is always an exponentially small, but finite energy splitting between these states at a given value of k , with the exception of the two zero-energy states at $k = \pm \frac{\pi}{a}$. There the effective Dirac theory breaks down, as seen in the previous section. Moving closer to the K -points, the splitting becomes larger, until the edge states join the extended, bulk-like states in the Dirac cones (see also Section G.3.1). The existence of this splitting has important consequences that we discuss below. Note that the edge state band merges smoothly into a band of bulk states: The states for $\tau q > -1/W$ are no longer localized to the edge, but extend through the whole width of the ribbon.

In Fig. 6.6(b) we show the probability density profile of an edge state across the nanoribbon. For the chosen value of the wave vector k , the edge state is localized to the edges, and resides on a single sublattice close to the edges, as expected from Eq. (6.38). The inset of Fig. 6.6(b) shows the probability density of the edge state in a logarithmic scale, where the exponential decay is more apparent. Close to the edges, we observe a small "kink" in the decay, as the state must also satisfy the boundary condition on the opposite edge.

A picture like Fig. 6.6(b) might lead to the impression that the two edges in the zigzag nanoribbon carry two independent edge states. This is not true: Fig. 6.6(b) shows a single eigenstate of the nanoribbon. The nanoribbon in nearest-neighbor tight-binding approximation is particle-hole symmetric under transformation with $\mathcal{P}_z$, and thus every eigenstate with nonzero energy must occupy both sublattices with equal probability, as discussed in Section 6.2.2. It is thus not possible to find an eigenstate with nonzero energy that resides at a single edge and a single sublattice only. Thus, any edge state with $k \neq \pm \frac{\pi}{a}$ is located with equal probability at both edges. Since the probability densities at both edges arise from a single state, they are inevitably coupled: whatever happens at one edge, will also influence the probability density at the other edge. This must be kept in mind when discussing the edge state in a zigzag nanoribbon ⁵.

Note that the edge state solutions (6.35) and (6.38) are of the form

$$\begin{pmatrix} \phi_A(y) \\ 0 \end{pmatrix} \pm \begin{pmatrix} 0 \\ \phi_B(y) \end{pmatrix}, \quad (6.40)$$

i.e. a "bonding" and "anti-bonding" linear combination of wave functions on a single sublattice. Combining "bonding" and "anti-bonding" solutions, it is in fact possible to construct a wave function localized at single edge only. However, since these states have positive and negative energy respectively, the resulting wave function is not an eigenstate of the system.

Within the effective Dirac Hamiltonian, the "bonding" ("anti-bonding") solution has a positive (negative) energy. In order to obtain the lattice wave function from the envelope wave functions, we again have to employ (6.8). The lattice structure

of the zigzag nanoribbon then leads to a peculiar even-odd effect with respect to the nanoribbon width [188], that is of crucial importance in valleytronics [188, 189], but does not play a role in this work.

As seen in this section, due to the symmetries of the graphene Hamiltonian it is not possible to find a state localized at a single edge in a zigzag nanoribbon in nearestneighbor tight-binding approximation ⁶. In the next sections we therefore examine the behavior of the edge state including additional terms in the Hamiltonian that break this symmetry.

19.5.8 6.3.3 Effects of a staggered potential

We first consider the effect of a staggered potential M . In a half-infinite graphene sheet, the edge state resides on a single sublattice only. The staggered potential then simply shifts

¹²⁵ In fact, calling this state "edge state" is already a possible source of confusion.

the edge state energy by $-M$, if the edge state resides on sublattice B, and $+M$, if the edge state resides on sublattice A.

In a nanoribbon, the situation is more complex, as the edge state is in principle located at both edges and both sublattices, as discussed in the previous section. However, since a staggered potential breaks the particle-hole symmetry $\mathcal{P}_z$, the eigenstates of the zigzag nanoribbon need not occupy both sublattices equally, and the system can localize to a single edge, with the energy shifted by $\pm M$, as we show within the framework of the effective Dirac Hamiltonian below.

The analysis of Ref. [53] that lead to the results presented in the previous section carries over straightforwardly to a zigzag nanoribbon with a constant staggered potential. We find the edge state solution as

$$\phi(y) \cong \begin{pmatrix} \sqrt{\frac{E+M}{E}} \sinh(zy) \\ \text{sgn}(E+M) \sqrt{\frac{E-M}{E}} \sinh(z(W-y)) \end{pmatrix} \quad (6.41)$$

where the decay constant z is again a solution of

$$q = -\tau \frac{z}{\tanh(zW)}. \quad (6.42)$$

Thus, decaying states exist for $q < -\tau/W$, as before. The energy of these states is then given by

$$E = \pm \hbar v_F \sqrt{\left(\frac{M}{\hbar v_F}\right)^2 + \left(\frac{z}{\sinh zW}\right)^2}. \quad (6.43)$$

Hence, we find edge state solutions for energies $|E| \geq |M|$. Thus, the staggered potential opens a band gap of $2|M|$ for the edge state, just as for the bulk states (see Section

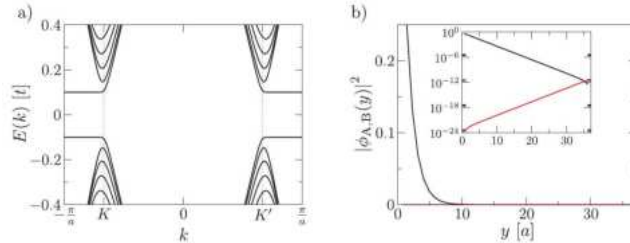


Figure 6.7 - A zigzag graphene nanoribbon in nearest-neighbor tight-binding approximation including a staggered potential $M = 0.1t$. (a) shows the band structure of the ribbon and (b) the probability density profile of the edge state across the nanoribbon, for $k = -2.4/a$ and $E \approx -M$. The inset shows the same density profile on a logarithmic scale. The probability density for sublattice A is shown in red, for sublattice B in black. In both (a) and (b) the nanoribbon had a width $W = 64a/\sqrt{3}$.

G.2.2). For $zW = -\tau qW \gg 1$ the energy dispersion simplifies to

$$E \approx \pm \hbar v_F \sqrt{\left(\frac{M}{\hbar v_F}\right)^2 + 4q^2 e^{-2\tau qW}} \quad (6.44)$$

which therefore is exponentially close to $\pm|M|$.

In Fig. 6.7(a) we show the band structure of a zigzag graphene nanoribbon in nearestneighbor tight-binding approximation with a constant staggered potential M . As predicted from the effective Hamiltonian, the staggered potential opens a band gap of size $2|M|$, so that we find two sets of almost dispersionless bands close to $E = \pm|M|$.

¹²⁶ As discussed above, this is not true for the zero-energy state at $k = \pm \frac{\pi}{a}$. However, since this is a single state, it will only provide a negligible contribution when we are studying transport in later sections.

Note that the functional form of the edge state in a staggered potential (6.41) is the same as for the edge state with $M = 0$, Eq. (6.35), but the wave functions for sublattice A and B are multiplied with different weights: For $E \approx M$, sublattice B is suppressed, whereas for $E \approx -M$ sublattice A is suppressed. Since $|E|$ is exponentially close to $|M|$ for $-\tau q W \gg 1$, the respective sublattices are also suppressed exponentially in this case. Hence we find an edge state solution that is predominantly localized at a single edge, as shown in Fig. 6.7(b). This exponential suppression of one sublattice (in this case sublattice A) and the fact that the functional form of the edge state solution does not change with $M \neq 0$ is most obvious in the inset of Fig. 6.7(b).

Hence, we find that a staggered potential localizes the edge state towards a single edge in a zigzag graphene nanoribbon. In this respect, the behavior of the edge state with a staggered potential is much more intuitive than without. At first glance such a staggered potential seems rather artificial, but we will discuss in Chapter 7 how antiferromagnetic ordering naturally induces such a staggered potential.

19.5.9 6.3.4 Effects of next-nearest neighbor hopping

Finally, we turn to the effects of next-nearest neighbor hopping and start by considering the half-infinite graphene sheet with a single zigzag edge.

For this case, Eqs. (6.19) and (6.20) can be extended to include next-nearest neighbor hopping:

$$E\phi_A(j) = -t\phi_B(j) + te^{-\gamma}\phi_B(j-1) + t'e^{-\gamma}\phi_A(j+1) + t'e^{-\gamma}\phi_A(j-1) + t'(2 - e^{-2\gamma})\phi_A(j), \quad (6.45)$$

$$E\phi_B(j) = -t\phi_A(j) + te^{-\gamma}\phi_A(j+1) + t'e^{-\gamma}\phi_B(j+1) + t'e^{-\gamma}\phi_B(j-1) + t'(2 - e^{-2\gamma})\phi_B(j), \quad (6.46)$$

where we used the identity $2 - e^{-2\gamma} = -2\cos(ka)$. At the boundary, these equations reduce to

$$E\phi_A(1) = -t\phi_B(1) + te^{-\gamma}\phi_B(0) + t'e^{-\gamma}\phi_A(2) + t'(2 - e^{-2\gamma})\phi_A(1), \quad (6.47)$$

and

$$E\phi_B(0) = te^{-\gamma}\phi_A(1) + t'e^{-\gamma}\phi_B(1) + t'(2 - e^{-2\gamma})\phi_B(0). \quad (6.48)$$

Unlike for the case of nearest-neighbor hopping, an analytical solution of these equations does not seem feasible. Thus, we will use perturbation theory to study the effects of next-nearest neighbor hopping, as in Ref. [183]. To this end, note that the edge state solution for $t' = 0$ (Eq. (6.24)),

$$\phi_A(j) = 0, \quad \phi_B(j) = \sqrt{1 - e^{-2\gamma}}e^{-\gamma j} \quad (6.49)$$

is almost a solution for $t' \neq 0$, too. In particular, it solves Eqs. (6.45), (6.46) and (6.47) for $E = 3t'$.⁷ However, it solves Eq. (6.48) only for $E = 2t'$, and thus is not an eigenstate. Therefore, apart from a trivial overall energy shift, the next-nearest neighbor hopping only influences the edge state close to the graphene edge. In fact, the previous analysis shows that for the edge state, next-nearest neighbor hopping corresponds to putting an additional on-site

potential $-t'$ on the edge atoms. However, since the edge state is localized at the edge, it is then influenced significantly by $t' \neq 0$.

In first order perturbation theory, the energy of the edge state is shifted by

$$\begin{aligned}\Delta E(k) &= \langle \varphi | \Delta H | \varphi \rangle / \langle \varphi | \varphi \rangle \\ &= \left(2t' + 3t' \sum_{j=1}^{\infty} e^{-2\gamma j} \right) (1 - e^{-2\gamma}) \\ &= \left(-t' + 3t' \frac{1}{1 - e^{-2\gamma}} \right) = 4t' + 2t' \cos(ka)\end{aligned}\tag{6.50}$$

Thus, the edge state energy is shifted by the bulk value $3t'$ at the K -points, but only by $2t'$ at $k = \pm \frac{\pi}{a}$ where the edge state is maximally localized. The previously perfectly flat edge state band hence acquires a dispersion given by $\Delta E(k)$.

Close to the K -points, the energy shift $\Delta E(k)$ can be expanded as

$$\Delta E(k) = 3t' \pm \sqrt{3}at' \left(k \pm \frac{2\pi}{3a} \right) + \mathcal{O} \left(\left(k \pm \frac{2\pi}{3a} \right)^2 \right)\tag{6.51}$$

i.e. the next-nearest neighbor hopping introduces corrections to the edge state energy dispersion linear in $q = k \pm \frac{2\pi}{3a}$. Note that we could neglect t' for bulk states in the lowenergy regime, since next-nearest neighbor hopping only introduced terms quadratic in $\mathbf{q}$, as discussed in Appendix G. However, since t' enters the edge state dispersion already to linear order in q , next-nearest neighbor hopping may not be neglected in this case.

We do not attempt to do perturbation theory for the zigzag nanoribbon, but we can extrapolate the results for a single edge to the case of two edges in a nanoribbon. Hence, we expect that the almost dispersionless bands, that we obtained in nearest-neighbor tight-binding, will also obtain a finite dispersion for $t' \neq 0$. This dispersion will compete with the energy splitting due to the finite width of the ribbon, Eq. (6.39). However, since this splitting decreases exponentially with the nanoribbon width, the widthindependent dispersion from next-nearest neighbor hopping will usually dominate.

We demonstrate these predictions in Fig. 6.8(a), where we show the band structure of a zigzag nanoribbon in next-nearest neighbor tight-binding approximation. Note that we have subtracted the trivial energy shift $3t'$ here, and we will use this convention in the remainder of this work, when considering next-nearest neighbor hopping. In the band structure, we observe two edge state bands (arising from the two edges of the nanoribbon) that now have a significant dispersion, in contrast to the almost flat bands for nearest-neighbor hopping only (see Fig.6.6(a)).

In Fig. 6.8(b) we show the probability density of the edge state across the nanoribbon for a given value of k . Since the ribbon has a reflection symmetry between the two edges, the Bloch eigenstate for a fixed value of the wave vector k still occupies both edges with equal probability. However, due to the dispersion of the edge state band, there is an energy range where in addition to the "bonding state" we also have an "anti-bonding" state at the same energy (and, additionally, an extended state). Thus,

¹²⁷ Note in this respect that the next-nearest neighbor hopping also shifts the bulk energy spectrum by $3t'$, as discussed in Section G.2.2.

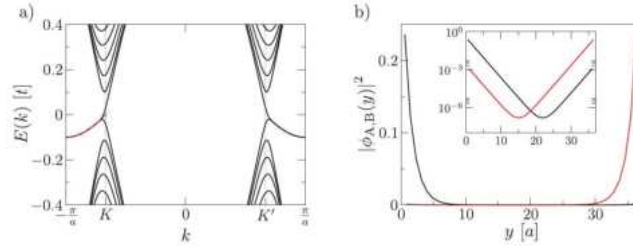


Figure 6.8 - A zigzag graphene nanoribbon in next-nearest neighbor tight-binding approximation for $t' = 0.1t$. (a) shows the band structure of the ribbon, where we have subtracted the trivial energy shift $3t'$. For comparison, the red dashed curve gives the energy dispersion of the edge state obtained from perturbation theory, Eq. (6.50). In (b) we show the probability density profile of the edge state across the nanoribbon, for $k = -2.4/a$. The inset shows the same density profile on a logarithmic scale. Again, the probability density for sublattice A is shown in red, for sublattice B in black. In both (a) and (b) the nanoribbon had a width $W = 64a/\sqrt{3}a$.

we can form linear combinations of these states to obtain eigenstates that are located at a single edge⁸. In addition to that, the next-nearest neighbor hopping also changes the character of the wave function, as seen in the inset of Fig. 6.8(b): Whereas the edge state resided on a single sublattice close to the boundary in the examples of the previous section, $t' \neq 0$ leads to an increased (but still small) occupation probability on the other sublattice. This has important consequences, that are discussed in the next section.

We conclude the characterization of the graphene edge state by discussing how to include the effects of next-nearest neighbor hopping in the effective Hamiltonian. Note that we have argued that the influence of the next-nearest neighbor hopping on the edge state is equivalent to a potential $-t'$ at the zigzag edge atoms. In the effective Hamiltonian we can include such a potential as

$$V_{t'}(y) = 3t' - \frac{\sqrt{3}a}{2}t'\delta(y - y_{\text{edge}}) \quad (6.52)$$

where y_{edge} is the y -coordinate of the boundary. Note that $\sqrt{3}a/2$ is the length of the graphene unit cell in y -direction. Such a potential will influence strongly a state localized at the boundary, but only little a state extended through the whole system. For a single zigzag edge, we obtain the energy shift in first order perturbation theory within the effective Hamiltonian as

$$\Delta E(q) = -2\tau q \int_0^\infty e^{-2\tau q y} V_{t'}(y) dy = 3t' - \sqrt{3}at'\tau q \quad (6.53)$$

in agreement with Eq. (6.51).

19.5.10 6.4 Edge state transport in zigzag nanoribbons

19.5.11 6.4.1 Model-dependence of edge state transport-some peculiar examples

As mentioned in the introduction, numerical studies of edge state transport in zigzag nanoribbons have, up to now, only considered graphene in nearest-neighbor tightbinding ap-

¹²⁸ Note that the "anti-bonding" state has a slightly different wave vector than the "bonding state", and thus a slightly different shape of the wave function. This difference in the wave vectors is due to the energy splitting because of the finite width of the nanoribbon. However, since this energy splitting decreases exponentially with the ribbon width, so does the difference in the wave vector. Thus, the linear combination of "bonding" and "anti-bonding" state is to a good approximation localized at a single edge only.

proximation [177-180]. The main finding of these accounts is that edge state transport is influenced only little by edge defects. This is a rather counterintuitive result: Since the probability density of the edge state is strongly localized to the edges of the ribbon, one would naively expect it to scatter strongly from defects at the edge. In fact, edge state transport holds even more surprises, as we will demonstrate on illustrative examples in this section.

For this, we use the numerical framework introduced in the first part of this work to calculate transmission probabilities and density as well as current profiles of zigzag nanoribbons. However, before presenting the calculations, it is worthwhile to spare some thoughts about general aspects of numerical studies on edge state transport.

Usually, numerical transport studies of nanoribbons compute the transmission probability as a function of Fermi energy E_F . The range of energies considered is commonly on the order of t , implying that the almost flat edge state band in nearest-neighbor approximation only contributes a few- or often only a single - point in the numerically computed transmission spectrum. Computing the transmission as a function of Fermi energy is based on the idea of mimicking the effect of a gate voltage. In an ordinary two-dimensional electron gas in a semiconductor heterostructure, the gate voltage V_{gate} and the Fermi energy E_F are in fact linearly proportional, since a capacitively coupled gate induces a charge $Q = CV_{\text{gate}}$ (C is the capacitance of the structure) and the density of states is a constant. This is not the case anymore in bulk graphene, where the density of states is linear in energy. Even worse in nanoribbons, the edge states in nearest-neighbor tight-binding have a very high density of states due to the almost flat edge state band. Hence, changing the Fermi energy linearly will grossly underestimate the gate voltage region where the edge states play a dominant role in transport. Note that this problem is less severe, if the edge states acquire a dispersion, such as due

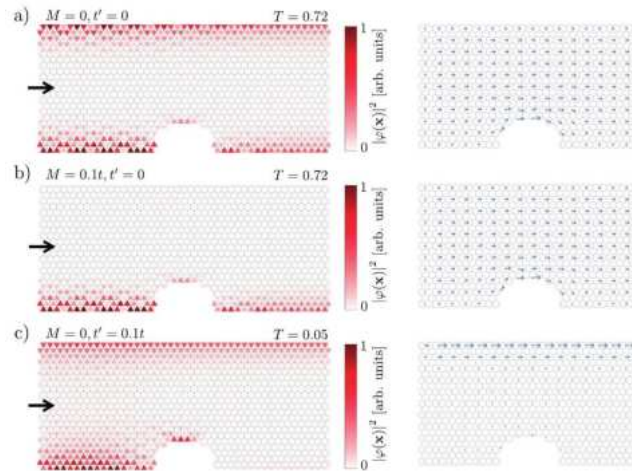


Figure 6.9 - Density profile (left panels) and current density profile (right panels) of the edge state ($k = -2.4/a$) in a zigzag nanoribbon (width $W = 34a/\sqrt{3}$) with a large obstacle on the lower edge. (a) shows the edge state in nearest-neighbor approximation ($t' = 0$ and $M = 0$), (b) for a finite staggered potential ($t' = 0$ and $M = 0.1t$), and (c) in next-nearest neighbor approximation ($t' = 0.1t$, $M = 0$). In all examples, the channels are incoming from the left. to next-nearest neighbor hopping or other perturbations. Furthermore, we have seen in the previous sections that the extent of the edge state wave function is determined by the Bloch vector k , and is largely independent of the particular model employed to describe graphene. Thus, instead of Fermi energy, we will use k in order to characterize the edge state in this and the following sections.

In order to demonstrate the peculiarities of edge state transport, we consider a zigzag nanoribbon with a rather large edge defect, as shown in Fig. 6.9. The wave vector $k = -2.4/a$ of the edge state was chosen such that the extent of the edge state in the perpendicular direction is smaller than the size of the defect.

We begin by considering the case of graphene in nearest-neighbor approximation. Even though the extent of the probability density at one of the edges is smaller than the defect itself, as shown in Fig. 6.9(a), the wave function also comprises the density at the opposite edge, and in this sense extends through the whole width of the ribbon. Since the edge state in this situation is situated on both edges at the same time, it is hard to predict a priori the effect of a disturbance on a single edge only. Still, intuition would suggest that an edge defect should result in strong scattering. Yet, we find a rather large transmission probability of $T = 0.72$. Furthermore, if we made the defect smaller, we would find almost perfect transmission, if only a single lattice site was removed [177-179].

In order to explain this seeming contradiction between a state localized at the edge and its insensitiveness to edge disorder, Refs. [177] and [179] considered the current density of the edge state. Here, we will follow their arguments using the effective Hamiltonian of graphene.

Within the effective Dirac Hamiltonian, the velocity operator is given as

$$\mathbf{v} = v_F \begin{pmatrix} \tau \sigma^x \\ \sigma^y \end{pmatrix} \quad (6.54)$$

Hence, we find the current density of a state φ as

$$\mathbf{j}(\mathbf{x}) = e\varphi^\dagger(\mathbf{x})\mathbf{v}\varphi(\mathbf{x}) \quad (6.55)$$

From the edge state solution (6.35) we find the current density in ribbon direction as

$$\begin{aligned} j_x(y) &\cong \phi_A(y)\phi_B(y) \\ &= \frac{1}{2}(\cosh(zW) - \cosh(2z(W/2 - y))). \end{aligned} \quad (6.56)$$

In Fig. 6.10 we show both the probability and current density profiles of the edge state. Note that the current densities have been normalized such that the integrated current carries one quantum of conductance⁹. The current density always has a maximum in the middle of the ribbon, whereas it vanishes at the edges, where the wave function of one of the sublattices must be zero, as seen from the zigzag boundary conditions (6.34). When the edge state decays slowly and extends far into the ribbon, the current density profile is nearly sinusoidal (more precisely, it is parabolic for $z = 0$), and is almost constant throughout the whole ribbon, if the edge state is strongly localized to the edge.

This behavior may be understood from the lattice structure of graphene: In nearest neighbor tight-binding approximation, current flow requires hopping from one sublattice to the other. Since the edge state close to a single edge is localized on a single

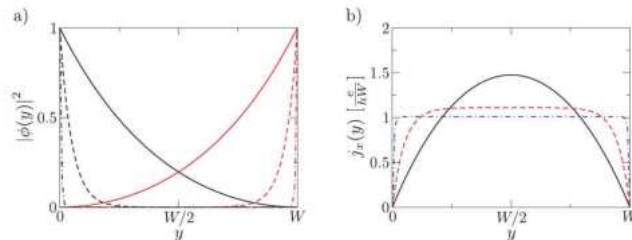


Figure 6.10 - (a) Probability density and (b) current profile of the edge state as a function of the perpendicular position y in a zigzag nanoribbon of width W . (a) and (b) show the behavior

¹²⁹ The rationale behind this normalization is that in ballistic transport every channel contributes $\frac{e^2}{h}$ to the conductance: The current density depends on the velocity of the channel, but this dependence is cancelled exactly by the density of states. In this normalization, the current density $j_x(y)$ of the state becomes essentially the energy-dependent current density $j_x(y, E)$ introduced in Appendix A.

for different decay constants of the edge state: $zW = 1$ (solid lines), $zW = 10$ (dashed lines), and $zW = 100$ (dash-dotted lines). In (a), sublattice A (B) is encoded in red (black), and the wave functions have been normalized to unity at the edges.

sublattice only, current is suppressed at the edges, and reaches a maximum in the middle, where the tails of the exponential decay overlap¹⁰. The fact that current is flowing mainly in the middle of the ribbon explains the fact that edge defects have only little influence of the edge state. As seen in Fig. 6.9(a) the current density merely meanders around the obstacle. Furthermore, the value of the transmission, $T \approx 3/4$ can be understood from the ratio of the extent of the obstacle and the ribbon width which is around $1/4$. In fact, we would still find a sizeable transmission, if we introduced defects at both edges.

Thus, although the edge state density is strongly localized at the edges, its current density rather behaves as expected from a bulk state. Note that the current density is zero for all states at the edges of a zigzag nanoribbon in nearest-neighbor approximation, since this is dictated by the boundary conditions (6.34). Thus, we also find extended states, where the density profile differs from the current profile. However, the difference is most drastic for the edge state, where the density at the edges is exponentially larger than in the middle.

We next turn to the case of a finite staggered potential $M = 0.1t$, depicted in Fig. 6.9(b). In contrast to the previously considered case $M = 0$, the edge state is now localized to a single edge only. Here, we have chosen the edge state at $E \approx -M$ which is localized at the same edge as the defect. Even more than in the previous case, intuition

would suggest that the state should be scattered back with a large probability. Yet, we find that the channel is transmitted again with a very high transmission probability $T = 0.72$, which is identical to the case $M = 0$. This can be understood from the shape of the wave function (6.41), which is, up to different weights for sublattice A and B, identical to the wave function for $M = 0$, Eq. (6.35). Thus, the current density profiles are equal in both cases, as can be seen in Fig. 6.9(b). Even the pattern of the reflected wave is identical in both situations. In fact, in this situation the edge state would be scattered as much from a defect on the opposite edge. Hence, we find the seemingly paradoxical situation that a state localized at a single edge, and thus being most asymmetric with regard to both edges, features a perfectly symmetric current density, and thus is affected equally by defects on either edge.

Finally, we consider the case of finite next-nearest neighbor hopping $t' = 0.1t$, which also allows for edge state solutions located at a single edge, just as $M \neq 0$. In this situation we find that the edge state located at the defective edge is scattered back almost perfectly ($T = 0.05$), whereas the state on the opposite edge can pass unaffected (see Fig. 6.9(c)). In addition, the current density is mainly located at the edge, in accordance with the probability density. Thus, the transport behavior of the edge state for $t' \neq 0$ differs fundamentally from the previous examples, although the probability densities look much alike.

As discussed above, the edge state in a ribbon with nearest-neighbor hopping only is comprised of contributions from both edges, because of the particle-hole symmetry associated with $\mathcal{P}_z$. Because of this symmetry, both edges are inevitably coupled. From this point of view, the counter-intuitive transport properties of the edge state, namely the insensitiveness to edge disorder, may not seem too surprising. Yet, even after breaking this particle symmetry with a mass term, localizing the edge state to a single edge only, the transport properties of the edge state remain unchanged. In contrast, next-nearest neighbor hopping, which also breaks $\mathcal{P}_z$, changes the transport properties fundamentally. It is thus not enough to break the particle-hole symmetry $\mathcal{P}_z$ alone to change edge state transport. In the examples of this section, the shape of the current density was most useful to understand the peculiar transport features.

¹²¹⁰ The maximum of the current density is in the middle, where the contributions from both sublattices are already very small, due to the exponential decay. The resulting velocity (and also the total current) of the state is thus exponentially small, which is reflected in the almost dispersionless edge state band.

Thus, in the next section, we will examine the shape of the current density systematically.

19.5.12 6.4.2 Where does the current flow?

In order to investigate the behavior of the current density profile of the edge state in dependence of additional terms to the graphene Hamiltonian in nearest-neighbor tight-binding approximation, we consider the following models of a zigzag nanoribbon:

- I. nearest-neighbor hopping only.
- II. nearest-neighbor hopping with a constant staggered potential M .
- III. finite next-nearest neighbor hopping t' ,
- IV. nearest-neighbor hopping with a position-dependent staggered potential $M(\mathbf{x})$.

$$M(\mathbf{x}) = \begin{cases} M_0 e^{-\bar{\gamma}y} & \text{for } y < W/2 \\ M_0 e^{-\bar{\gamma}(W-y)} & \text{for } y > W/2 \end{cases} \quad (6.57)$$

Thus, the magnitude of the staggered potential decays exponentially away from the edges. Such a model may be appropriate for describing edge magnetism (see Chapter 7).

- V. nearest-neighbor hopping with a position-dependent potential $V(\mathbf{x})$:

$$V(\mathbf{x}) = V_0 \left(\frac{1}{2} - \frac{y}{W} \right). \quad (6.58)$$

This is a potential that changes linear from $-V_0/2$ at one edge to $V_0/2$ at the other edge. This may model the effects of side gates (see Ref. [176]).

Note that models II-V break the particle-hole symmetry associated with $\mathcal{P}_z$ and thus allow the localization of the edge state on a single edge only.

In Fig. 6.11 we show a comparative overview of the band structure and the edge state probability as well as current density profile for all five models (for a discussion of the respective observables, see Appendix A).

Fig. 6.11(a) and (b) show a almost constant current density profile with a maximum in the middle, as discussed in the previous section, whereas in (c)-(e), the current density profile has distinct maxima at the edges, in accordance with the probability density. In these calculations, we chose the Fermi energy such that we have an edge state with a fixed value of $k = -2.4/a$. Due to the band-bending in models III-V there are three channels at the Fermi energy in Fig. 6.11(c) (two edge states and one extended state) and two channels in (d) and (e) (one edge state and one extended state). The finite current density in the middle of the nanoribbon in (c)-(e) is due to the extended state. In contrast to the current density that differs fundamentally between the different models, the probability density of the edge state wave function is almost identical for all models - only on a logarithmic scale it is possible to observe differences.

When we observe a difference in the current density profile for the various models, this can be either due to a change in the wave function, or due to a change in the current density operator. The current density operator is identical for models I, II, IV and V, since all these comprise only nearest-neighbor hopping¹¹. The difference in these models thus must be due to changes in the wave function. In contrast, nextnearest neighbor hopping (model III) also introduces changes to the current density operator. However, it turns out that also in this case the change in the wave function is the dominant source for altering the current density profile. This can be seen by

¹²¹¹ The current density operator in tight-binding only depends on hopping elements, not on the on-site energies (see Appendix A).

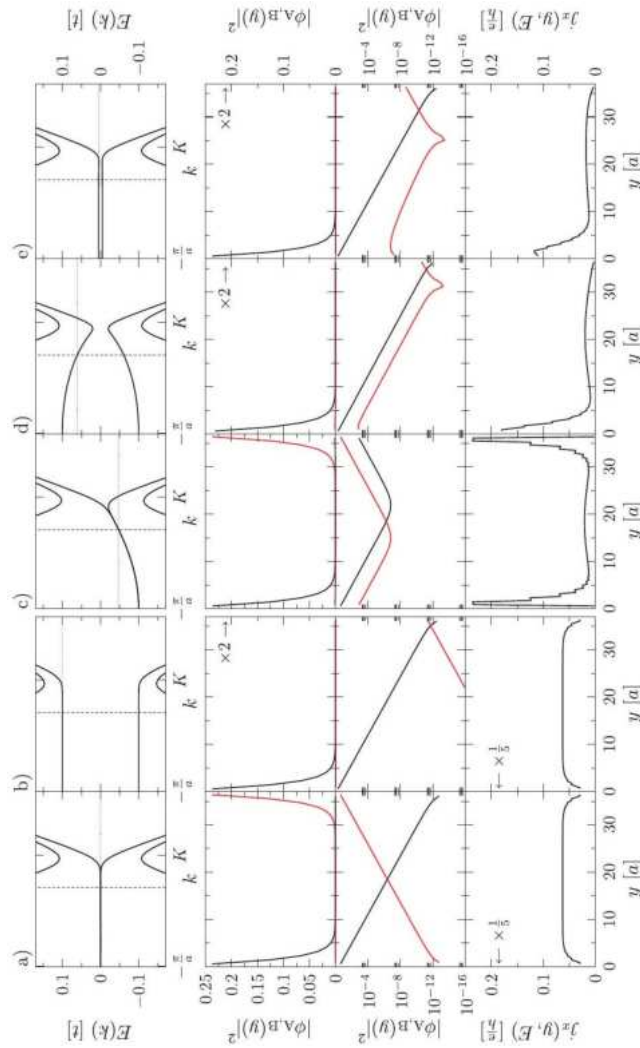


Figure 6.11 - Overview over the edge state behavior for different models of a zigzag nanoribbon (width $W = 64a/\sqrt{3}$): (a) nearest neighbor hopping only (model I), (b) a constant staggered potential with $M = -0.1t$ (model II), (c) next-nearest neighbor hopping $t' = 0.1t$ (model III), (d) a position-dependent staggered potential $M(\mathbf{x})$ with $M_0 = -0.1t$ and $\bar{\gamma} = 1/a$ (model IV), and (e) a position-dependent potential $V(\mathbf{x})$ with $V_0 = 0.01t$ (model V). The upper panels show the band structure in the vicinity of the edge state. We consider edge states with wave vector $k = -2.4/a$; in the plot of the band structure, this k -value is indicated as a dashed line and the corresponding Fermi energy as a dotted line. The second and third row of panels show the probability density profile of the edge state on a linear and logarithmic scale, respectively, whereas the lower panels show the current density profiles for all states at the Fermi energy.

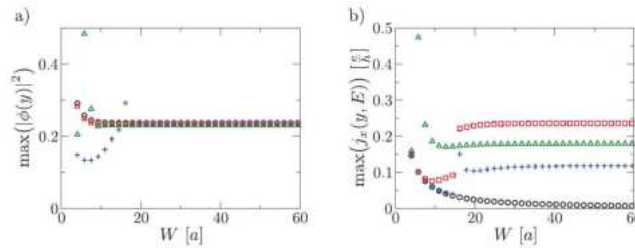


Figure 6.12 - Maximum value of the (a) probability and (b) current density of the edge state ($k = -2.4/a$) as a function of the ribbon width W , for nearest-neighbor hopping only ($\circ$), next-nearest neighbor hopping ($\square$), a position-dependent staggered potential ($\triangle$), and a position-dependent potential ($+$). Note that we have divided the probability density of the models IV and V by 2, since there the edge state is localized on a single edge only.

comparing the current densities computed on the one hand with a current operator including t' and, on the other hand, with a current operator where t' is set to zero artificially.

In Fig. 6.11(a) and (b) we observe that the occupation probability of a sublattice decreases monotonically from one edge to the other. This is not the case for (c)-(e): There, the probability density increases again towards the opposite edge, so that for every edge state that occupied a single sublattice only, there is an enhanced occupation probability of the opposite sublattice. Furthermore, the probability densities show a similar functional behavior in the vicinity of the edges, and seem in a sense "linked" to each other.

Before pursuing these observations further, we briefly revisit the statement of the previous section that the current density of the edge state in nearest-neighbor approximation behaves like a bulk state. A bulk state extends through the whole system. In particular this means that the occupation probability for a single lattice site scales inversely proportional with the width of the nanoribbon. In contrast, a true edge state remains unchanged when the width of the nanoribbon is varied. In Fig. 6.12(a) we show the maximum value of the probability density for the models I and III-V (As discussed in the previous section, model II behaves just like I in this regard, and is thus not shown explicitly). Except for small widths, where finite size effects play a role, all models show a constant maximum of the probability density, independent of the ribbon width. This is the signature of an edge state. In contrast, only models III-V show a constant maximum of the current density in Fig. 6.12(b) whereas the maximum of the current density in model I decreases as the inverse of the nanoribbon width. Thus, again the current density of the edge state in nearest-neighbor tight-binding approximation shows bulk-state-like behavior, whereas models III-V behave like a true edge state in every sense.

The enhanced occupation probability of the opposite sublattice and a current flowing at the edges are necessarily correlated: In order to obtain current flowing close to the edge, the value of the wave function on the sublattice that was close to zero before must be increased, as the current density is given by the product of the wave functions on both sublattices, Eq. (6.56). Furthermore, it seems that a maximum in the current density at the edge is always linked to a finite bending of the edge state band, in particular a bending beyond the exponentially small energy splitting due to the finite width. In the next section, we will further elaborate on these connections.

19.5.13 6.4.3 Effects of perturbations on the current flow

In order to better understand, under which circumstances the current density changes from bulk-like to edge-like, we study the corrections to the edge state of a single edge in perturbation theory within the effective Dirac Hamiltonian. The results will then be interpolated to the case of a nanoribbon.

To this end we consider again a half-infinite graphene sheet with a single zigzag edge as in Fig. 6.2(b), governed by the Dirac Hamiltonian H_τ , Eq. (6.6), including a perturbation:

$$H = H_\tau + V^\mu \quad (6.59)$$

where

$$V^\mu = g(\partial_x, \partial_y, y) \sigma^\mu \quad (6.60)$$

Here, $\mu = 0, x, y, z$, and $g(\partial_x, \partial_y, y)$ is an function that may contain arbitrary derivatives, but only depends on the coordinate y perpendicular to the zigzag edge. This perturbation potential includes examples such as a magnetic vector potential ($V^\mu \sim \sigma^x, \sigma^y$), a staggered potential ($V^\mu \sim \sigma^z$), a potential or next-nearest neighbor hopping ($V^\mu \sim \sigma^0$). The system then

still exhibits translational symmetry in the x -direction, and the solution has the Bloch form (6.28). As a consequence, Eq. (6.59) turns into a one-dimensional differential equation for $\phi(y)$:

$$(H_\tau(q) + V^\mu(q)) \phi(y) = E \phi(y), \quad (6.61)$$

where

$$H_\tau(q) = \hbar v_F (\tau q \sigma^x + (-i \partial_y) \sigma^y) \quad (6.62)$$

is the unperturbed Dirac Hamiltonian and

$$V^\mu(q) = g (iq, \partial_y, y) \sigma^\mu \quad (6.63)$$

the perturbation for a Bloch vector q . Within this subspace, we can then apply nondegenerate perturbation theory (for example, see Ref. [153, Chap. 11]).

The unperturbed Hamiltonian $H_0(q)$ together with the zigzag boundary condition $\phi_A(0) = 0$ has the already known edge state solution, Eq. (6.24),

$$\phi^{\text{edge}}(y) = \sqrt{-2\tau q} e^{\tau q y} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{with energy } E^{\text{edge}} = 0 \quad (6.64)$$

as well as the bulk solutions

$$\phi^\pm(z, y) = \frac{1}{\sqrt{\pi}} \begin{pmatrix} \sin(zy) \\ \pm \sin(zy + \theta) \end{pmatrix} \quad \text{with energy } E^\pm(q, z) = \pm \hbar v_F \sqrt{q^2 + z^2} \quad (6.65)$$

where the angle θ is determined by the equation

$$e^{i\theta} = \frac{\tau q + iz}{\sqrt{q^2 + z^2}} = \frac{\sqrt{q^2 + z^2}}{\tau q - iz} \quad (6.66)$$

and z is a positive real number. The eigenstates have been normalized such that the completeness relation

$$\phi^{\text{edge}}(y) (\phi^{\text{edge}}(y'))^\dagger + \sum_{\tau=\pm} \int_0^\infty dz \phi^\tau(z, y) (\phi^\tau(z, y'))^\dagger = \mathbb{1} \delta(y - y') \quad (6.67)$$

holds.

Then, the energy shift $\Delta E(q)$ in first order perturbation theory is given as

$$\begin{aligned} \Delta E(q) &= \langle \phi^{\text{edge}} | V^\mu | \phi^{\text{edge}} \rangle \\ &= -2\tau q \int_0^\infty dy' e^{\tau q y'} \begin{pmatrix} 0 \\ 1 \end{pmatrix}^\dagger \sigma^\mu \begin{pmatrix} 0 \\ 1 \end{pmatrix} g(iq, \partial_{y'}, y') e^{\tau q y'}. \end{aligned} \quad (6.68)$$

The expression for $\Delta E(q)$ contains an expectation value of a Pauli matrix σ^μ with an eigenstate of σ^z . Hence, we find that the energy shift $\Delta E(q) = 0$ for $V^\mu \sim \sigma^x, \sigma^y$. This is not surprising, as such a perturbation does not change the off-diagonal form of the Hamiltonian. Thus, the differential equations for the two sublattices still decouple for $E = 0$ and there will be a zero-energy solution. On the contrary, if $V^\mu \sim \sigma^0, \sigma^z$, we find a finite (q -dependent) energy shift

$$\Delta E(q) = \mp 2\tau q \int_0^\infty dy' e^{2\tau q y'} g(iq, q, y') \quad (6.69)$$

where the upper sign refers to $V^\mu \sim \sigma^0$ and the lower sign to $V^\mu \sim \sigma^z$. Thus, any perturbation $V^\mu \sim \sigma^0, \sigma^z$ will in general lead to band bending.

The first-order correction to the wave function is given as

$$\Delta\phi(y) = \sum_{\tau=\pm} \int_0^\infty dz \frac{\langle \phi^\tau(z) | V^\mu | \phi^{\text{edge}} \rangle}{0 - E^\tau(q, z)} \phi^\tau(z, y) \quad (6.70)$$

which is evaluated to

$$\begin{aligned} \Delta\phi_A(y) &= \frac{1}{\sqrt{\pi}} \int_0^\infty dz \frac{1}{E^+(q, z)} (\langle \phi^-(z) | V^\mu | \phi^{\text{edge}} \rangle - \\ &\quad \langle \phi^+(z) | V^\mu | \phi^{\text{edge}} \rangle) \sin(z y) \\ \Delta\phi_B(y) &= -\frac{1}{\sqrt{\pi}} \int_0^\infty dz \frac{1}{E^+(q, z)} (\langle \phi^+(z) | V^\mu | \phi^{\text{edge}} \rangle + \\ &\quad \langle \phi^-(z) | V^\mu | \phi^{\text{edge}} \rangle) \sin(z y + \theta), \end{aligned} \quad (6.71b)$$

where we made use of the relation $E^+(q, z) = -E^-(q, z)$. These equations can be simplified further by the observation that

$$\begin{aligned} \langle \phi^+(z) | V^\mu | \phi^{\text{edge}} \rangle &= \sqrt{\frac{-2\tau q}{\pi}} \int_0^\infty dy' \left(\frac{\sin(z y')}{\sin(z y' + \theta)} \right)^\dagger \sigma^\mu \begin{pmatrix} 0 \\ 1 \end{pmatrix} g(iq, \partial_{y'}, y') e^{-\tau q y'} \\ &= \begin{cases} + \langle \phi^-(z) | V^\mu | \phi^{\text{edge}} \rangle & \text{for } V^\mu \sim \sigma^x, \sigma^y, \\ - \langle \phi^-(z) | V^\mu | \phi^{\text{edge}} \rangle & \text{for } V^\mu \sim \sigma^0, \sigma^z. \end{cases} \end{aligned} \quad (6.72)$$

Inserting these expressions into Eqs. (6.71) we find that $\Delta\phi_A(y) = 0$ and $\Delta\phi_B(y) \neq 0$ for $V^\mu \sim \sigma^x, \sigma^y$, i.e. the edge state remains localized on a single sublattice. Again, this is not surprising, since σ^x and σ^y do not change the off-diagonal structure of the Dirac Hamiltonian, and the corresponding zero-energy solutions are located on a single sublattice only. In particular, this applies to the case of a weak magnetic field, where the magnetic vector potential creates a perturbation $\sim \sigma^x, \sigma^y$.

In contrast, if $V^\mu \sim \sigma^0, \sigma^z$, we find

$$\Delta\phi_A(y) = -\frac{2}{\sqrt{\pi}} \int_0^\infty dz \frac{1}{E^+(q, z)} \langle \phi^+(z) | V^\mu | \phi^{\text{edge}} \rangle \sin(z y + \theta), \quad (6.73a)$$

$$\Delta\phi_B(y) = 0. \quad (6.73b)$$

Thus, terms proportional to σ^0 and σ^z will in general introduce a finite occupation probability on the sublattice that was strictly zero in the unperturbed Hamiltonian.

We now turn to evaluating Eq. (6.73a):

$$\begin{aligned} \Delta\phi_A(y) &= -\frac{2}{\pi} \sqrt{-2\tau q} \int_0^\infty dz \int_0^\infty dy' \frac{1}{\hbar v_F \sqrt{q^2 + z^2}} \\ &\quad \left(\frac{\sin(z y')}{\sin(z y' + \theta)} \right)^\dagger \sigma^{0,z} \begin{pmatrix} 0 \\ 1 \end{pmatrix} g(iq, \partial_{y'}, y') e^{-\tau q y'} \sin(z y) \\ &= \pm \frac{\sqrt{-2\tau q}}{2\pi \hbar v_F} \int_0^\infty dy' g(iq, q, y') e^{\tau q y'} \int_{-\infty}^\infty dz \left(\frac{e^{iz(y'+y)}}{\tau q - iz} - \frac{e^{-iz(y'-y)}}{\tau q + iz} \right), \end{aligned} \quad (6.74)$$

where the upper sign refers to σ^0 and the lower sign to σ^z . In order to arrive at this expression, we used $\sin(x) = \frac{1}{2i}(e^{ix} - e^{-ix})$ and Eq. (6.66). We can simplify this expression further by making use of the identities

$$\int_{-\infty}^{\infty} dz e^{-izy} \frac{1}{\tau q + iz} = \int_{-\infty}^{\infty} dz e^{izy} \frac{1}{\tau q - iz} = -2\pi\Theta(y)e^{\tau q y} \quad (6.75)$$

where $\Theta(y)$ is the Heaviside step function. Eq. (6.75) can be easily obtained by contour integration. We finally obtain

$$\Delta\phi_A(y) = \mp \frac{\sqrt{-2\tau q}}{\hbar v_F} \left(e^{\tau q y} \int_0^{\infty} dy' g(iq, q, y') e^{2\tau q y'} - e^{-\tau q y} \int_y^{\infty} dy' g(iq, q, y') e^{2\tau q y'} \right) \quad (6.76)$$

where again the upper sign is for $V^\mu \sim \sigma^0$ and the lower sign for $V^\mu \sim \sigma^z$. In order to get a feel for the nature of the perturbed wave function we now evaluate Eq. (6.76) for the models of the previous section:

- $V^\mu = V_0\sigma^0$ or $V^\mu = M\sigma^z$: For a constant potential or staggered potential we obtain

$$\Delta\phi_A(y) = 0 \quad (6.77)$$

This is obvious for the constant potential, as this only leads to an overall energy shift. For the staggered potential this is in agreement with the findings of Section 6.3.3, where the occupation probability on sublattice A originated solely due to the second edge in the ribbon.

- $t' \neq 0, V^\mu = V_{t'}(y)\sigma^0 = 3t' - \frac{\sqrt{3}a}{2}t'\delta(y - y_{\text{edge}})$: For finite next-nearest neighbor hopping we obtain

$$\Delta\phi_A(y) = \frac{t'}{t} \sqrt{-2\tau q} e^{\tau q y} = \frac{t'}{t} \phi_B(y). \quad (6.78)$$

- $V^\mu = M_0 e^{-\bar{\gamma} y} \sigma^z$: For a position-dependent staggered potential we obtain

$$\Delta\phi_A(y) = \frac{M(e^{-\bar{\gamma} y} - 1)}{\hbar v_F(2\tau q - \bar{\gamma})} \phi_B(y) \quad (6.79)$$

- $V^\mu = V_0/W(W/2 - y)\sigma^0$: For a position-dependent potential we obtain

$$\Delta\phi_A(y) = \frac{V_0 y}{\hbar v_F W(-2\tau q)} \phi_B(y) \quad (6.80)$$

Hence, we find that the wave function of sublattice A always is of the form $f(y)\phi_B(y)$, where $f(y)$ is some function of y . In fact, this form is rather general, as can be seen from Eq. (6.76): Provided that the perturbing function $g(iq, q, y)$ is sufficiently benign-and only for such functions perturbation theory is applicable - partial integration of the second term in Eq.

(6.76) will always lead to terms proportional to $\phi_B(y) = \sqrt{-2\tau q}e^{\tau q y}$.

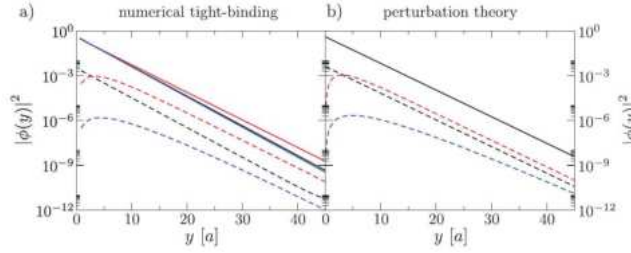


Figure 6.13 - Comparison of edge state probability densities for (a) the full tight-binding model and (b) for a perturbative treatment of the Dirac Hamiltonian. In order to simulate a half-infinite graphene sheet in the numerics, we only show the region close to one of the edges of a wide nanoribbon (width $W = 154a/\sqrt{3}$). The edge state ($k = -2.4/a$) is shown for the case of next-nearest neighbor hopping with $t = 0.1t$ (black), a position-dependent staggered potential with $M_0 = -0.1t$ and $\bar{\gamma} = 0.3/a$ (red) and a position-dependent potential with $V_0 = 0.4$ (blue). The wave function on sublattice B (A) is shown as a solid (dashed) line.

Thus, the wave functions on sublattice A and B in the vicinity of the edge are indeed "linked", as observed in the numerical results of the previous section.

In Fig. 6.13 we show a comparison of the probability density of the edge state wave function at a zigzag edge from a numerical calculation within the tight-binding model (Fig. 6.13(a)) and from Eqs (6.78), (6.79), and (6.80) obtained perturbatively within the effective Dirac Hamiltonian. Apart from the fact that the Dirac Hamiltonian in general slightly underestimates the strength of the edge state decay, we find excellent agreement between the results from the full tight-binding model and the perturbation theory of this section.

Hence, next-nearest-neighbor hopping and arbitrary position-dependent potentials and staggered potentials lead to a bending of the edge state band and an increased occupation probability on the sublattice that is not occupied without these corrections to the Hamiltonian. Although these calculations have only been performed for a single edge, we can extrapolate the results to the case of a nanoribbon of finite width W . For simplicity, we will first concentrate on the case of next-nearest neighbor hopping.

If $-\tau q W \gg 1$, i.e. when the edge state is strongly localized towards the edges of the nanoribbon, the wave function of the edge state including the effects of a finite t' can be written to a good approximation as

$$\phi(y) \cong \begin{pmatrix} \sinh(zy) + \frac{t'}{t} \sinh(z(W-y)) \\ \sinh(z(W-y)) + \frac{t'}{t} \sinh(zy) \end{pmatrix}. \quad (6.81)$$

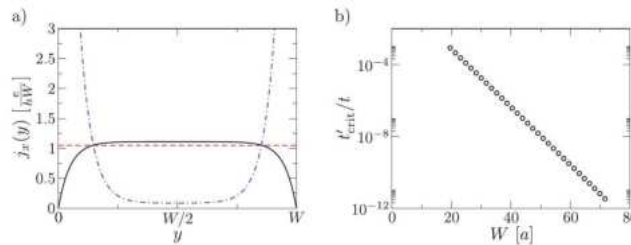


Figure 6.14 - (a) Current density $j_x(y)$ from Eq. (6.82) for $zW = 10$ and different values of t' : $t' \ll t'_{\text{crit}}$ (black solid line), $t' = t'_{\text{crit}}$ (red dashed line), and $t' \gg t'_{\text{crit}}$ (blue dashdotted line). (b) Numerical calculation of t'_{crit}/t as a function of nanoribbon width W for $k = -2.4/a$.

Here we extrapolated Eq. (6.78) in order to formulate Eq. (6.35) for $t' > 0$: Close to the edges, the wave functions of sublattice A and B are proportional to each other, with a proportionality factor t'/t . From this wave function we can calculate the current density profile as

$$\begin{aligned}
j_x(y) &\cong \frac{1}{2}(\cosh(zW) - \cosh(2z(W/2 - y))) \\
&\quad + \frac{t'}{t} ((\sinh(zy))^2 + (\sinh(zy))^2) + \mathcal{O}((t'/t)^2) \\
&= j_{x, \text{ n.n. }}(y) + \frac{t'}{t} (|\phi_A(y)|^2 + |\phi_B(y)|^2) + \mathcal{O}((t'/t)^2)
\end{aligned} \tag{6.82}$$

where $j_{x, \text{ n.n. }}(y)$ is the current density in nearest-neighbor approximation as given in Eq. (6.56). Thus, in addition to the current density in nearest-neighbor approximation with a maximum in the middle of the ribbon, the current density for $t' \neq 0$ contains an additional term that is proportional to the edge state probability density, located at the edges of the ribbon. In order to quantify which term will dominate the current density, we define the critical next-nearest neighbor hopping t'_{crit} such that

$$j_{x, \text{ nn}}(W/2) = \frac{t'_{\text{crit}}}{t} (|\phi_A(0)|^2 + |\phi_B(0)|^2) = \frac{t'_{\text{crit}}}{t} (|\phi_A(W)|^2 + |\phi_B(W)|^2), \tag{6.83}$$

i.e. the current flow at the edges equals the current flow in the middle. Then, for $t' \ll t'_{\text{crit}}$, current flow will be mainly through the middle of the nanoribbon, whereas for $t' \gg t'_{\text{crit}}$ it will be dominated by the contributions from the edge. Fig. 6.14(a) shows these different behaviors.

We can solve Eq. (6.83) for t'_{crit}/t (neglecting terms of order $(t'_{\text{crit}}/t)^2$ and higher and obtain

$$\frac{t'_{\text{crit}}}{t} = \frac{1}{2} \frac{\cosh(zW) - 1}{\cosh(zW)^2 - 1} \tag{6.84}$$

For $zW \gg 1$ this expression simplifies to

$$\frac{t'_{\text{crit}}}{t} \approx \frac{1}{2} e^{-zW} = \frac{1}{2} e^{\tau q W} \tag{6.85}$$

Thus, it is enough to have a next-nearest neighbor hopping that is exponentially small with the width of the nanoribbon in order to change the current density profile fundamentally. In Fig. 6.14(b) we show the value for t'_{crit} within the tight-binding model, obtained numerically by comparing current densities at the edge and in the middle of the nanoribbon for different values of t' . Indeed, we find that t'_{crit} decreases exponentially with the width of the nanoribbon, as predicted from the effective Dirac Hamiltonian.

Note that the exact value of t' for graphene is not known: In the literature one can find values $t' \approx 0.02t' - 0.2t'$ [35, 183, 184, 190]. In this respect, it is reassuring that the critical value of t' decreases exponentially with the ribbon width: For all but the smallest ribbons, all reported values of t' will lead to the same transport properties for the edge state.

In the case of next-nearest neighbor hopping the extrapolation to the nanoribbon was particularly easy, as there the wave functions on sublattice A and B were identical up to a prefactor. This is more complicated for the other examples of this section, where there are other functions multiplying the exponential decay (Eqs. (6.79) and (6.80)). Still, eventually the exponential decay dominates as seen in Fig. 6.13, and we expect a similar behavior as for next-nearest neighbor hopping. In particular, we predict that an exponentially small perturbation will again be enough to change the edge state transport properties fundamentally.

19.5.14 6.4.4 Scattering from edge defects

In the previous sections we have shown that the current density profile of zigzag nanoribbons can change drastically upon certain additional terms in the nearest-neighbor tight-binding model. However, the scattering properties of the edge state are influenced not only by the

current density¹² : For example, the availability of states that the edge state can scatter to is of importance. Thus, here we will demonstrate that not only the current density changes fundamentally, but also the scattering properties.

For this we consider a zigzag nanoribbon, where a single edge atom is missing at one of the edges, and calculate numerically the transmission probabilities. Depending

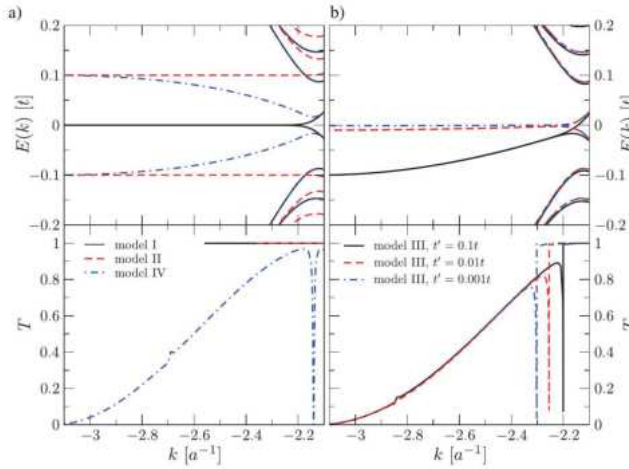


Figure 6.15 - Band structure $E(k)$ (upper panels) and transmission probability T (lower panels) as a function of the edge state wave vector k in a zigzag nanoribbon (width $W = 79a/\sqrt{3}$) where a single edge atom was removed at one of the edges. (a) shows data for nearest-neighbor hopping only (black solid line), including a constant staggered potential (red dashed line), and a position-dependent staggered potential with $M_0 = -0.1t, \bar{\gamma} = 1/a$ (blue dash-dotted line). (b) presents data including next-nearest-neighbor hopping for $t' = 0.1t$ (black solid line), $t' = 0.01t$ (red dashed line), and $t' = 0.001t$ (blue dash-dotted line).

on the situation, there may be more than one channel, and, in addition to that, a state localized to a single edge may be a linear combination of channels as in the case of next-nearest neighbors. Because of that we calculate the transmission eigenvalues T_n , obtained by diagonalizing $tt^\dagger$, where t is the transmission matrix containing all transmission amplitudes t_{mn} . The smallest transmission eigenvalue is related to the edge state, and all other transmission eigenvalues are found to be one. Thus, we simply refer to the smallest transmission eigenvalue as the "transmission probability of the edge state".

In Fig. 6.15 we show the transmission probability as a function of the edge state wave vector k . As discussed in Section 6.3, k determines the decay constant of the edge state, and thus how strongly it is localized to the edges. For the case of nearest-neighbors only or nearest-neighbors including a constant mass term (Fig. 6.15(a)) the transmission probability is very close to one and does not show a noticeable dependence on k , i.e. no dependence on the localization to the edges. Note that the data does not extend down to $k = -\pi/a$ in these cases: A numerical calculation cannot access a band edge, where the density of states diverges. In particular, it is not possible to reach $E = 0$ for the edge state in nearest-neighbor approximation. Even more, in nearest-neighbor approximation the energy of the edge state is exponentially small, and hence there will be a whole range of k -values around $k = \pm\pi/a$ that cannot be accessed numerically, because of the finite numerical precision. Thus, any perturbation that leads to an edge state dispersion is beneficial for numerical simulations in this respect, too.

Fig. 6.15(a) also shows the transmission probability for a position-dependent staggered potential. In this case, we see a pronounced dependence on k : As the edge state localizes stronger to the edge, the probability of scattering increases. In addition to this monotonic

¹²¹² In fact, it is not obvious how the current density enters when calculating a scattering probability. For example, in a perturbative treatment for very weak disorder, only the probability density enters (Fermi's golden rule). In contrast, here we consider strong scattering (missing atoms in the graphene lattice).

behavior, we observe some other features related to the band structure. In particular, the transmission probability is strongly suppressed (the edge state scatters strongly from the defect) at the extremum of the edge state band, close to the K -point ($k \approx -2.15/a$). The extremum of the edge state band separates left- and right-going states (corresponding to a positive and negative slope of the energy band), as does any band minimum or maximum. Close to this point, the wave functions of left- and right-going states have a good overlap, leading to a large probability of back-scattering from the defect. In addition, the back-scattering occurs within a single K -valley in this case. Aside from this pronounced effect, we also observe an additional small "wiggle" in the transmission probability ($k \approx -2.7/a$). This occurs whenever a new channel opens, as seen by comparing the respective energies with the plot of the band structure in Fig. 6.15(a).

For finite next-nearest neighbor hopping we observe an analogous behavior, shown in Fig. 6.15(b). In this case, the dependence of the transmission probability on k is almost independent of the exact value of t' , except for a shift in the position of the edge state band extremum. Note that t' covers two orders of magnitude in the examples shown in Fig. 6.15(b). This may be explained by the fact that the shape of the current density is largely independent of t' , as soon as $t' \gg t'_{\text{crit}}$. t' only determines the energy window where the edge state exists, as seen from the band structure plots in Fig. 6.15(b). This is another example demonstrating that the nature of the edge state is better characterized by the wave vector k , rather than the energy. Furthermore, although the calculated transmission probabilities for next-nearest neighbor hopping and the position-dependent staggered potential differ quantitatively, they show the same qualitative behavior.

Hence, next-nearest neighbor hopping or a position-dependent (staggered) potential not only leads to a current distribution located at the edges, but also results in substantial scattering from edge defects. The scattering is stronger, the more the edge state is localized to the edge.

19.5.15 6.5 Summary

In this chapter we have presented an exhaustive study on the properties of the graphene edge state. Within the commonly used nearest-neighbor tight-binding approximation for graphene, the edge state in a zigzag nanoribbon always occupies both edges, and hence both edges are maximally correlated. Moreover, the current carried by the edge state is mainly located in the center of the ribbon, opposite to the charge density localized at both sides of the ribbon. Thus, the edge state is affected only little by defects at the nanoribbon edges. With respect to transport, the edge state in this approximation rather behaves like a bulk state. These findings are in agreement with previous studies [177-180]. Introducing a constant staggered potential localizes the state to a single edge only, but does not change the transport properties. In this case we find the paradoxical situation that the charge density is asymmetric (localized at a single edge only), but the current density is symmetric with respect to the ribbon axis (flowing through the middle of the ribbon).

However, when next-nearest neighbor hopping, a position-dependent potential or a position-dependent staggered potential is included, the transport properties of the edge state change fundamentally: The current density is now also localized to the edges, in accordance with the charge density. As a consequence, the edge state scatters substantially from edge defects. This scattering is stronger, the more the state is localized to the edge. Furthermore, the edge state at either side of the nanoribbon is not affected by the opposite edge, and hence both edges are uncorrelated. Moreover, these fundamental changes prevail even if the perturbation induced by the above-mentioned terms is exponentially small with the nanoribbon width. Thus, the properties of the edge state that were complicated in a simple model of graphene, namely nearest-neighbor tight-binding, become simple and intuitive in a more complex model.

The fact that current and charge density are drastically different for zigzag nanoribbons in nearest-neighbor tight-binding approximation leads to many counter-intuitive effects and is unique to graphene. Yet, already an exponentially small perturbation leads to a fundamentally different behavior of the edge state. Hence, the nearestneighbor tight-binding approximation is a rather singular model with respect to the edge state, and does not apply to any realistic system.

19.6 Chapter 7 Edge state based spintronics in graphene

19.6.1 7.1 Introduction

Amongst its many unusual electronic properties, graphene is also a good candidate for spintronics applications: The spin relaxation time in graphene is expected to be very long, since spin-orbit effects are expected to be weak [39-41, 191], and because the predominant isotope in natural carbon is C^{12} which has zero nuclear spin, resulting in a vanishing hyperfine interaction [38]. In fact, spin injection from ferromagnetic metal contacts into graphene and coherent spin transport in graphene have been already demonstrated experimentally [42-46].

For ohmic contacts to the ferromagnetic electrodes, the conductance mismatch between the three-dimensional ferromagnetic metal contacts and the two-dimensional graphene sheet prevents efficient spin injection [192]. In order to overcome this obstacle in the experiment, a tunnel barrier is introduced between the ferromagnetic contact and the graphene sheet as suggested in Refs. [193, 194]. Thus, using ferromagnetic metals as a source of spin-polarized electrons for graphene involves a complicated layered structure with very different materials: Ferromagnetic metals as electrodes, some insulating material (Al_2O_3 in Refs. [43, 45, 46]), and finally graphene including the underlying substrate. In order to avoid the conductance mismatch problem, it would be desirable to arrive at a spin-injection design based on graphene only.

The graphene edge state at a zigzag edge is a promising candidate for magnetism, as it was predicted to exhibit a locally ferrimagnetic state [170]: Spins within the same sublattice are aligned parallel, whereas spins on different sublattices are aligned antiparallel. As a result, a zigzag nanoribbon shows a magnetic structure, where the spins at one of the edges are predominantly aligned in one direction, and the spins

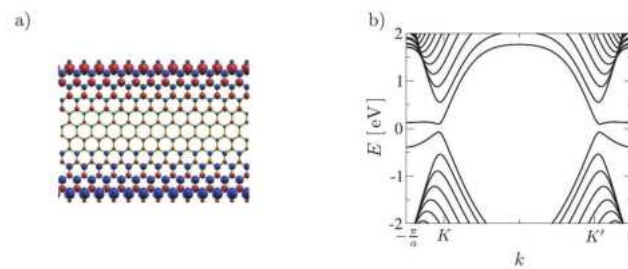


Figure 7.1 - (a) Ground state spin density of a zigzag graphene nanoribbon as obtained from DFT [195]. Spin up (down) is shown in blue (red). (b) Band structure of a zigzag graphene nanoribbon as obtained from DFT [196].

on the opposite edge in the antiparallel direction, as shown in Fig. 7.1(a). These predictions have been confirmed by density functional theory (DFT) calculations [164, 175, 176, 197], but although the existence of edge states has been verified through scanning tunneling spectroscopy, the existence of edge magnetism has not been shown experimentally yet. The antiferromagnetic ordering between the different sublattices opens a gap in the band structure of the zigzag nanoribbon, as shown in Fig. 7.1(b). Hence, edge magnetism is one possible explanation for the experimentally observed band gap of graphene nanoribbons (GNRs) [158-160].

The magnetic properties of the edge state have been studied in great detail in several

theoretical works. For example, the antiferromagnetic ordering between the different sublattices can be understood from a theorem about the Hubbard model on a bipartite lattice [198]. In addition, the ferromagnetic ordering at a single edge can be understood as a consequence of the Stoner criterion [199], as the edge states give rise to a very high density of states (see the discussion in the previous chapter). Still in general, a high density of states at the Fermi energy does not necessarily lead to magnetism, but may also result in some other instability, such as a geometrical distortion. However, DFT calculations favor magnetism over other instabilities [197]. Furthermore, from a general theorem that excludes ferromagnetic or antiferromagnetic ordering in onedimensional systems [200] it may seem that a zigzag graphene edge cannot exhibit true long-range magnetic ordering. However, long-range order has been shown to prevail at experimentally relevant, finite temperatures due to intrinsic or substrate-induced spin-orbit effects [201]. In addition to that, edge magnetism has been shown to exist not only in clean graphene nanoribbons, but also at disordered edges, provided that zigzag pieces with a length of at least 3-4 unit cells exist [195, 202-205]. Moreover, since edge states are expected to exist also for edge orientations other than zigzag [171],

edge magnetism also prevails in nanoribbons that are tilted with respect to the zigzag direction [204]. Finally, in order to achieve transport, the zigzag GNR must be doped to bring the Fermi energy into a regime with open conduction channels. This can be achieved in practice by a gate voltage or chemical doping. As the Fermi energy moves away from the high density of states created by the edge states, the magnetization decreases - only the edge states are spin-polarized, not the bulk states - until finally the magnetic ordering is destroyed. Within DFT calculations, the critical value of this doping is found to be ≈ 0.5 electrons (≈ 0.4 holes) per zigzag edge atom [195]. Hence, the magnetic structure survives a finite amount of doping necessary for transport.

In summary, many different theoretical predictions strongly support the existence of edge magnetism. In the remainder of this chapter we will therefore discuss the implications of this edge magnetization for graphene spintronics. Due to the nature of the magnetic ordering - antiferromagnetic between the two sublattices - a zigzag GNR does not exhibit global, but only local spin-polarization. Hence, the edge magnetism in clean GNRs cannot be used directly to generate spin-polarized currents. The objective of this chapter is to show new concepts how to generate spin-polarized and pure spin currents using the edge magnetism in zigzag graphene nanoribbons.

To this end, the chapter is organized as follows: First, we introduce a mean-field model to include the edge magnetism in the tight-binding model for graphene. Then, we show that zigzag graphene nanoribbons with rough edges exhibit universal spin conductance fluctuations and are thus a natural source of spin-polarized electrons. Based on these results we suggest an experiment to prove the existence of edge magnetism, using an all-electrical measurement. Finally, we show that the edge state can lead to a spin Hall effect in ballistic structures. In particular, this spin Hall effect can be used to generate spin-polarized currents in a three-terminal device, and a pure spin current in a four-terminal geometry. Since spin-polarized and pure spin currents are a prerequisite for spintronics applications, these results may serve as a foundation for an all-graphene spintronics.

19.6.2 7.2 Mean-field model for edge-state magnetism

We describe the edge magnetism in mean-field approximation within the tight-binding model for graphene. In this approximation, the Hamiltonian contains an additional term [170]

$$H_{\text{mag}} = \sum_{i,\alpha} \frac{1}{2} \mathbf{M}_{i\alpha} \cdot \mathbf{s} |i\alpha\rangle \langle i\alpha|, \quad (7.1)$$

where again (i, α) denotes the lattice site ($\alpha = A, B$), $\mathbf{M}_{i\alpha}$ the local magnetization and $\mathbf{s}$ the vector of Pauli matrices in (real, not pseudo-) spin space. The local magnetization can be determined self-consistently within the mean-field approximation to the Hubbard model [170], or obtained from DFT calculations. In this work however, we choose to employ a phenomenological model for $\mathbf{M}_{i\alpha}$, capturing the essential physics.

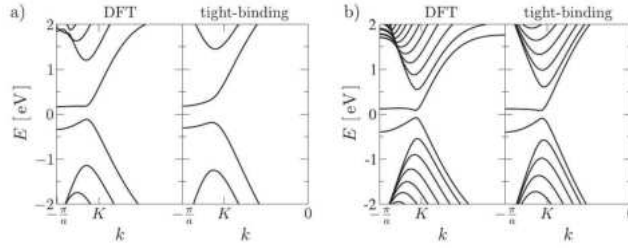


Figure 7.2 - Comparison between the band structure of a zigzag graphene nanoribbon, as obtained from DFT and the mean-field model within the tight-binding approximation. (a) shows the band structures of a GNR with width $W = 13a/\sqrt{3}$ and for the constant magnetization approximation with $M = 0.18t$. (b) Band structures for a ribbon with width $W = 31a/\sqrt{3}$ and the position-dependent magnetization from Eq. (7.4) with $M_0 = 0.192t$ and $t' = 0.067t$.

As seen in the previous section, spins are aligned in parallel within a single sublattice, but antiparallel between different sublattices. Hence, we can write the magnetic term in the Hamiltonian as

$$\langle i\alpha | H_{\text{mag}} | i\alpha \rangle = \begin{cases} +\frac{M_i}{2} s^z & \text{for } \alpha = A, \\ -\frac{M_i}{2} s^z & \text{for } \alpha = B. \end{cases} \quad (7.2)$$

Since the local magnetization forms a collinear texture, we can choose the magnetization to be in z -direction without loss of generality. Thus, the system can be decomposed into two independent subsystems for spin up and spin down. Within these subspaces, the magnetization term reduces to a position-dependent staggered potential, with opposite sign for spin up and spin down. Within the effective Dirac Hamiltonian, the magnetic term takes the form

$$H_{\text{mag}} = \frac{1}{2} M(\mathbf{x}) s^z \sigma^z, \quad (7.3)$$

where again σ^z acts in pseudospin (sublattice) space.

Finally, we have to specify the position dependence of the magnetization M_i . In the simplest approximation, we neglect the position dependence and set $M_i = M$, where M is a constant magnetization. Using the tight-binding model in next-nearest neighbor approximation with $t = 2.7\text{eV}$ and $t' = 0.1t$ we fit the magnetization M to the results of the DFT calculations for a narrow nanoribbon with width $W = 13a/\sqrt{3}$. We find a very good agreement between the band structures of the tight-binding model and the DFT calculations for $M = 0.18t$, as shown in Fig. 7.2(a).

constant M	$t = 2.7\text{eV}$	$t' = 0.1t$	$M = 0.18t$
position-dependent M_i (Eq.(7.4))	$t = 2.7\text{eV}$	$t' = 0.067t$	$M_0 = 0.192t$

Table 7.1 - Mean-field model parameters fitted to ab-initio DFT calculations.

A more realistic model, that also seeks to describe all details of the band structure in wider ribbons, must take into account the position dependence of the magnetization. Since the magnetization originates from spin-polarized edge states, it is largest at the edges and decays towards the middle of the graphene nanoribbon. We can estimate the magnetization close to an edge by summing over all edge state solutions, Eq. (6.22), such that

$$M_j = M_0 \times \frac{3a}{2\pi} \int_{\frac{2\pi}{3a}}^{\frac{4\pi}{3a}} dk |\phi_{\text{edge}}(k, j)|^2 = M_0 \times \frac{3a}{2\pi} \int_{\frac{2\pi}{3a}}^{\frac{4\pi}{3a}} dk e^{2\gamma(k)j} \quad (7.4)$$

where j is an index numbering the atoms relative to the edge, and $\gamma(k)$ is the decay constant of the edge state, that depends on the lateral Bloch vector k .¹ The integral has been normalized such that the magnetization has the value M_0 at the edge atom.

We now fit both t' and M_0 to the results of DFT calculations for a nanoribbon with width $W = 31a/\sqrt{3}$, which is the widest nanoribbon for which we have access to DFT data. For $t' = 0.067t$ and $M_0 = 0.192t$ we find very good agreement between the band structures from DFT and the tight-binding model, as shown in Fig. 7.2(b). At $k = \pm\pi/a$ the band structures exhibit a gap of size M_0 , since the edge state is then localized on the outermost atoms only. Moving closer to the K and K' -point, the band gap decreases. In this situation the edge states decay further into the nanoribbon and thus feel an average magnetization smaller than M_0 .

The parameters for the two models are summarized in Table 7.1. As we show below, the characteristics of the spin-polarized edge state transport does not depend on the details of the position dependence of the magnetization. The position dependence only modifies the energy window, where transport is dominated by the spin-polarized edge states. For this reason, we will employ the constant magnetization approximation also for wide ribbons, unless explicitly stated otherwise.

It is important to note that in both models the transport properties of the edge state are qualitatively the same: In both cases, current density is highest at the edge, and the edge state transport is strongly influenced by edge impurities, as discussed in Chapter 6. In contrast, if we had chosen a constant magnetization but $t' = 0$, the transport properties characteristics would be radically different. Such a singular model of graphene transport would also lead to radically different spin transport properties, as we demonstrate below.

19.6.3 7.3 Spin currents in rough graphene nanoribbons

19.6.4 7.3.1 Basic ideas

In the following, we work in the linear response regime and assume that the phase coherence length is much larger than the system size. The spin conductance of a zigzag GNR is defined as

$$G_s = \frac{e}{4\pi} (T_{\uparrow} - T_{\downarrow}), \quad (7.5)$$

where $T_{\uparrow(\downarrow)}$ is the transmission probability for spin up (down). Note that since the magnetization is collinear, the system decomposes into two independent subsystems, and no spin flip can occur. In contrast to the spin conductance, the total (charge) conductance of the system is given as (Eq. (2.120))

$$G_{\text{tot}} = \frac{e^2}{h} (T_{\uparrow} + T_{\downarrow}). \quad (7.6)$$

Note that the units of spin conductance are such that the spin current has units angular momentum per time, whereas the charge current has units charge per time.

Since the edge magnetization has the form of a staggered potential, the edge state is localized to a single edge only, as discussed in Section 6.3.3. Since spin up and down feel a staggered

¹²¹ Note that we used the $2\pi/a$ -periodicity of the Bloch vector k in order to write the integral in a compact form.

potential of opposite sign, Eq. (7.2), the different spins are localized to opposite edges, as shown in Fig. 7.3(a).

Due to the spatial localization of the edge state towards the zigzag edge, the edge state is scattered strongly by defects at the same edge, whereas it remains unaffected by defects on the opposite edge². Hence, distinguishing a left (l) and right (r) edge of the nanoribbon, we can thus approximate $T_{\uparrow(l)}$ by $T_{1(r)}$, where $T_{1(r)}$ is the transmission probability of the corresponding edge state assuming that the opposite edge is not disordered. The transport properties of the zigzag GNR are thus essentially those of two independent one-dimensional wires, oppositely spin polarized and connected in parallel between the reservoirs. We will demonstrate the validity of this two-wire model using numerical calculations.

In a clean zigzag graphene nanoribbon, the symmetry between both edges implies $T_l = T_r$ such that the spin conductance of a perfect zigzag GNR vanishes. This is demonstrated in Fig. 7.3(a): both spins are transmitted equally into the doped bulk graphene³. In order to obtain a finite spin conductance, we need to break the symmetry between both edges. This can be achieved by a large transverse electric field, as shown in Ref. [176]. In this case, one of the edge states is shifted out of the Fermi window,

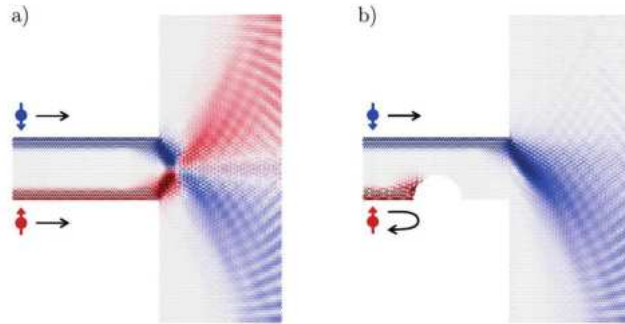


Figure 7.3 - Spin injection from (a) a perfect zigzag GNR and (b) a zigzag GNR with a large obstacle at one edge. Spin up (down) non-equilibrium densities (Eq. (A.22)) are shown in red (blue). In order to simulate spin injection into bulk graphene, where many transport channels are available, the band edge of the wider graphene ribbon has been shifted by a potential $V_{\text{gate}} = -0.6t$, mimicking the effect of a gate voltage. Note that the interference fringes are due to back reflection from the edge of the wide ribbon.

rendering the zigzag GNR completely spin-polarized. Here we will focus on alternative symmetry breaking method, that may be of more importance for the experiment: disorder at the graphene edges. If one of the edges is disordered more strongly, say the right edge, we can expect $T_r < T_l$, leading to a finite spin conductance. This is demonstrated in Fig. 7.3(b) where a large obstacle at one edge blocks the corresponding spin channel completely, leading to perfect spin injection. Hence, by controlling the roughness of the edges, zigzag GNRs can be designed to be a source of spin-polarized electrons.

19.6.5 7.3.2 Spin conductance of rough graphene nanoribbons

From an experimental perspective, unless nanoribbons are specifically fabricated with edges of different roughness, both edges will be similarly disordered. In this case, the average transmission probability of both edges will be equal $\langle T_l \rangle = \langle T_r \rangle$, resulting in a vanishing average spin conductance

¹²² As discussed in Chapter 6, this is the case for any realistic model of graphene. Here, "realistic" means going beyond the nearest-neighbor tight-binding approximation that behaves rather singular with respect to edge state transport

³ Although the nanoribbon has zero spin conductance in this situation, there may be a spin Hall effect-as seen in Fig. 7.3(a), different spins are deflected in opposite directions. We will discuss this effect in detail in Section 7.4.

$$\langle G_s \rangle = 0 \quad (7.7)$$

Here, $\langle \dots \rangle$ denotes an average over edge disorder configurations. However, since the disorder on both edges of the GNR is usually uncorrelated, a single graphene nanorib-

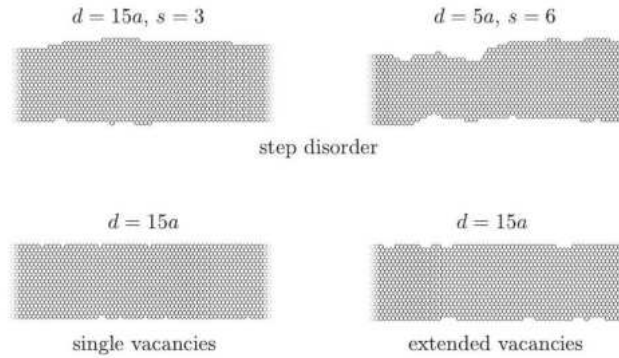


Figure 7.4 - Overview of the different disorder models used in this work. Step disorder: edge disorder created by a random walk, where the width of the nanoribbon is changed by one hexagon at every step. Steps are made with probability a/d and the maximum deviation of the width is $\leq s$ hexagons. Single vacancies: edge atoms are removed randomly with probability a/d . Extended vacancies: similar to single vacancies, but also neighboring edge atoms are removed. In all cases the disorder on the two edges is uncorrelated.

bon will typically exhibit $T_1 \neq T_r$ for a given energy, leading to mesoscopic spin conductance fluctuations. Within the two-wire model we find the variance $\text{Var } G_s$ of the spin conductance as

$$\text{Var } G_s = \left(\frac{e}{4\pi} \right)^2 (\text{Var } T_1 + \text{Var } T_r) = \left(\frac{h}{2e} \right)^2 \text{Var } G_{\text{tot}} \quad (7.8)$$

Since the transmission probabilities of both edges are uncorrelated, the variance of the spin conductance is simply given by the sum of the variances of the transmission probabilities for both edges. Hence, the variance of the spin conductance is equal to the variance of the charge conductance, up to a factor that is only due to the different units of spin and charge conductance.

In order to demonstrate the validity of the predictions of the two-wire model, we now perform numerical simulations of zigzag GNRs of different length L and width W . We employ several models of edge roughness, as summarized in Fig. 7.4, in order to show the independence of the effects on the type of edge disorder. In all disorder models, d denotes the average distance between scatterers. We first focus on the case of dilute disorder, $d \gg a$, and come back to the problem of strong disorder, $d \approx a$, in the next section.

We start by considering the ensemble averaged properties of a zigzag GNR, i.e. the average conductance and its fluctuations. Fig. 7.5(a) shows the typical behavior of the charge and spin conductances. In order to compare the average conductance and

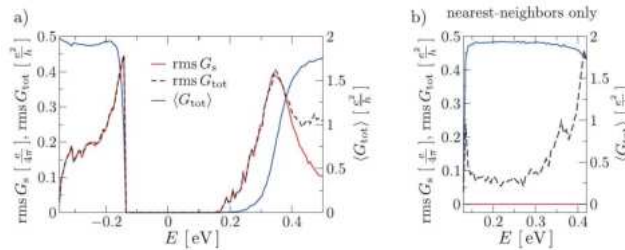


Figure 7.5 - (a) Average total conductance, $\langle G_{\text{tot}} \rangle$ (blue solid line), rms of the total conductance, $\text{rms } G_{\text{tot}}$ (black dashed line), and rms of the spin conductance, $\text{rms } G_s$ (red solid line), as a function of E ($E = 0$ is chosen to correspond to zero gate voltage). The data were averaged

over 1000 configurations of single vacancies with $d = 40a$, for a GNR with $W = 34a/\sqrt{3}$ and $L = 800a$. (b) shows the same quantities for a model with constant magnetization and nearest-neighbor hopping only ($t' = 0$).

the conductance fluctuations quantitatively, we characterize the fluctuations by the root-mean-square conductance, $\text{rms}G = \sqrt{\text{Var}G}$.

Note that over the whole energy region where the edge states are present we observe $\frac{\hbar}{2e}\text{rms}G_{\text{tot}} \approx \text{rms}G_s$, confirming the validity of the two-wire model. As the Fermi level is raised by gating or doping, the edge states extend further into the ribbon and start to feel both edges. Then, the assumption of uncorrelated channels breaks down, and $\frac{\hbar}{2e}\text{rms}G_{\text{tot}} > \text{rms}G_s$.

For an n-type GNR, when the Fermi level is near the band edge, the states are localized and both the average conductance and the fluctuations are suppressed exponentially. Raising the energy E , we observe in Fig. 7.5 (a) a crossover to the ballistic regime, where the conductance rises up to the quantum limit of conductance $2e^2/h$. Correspondingly, we see a maximum in the conductance fluctuations before they vanish again in the ballistic regime.

In the p-type regime, the situation is more complicated. Due to the finite dispersion of the edge state band, there are two open channels at the Fermi level (see Fig. 7.2(a)). One of these states is an edge state and strongly localized to the edge, whereas the other one extends further into the ribbon. As discussed in Section 6.4.4, scattering is strongest at the band edge, and correspondingly we observe a maximum in the spin conductance fluctuations close to the band edge. Although the n- and p-type regimes are seemingly quite different, they can be both described in a unified way, as we show in the next section.

Finally, we also consider the charge and spin conductance fluctuations for graphene

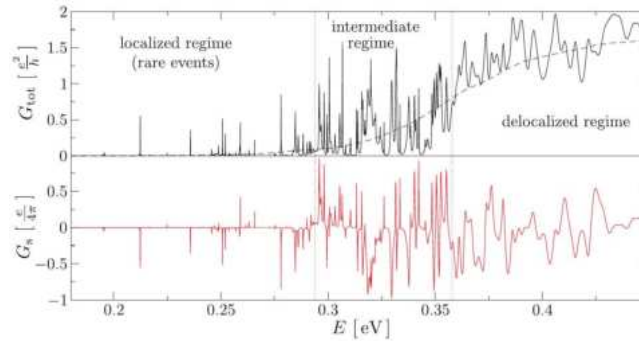


Figure 7.6 - (a) Charge conductance G_{tot} (black solid line) and (b) spin conductance G_s (red solid line) as a function of energy E for a single graphene nanoribbon with $W = 34a/\sqrt{3}$ and $L = 800a$. As in Fig. 7.5, the edge disorder model is single vacancies with $d = 40a$, but instead of an average over several disorder configurations, we only show the transport properties of a single configuration. For comparison, the grey dashed curve in (a) shows $\langle G_{\text{tot}} \rangle$ from Fig. 7.5(a).

in nearest-neighbor tight-binding approximation with a constant magnetization. In Chapter 6 we have shown that this is a rather singular model with respect to edge state charge transport. In particular, the edge state was influenced equally by scatterers on both edges. Fig. 7.5(b) shows that this also leads to a singular behavior with respect to spin transport properties: Whereas we obtain finite charge conductance fluctuations, the spin conductance fluctuations vanish. Since the edge state is influenced by both edges equally, $T_{\uparrow} = T_{\downarrow}$, and the spin conductance is always zero. Hence, also for spin transport properties it is essential to go beyond the nearest-neighbor tight-binding approximation.

After discussing ensemble-averaged properties, we now turn to the transport properties of a single, disordered nanoribbon. Experimentally, the average conductance would be observed when measuring an array of different nanoribbons at the same time. In this case, spin conductance would be quenched. However, if a single GNR is measured, the results will be dominated by the conductance fluctuations.

In Fig. 7.6 we show both the charge and the spin conductance of a single GNR as a function of energy E in the n-type regime. In order to compare the average charge conductance $\langle G_{\text{tot}} \rangle$ with the charge conductance of a single ribbon, we show both quantities in the upper panel of Fig. 7.6. We can clearly identify different regimes:

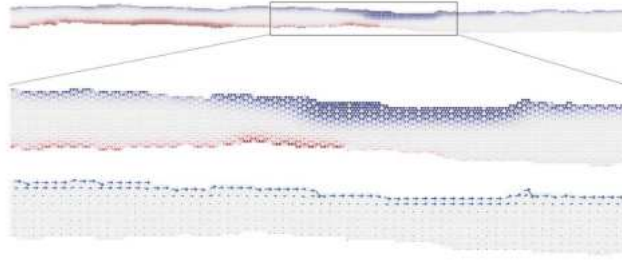


Figure 7.7 - Non-equilibrium spin (upper and middle panel) and current density (lower panel) of a zigzag GNR with $W = 34a/\sqrt{3}$ and $L = 600a$. The figure shows one configuration of a GNR with step disorder ($d = 10a$ and $s = 6$). In the upper two panels, Spin up (down) density is shown in blue (red). The lower two panels show a magnified part of the nanoribbon, as indicated in the upper panel.

For low energies, close to the band edge, the system is localized and the conductance is dominated by rare events that appear as sharp peaks in the transmission function with a large transmission probability. These are due to the long tail of the distribution function for the transmission probabilities, and the average conductance is not a representative quantity [206]. Remarkably, every one of these rare events leads to a large spin conductance. Since these large transmission probabilities are "rare", only one of the spin-polarized edges will transmit, whereas the other one is blocked, leading to a completely spin-polarized current. For large energies, where the edge state starts to extend further into the ribbon and scattering from the edge becomes weaker, the conductance of the single GNR follows the average conductance, albeit with sizable smooth fluctuations. These fluctuations somewhat resemble the universal conductance fluctuations (UCF) known from quasi-one-dimensional wires [207] with many channels. Strictly speaking, this is not true: The conductance arises from a single channel only at each edge, thus acting as a one-dimensional wire. A one-dimensional system however always localizes, and a metallic regime is not present [206, 207]. Again, these charge conductance fluctuations give rise to (smooth) spin conductance fluctuations. In the intermediate regime we observe a mixture of smooth fluctuations and rare events, that again carries over to the spin conductance. In fact, as seen from Fig. 7.5(a), the spin conductance fluctuations are largest in this case.

Finally, we illustrate the wave function associated with a rare event in Fig. (7.7). There, we chose to plot the non-equilibrium spin density and the charge current for an energy E where the transmission probability is strongly peaked with $T(E) \approx 1$. The edge state "meanders" along the strongly disordered edge, as seen both from the spin as well as current densities, that exactly follow the shape of the edge. While one of the edges transmits with a high probability, the edge state on the opposite side is eventually blocked, leading to a completely spin-polarized current. In fact, it is quite remarkable that the edge state, which is strongly localized towards the edge in this case, can follow variations of the ribbon edge larger than its own extent. This is due to the fact that a few hexagons in zigzag orientation are enough to support a local edge state [169]; in order to transmit, the edge state then only has to overcome a single step to arrive at the next available state. We will comment on this topic some more at the end of the next section.

19.6.6 7.3.3 Universal spin conductance fluctuations

After discussing the typical behavior of the spin conductance and verifying the predictions of the two-wire model for a given disorder model, we will now show that the properties of the spin conductance are in fact independent on the model for edge disorder. In particular, we will find a universal maximum value of the spin conductance fluctuations.

Within the two-wire model, we have stated that the conductance of a zigzag GNR is given by the conductances of two independent, disordered wires. The transmission eigenvalue statistics of disordered wires is governed by the Dorokhov-Mello-PereyraKumar (DMPK) equation [208, 209] (for a review, see Ref. [207]). For the onedimensional, single-channel case, the DMPK equation reduces to

$$l \frac{\partial}{\partial L} P(\lambda, L) = \frac{\partial}{\partial \lambda} \lambda(1 + \lambda) \frac{\partial}{\partial \lambda} P(\lambda, L) \quad (7.9)$$

where l is the mean free path, L the wire length, $\lambda = (1 - T)/T$, T the transmission probability and $P(\lambda, L)$ a probability distribution function for λ , normalized such that $\int_0^\infty d\lambda P(\lambda, L) = 1$. From this distribution function, we can calculate the expectation value of any function of the transmission probability, and this value is uniquely determined by the ratio L/l of the mean free path l and the length of the wire L .⁴ This fact is also referred to as single-parameter scaling.

Note that the differential equation (7.9), that is an equation for the distribution function, can be converted to a differential equation for expectation values. This is particularly useful for

$$\langle \ln T \rangle = - \int_0^\infty d\lambda \ln(1 + \lambda) P(\lambda, L) \quad (7.10)$$

as this results in a closed differential equation:

$$\begin{aligned} l \frac{\partial}{\partial L} \langle \ln T \rangle &= - \int_0^\infty d\lambda \ln(1 + \lambda) l \frac{\partial}{\partial L} P(\lambda, L) \\ &= - \int_0^\infty d\lambda \ln(1 + \lambda) \frac{\partial}{\partial \lambda} \lambda(1 + \lambda) \frac{\partial}{\partial \lambda} P(\lambda, L) \\ &= + \int_0^\infty d\lambda \lambda \frac{\partial}{\partial \lambda} P(\lambda, L) \\ &= - \int_0^\infty d\lambda P(\lambda, L) = -1, \end{aligned} \quad (7.11)$$

where we used Eq. (7.9), the normalization of $P(\lambda, L)$ and integration by parts twice. Together with the boundary condition $\lim_{L \rightarrow 0} \langle \ln T \rangle = 0$ (i.e. $T \rightarrow 1$) we arrive at

$$\langle \ln T \rangle = - \frac{2L}{\xi} \quad (7.12)$$

where $\xi = 2l$ is the localization length [207].

Apart from giving rise to differential equations for expectation values, the DMPK equation (7.9) can also be solved for the distribution function $P(\lambda, L)$ directly (see Refs. [210, Eq. (20)] and [211, Eq. (4.3)]):

$$P(\lambda, L) = \frac{1}{\sqrt{2\pi}} \left(\frac{l}{L} \right)^{\frac{3}{2}} e^{-\frac{L}{4l}} \int_{\text{arccosh}(1+2\lambda)}^\infty du \frac{u e^{-\frac{u^2 L}{4l}}}{\sqrt{\cosh u - 1 - 2\lambda}}. \quad (7.13)$$

¹²⁴ This can be seen by a change of coordinates $L \rightarrow L/l$.

From this explicit form of the distribution value, we can calculate arbitrary expectation values - at least numerically. In particular, we obtain the variance of the spin conductance within the two-wire model as

$$\text{Var } G_s = \left(\frac{e}{4\pi}\right)^2 2 \text{Var } T = \frac{2e^2}{(4\pi)^2} \left(\int_0^\infty d\lambda \frac{P(\lambda, L)}{(1+\lambda)^2} - \left(\int_0^\infty d\lambda \frac{P(\lambda, L)}{1+\lambda} \right)^2 \right) \quad (7.14)$$

As seen from the DMPK equation (7.9), $\text{Var } G_s$ only depends on l/L , or equivalently ξ/L . Hence, in order to compare the spin conductance fluctuations for different disorder models and n - as well as p-type ribbons, we extract the energy dependence of the localization length $\xi(E)$ from the numerical simulations using Eq. (7.12). The result for different disorder models is shown in Fig. 7.8. We observe a monotonic increase of the localization length ξ with increasing (decreasing) energy in the n-type (p-type) regime, as the edge state extends further into the ribbon and is scattered less from edge defects. Note that the disorder involving single vacancies only (red dashed curve in Fig. 7.8) has the largest distance d between scatterers, but on average the shortest localization length. At first glance this is somewhat unexpected, but can be understood from the fact that all other types of disorder always feature at least a few hexagons in zigzag orientation in a row, supporting a local edge state. This is not the case for

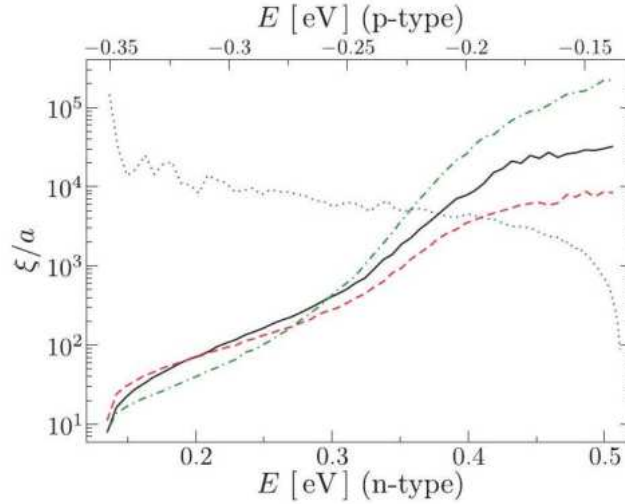


Figure 7.8 – ξ/a as a function of E for step disorder ($d = 30a, s = 3$; black solid line), single vacancies ($d = 40a$; n-doped: red dashed line, p-doped: blue dotted line) and extended vacancies ($d = 30a$, green dashed-dotted line).

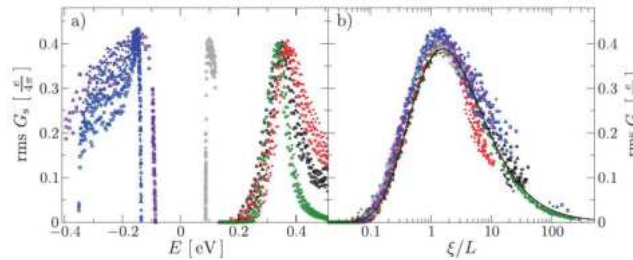


Figure 7.9 - Spin conductance fluctuations: $\text{rms } G_s$ as a function of (a) E and (b) ξ/L for n- and p-doped graphene. We show data within the constant magnetization approximation: step disorder for n-type, $d = 20a, s = 3$ (black), single vacancies for n- and p-type, $d = 40a$ (red and blue, respectively) and extended vacancies for n-type, $d = 30a$ (green); we also present data for the model including a position-dependent magnetization: single vacancies for n- and p-type, $d = 40a$ (grey and violet, respectively). The solid line corresponds to the DMPK prediction, Eq. (7.14). The data is shown for graphene nanoribbons with lengths $L = 800a(\circ), 1000a(\square), 1200a(\triangle), 1400a(+),$ and $1600a(\times)$, and width $W = 34/\sqrt{3}a$. The $\text{rms } G_s$ is estimated from 1000 disorder configurations.

the single vacancies which only involve one hexagon in the graphene lattice, and thus leads to stronger scattering.

In Fig. 7.9 we show $\text{rms}G_s$ for zigzag GNRs within the constant magnetization approximation and a more realistic model with a position-dependent magnetization (as discussed in Section 7.2) and all three disorder models (as shown in Fig. 7.4), with different values of d and a wide range of ribbon lengths L . Whereas the functional dependence of the rms conductance $\text{rms}G_s$ on the energy E is very different for the various models as shown in Fig. 7.9(a), all the data collapse onto a single curve when plotted as a function of ξ/L (Fig. 7.9(b)). The results of the numerical simulations agree very well with the DMPK prediction in Eq. (7.14) (shown as black solid line in Fig. 7.9(b)). The agreement is especially good for n-doped GNRs in constant magnetization approximation, where there is only one conductance channel per spin. In this case, the single-channel DMPK equation (7.9) holds exactly. For the p-doped ribbon in constant magnetization approximation and the ribbons with a position-dependent magnetization, there are two conductance channels per spin (c.f. Fig. 7.2), and we observe a slight increase in the rms conductance with respect to the single-channel case. In this situation, the predictions of the single-channel DMPK equation only hold approximately, and the increase in $\text{rms}G_s$ is presumably due to a beginning crossover to a multi-channel quasi-1D wire [207]. For all models, the maximum value of the spin conductance fluctuations fits well to $\text{rms}G_s \approx 0.4e/4\pi$. These considerations show that the spin conductance fluctuations are indeed independent from the model employed to describe the edge magnetization and the type of disorder, and are universal in this respect.

In Fig. 7.10 we concentrate on n-doped graphene for step disorder (upper panels of Fig. 7.4), that may be closest to a real, experimentally obtained nanoribbon. Fig. 7.10(a) shows again the universality of the spin conductance fluctuations with respect to a wide range of parameters characterizing edge roughness, ribbon length and width. There is little dependence on the ribbon width W , confirming that the observed effect is entirely due to the edges.

Up to now, we have focussed on the case of dilute disorder $d \gg a$. However, currently there is not much experimental control over the edges of nanoribbons. Hence, such nanoribbons may very well exhibit dense disorder, where d is on the order of the lattice constant, $d = \mathcal{O}(a)$. In Fig. 7.10(b) we show the maximum value of spin conductance fluctuations as a function of d/a . We observe that the maximum of the spin conductance fluctuations is independent from the maximum height of the steps s , and retains the maximum universal value $\text{rms}G_s = 0.4e/4\pi$ for $d \gtrsim 5a$. For $d < 3a - 5a$, the maximum value decreases rapidly. The maximum of the spin conductance fluctuations starts to decrease, when the system remains always in the strongly localized regime for the energy range corresponding to the edge state, since the conductance fluctuations are suppressed in this case. Still, even in this regime, rare events will lead to spin-polarized currents, as discussed in Section 7.3.2. For $d \gtrsim 5a$, the edge state still has an appreciable transmission probability, and therefore we observe the universal maximum of the spin conductance fluctuations. Again, this can be understood from

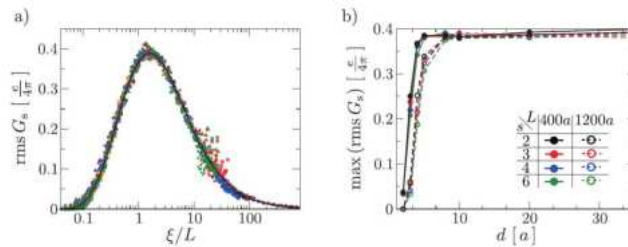


Figure 7.10 - (a) $\text{rms}G_s$ as a function of ξ/L for step disorder in n-doped graphene: $d = 20a$ and $s = 3$ (red; orange for $W = 94a/\sqrt{3}$), $d = 35a$ and $s = 2$ (black), $d = 35a$ and $s = 6$ (blue; violet for $W = 94a/\sqrt{3}$), $d = 20a$ and $s = 6$ (green). The solid line corresponds to the DMPK prediction, Eq. (7.14). The data is shown for GNR lengths $L = 800a(\bigcirc)$, $1000a(\square)$, $1200a$

(Δ), $1400a(+)$, and $1600a(\times)$, width $W = 34/\sqrt{3}a$ unless specified otherwise. The rms G_s is estimated from 1000($W = 32a/\sqrt{3}$) and 750($W = 92a/\sqrt{3}$) disorder configurations. (b) Maximum value of rms G_s as a function of d/a for the step disorder models.

the fact that a few (3-4) hexagons in zigzag orientation already support a local edge state [169]. Hence, the transport properties of the edge state are surprisingly stable with respect to edge disorder, just as the magnetic properties of the edge state at disordered edges [195, 202-205].

19.6.7 7.3.4 All-electrical detection of edge magnetism

Although there is evidence for the existence of the edge state at graphene edges on graphite through STM measurements [172-174], neither the existence of the edge state in exfoliated graphene nor the edge magnetism have been proven experimentally yet. In fact, both STM on exfoliated graphene edges and magnetic STM pose many experimental challenges [212]. Here, we propose an experimental alternative to STM for proving the existence of edge magnetism using an all-electrical measurement only. This measurement makes use of the finite spin conductance G_s of a zigzag nanoribbon, predicted in the previous sections, and is based on the ideas put forward in Ref. [213].

Since the edge magnetization breaks time-reversal symmetry, the finite spin conductance G_s can be in principle measured both by two- or four-probe setups involving ferromagnetic leads [214]. Fig. 7.11 shows a possible non-local four-probe measurement setup, similar to the experimental setups for spin injection into bulk graphene [43-46]: Current is driven through a ferromagnetic contact, injecting spins into bulk graphene. The spin-polarized charge current from the ferromagnet leads to a spin accumulation underneath the ferromagnetic contact. This spin accumulation can be understood as a

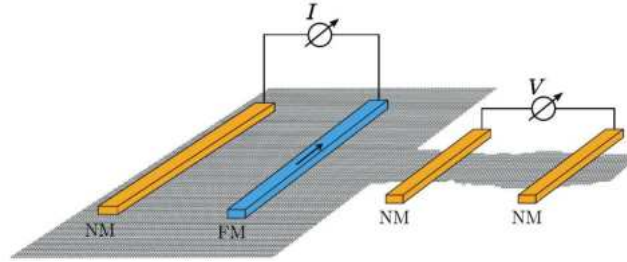


Figure 7.11 - Possible four-probe measurement setup for an all-electrical detection of edge magnetism: Spin polarized current is injected via a ferromagnetic (FM) contact (shown in blue), creating a spin accumulation. The diffusing spin accumulation causes a voltage drop due to the finite spin conductance G_s of the nanoribbon that can be measured by nonmagnetic (NM) contacts (shown in orange).

difference in chemical potential for spin up and down. Hence, such a spin accumulation leads to a pure spin current. In a nonmagnetic material, this spin current does not lead to a voltage drop, as both spin up and down are conducted equally. However, in the graphene nanoribbon with $G_s \neq 0$ the conductances for spin up and down are different. Hence, the pure spin current would be converted into a charge current, if the circuit were closed, but since no current may flow through the (unbiased) graphene nanoribbon, this leads to a voltage drop along the nanoribbon. This voltage drop can be measured with non-magnetic contacts. Moreover, the measured voltage depends on the orientation of the injected spins. Note that these mechanisms are explained in more detail for the case of diffusive systems in Ref. [215], and for the case of coherent transport in Ref. [213].

A measured voltage drop that depends on the orientation of the magnetization in the ferromagnet would be an unambiguous sign of a finite spin conductance. The existence of a finite spin conductance would then in turn also be evidence for edge magnetism.

19.6.8 7.4 Edge-state induced Spin Hall effect

19.6.9 7.4.1 Spin-dependent deflection of the edge state

We now come back to the spin-dependent deflection of the edge states already observed in Fig. 7.3, and shown again in Fig. 7.12(a). Here, a zigzag graphene nanoribbon with edge magnetization is coupled to a larger (non-magnetic) ribbon with a negative potential V_{gate} , in order to simulate doped bulk graphene with many channels.

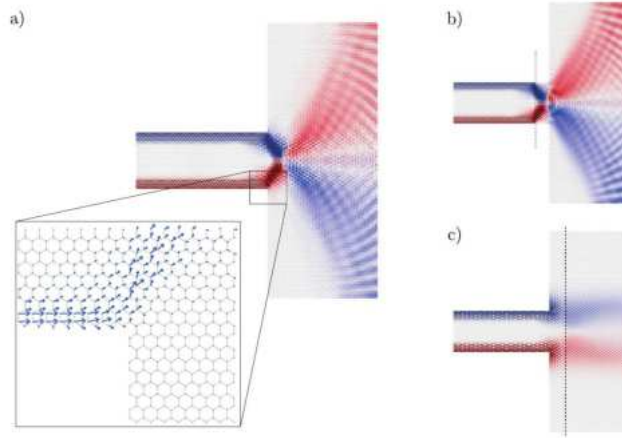


Figure 7.12 - Spin injection from a perfect zigzag GNR into a wider ribbon. Spin up (down) non-equilibrium densities are shown in red (blue). In order to simulate spin injection into bulk graphene, where many transport channels are available, the band edge of the wider graphene ribbon has been shifted by a potential $V_{\text{gate}} = -0.6t$, mimicking the effect of a gate voltage. This potential is switched on (a) directly at the interface between the narrow and the wide ribbon, (b) some distance within the narrow ribbon, and (c) some distance within the wide ribbon (as indicated by black dashed lines). The inset in (a) shows the current density close to one of the edges.

When the edge state enters the bulk, it must expand, as the solution strongly localized to the zigzag edge is not a bulk solution. Furthermore, because of the edge geometry at the contact between the narrow and the wide ribbon, one might expect quantum mechanical diffraction at the edge. At first glance we might thus expect a uniform expansion of the edge state, possibly enhanced in the direction of the edge due to diffraction. This is not observed in the numerical simulations: Instead, we see a clear deflection away from the graphene edge. Since spin up and down are located on opposite edges, they are deflected in opposite directions and hence separated, leading to a spin Hall effect, similar to systems with spin-orbit interaction [25-28].

In order to understand this effect, first it should be noted that even if the geometry of the system is symmetric with respect to the ribbon axis, as in Fig. 7.12(a), the staggered magnetization breaks this symmetry. The Hamiltonian of the system is then only invariant under reflection about the ribbon axis and $s^z \rightarrow -s^z$, i.e. exchange of spin up and spin down. This can be seen from Eq. (7.2) and the fact that a reflection about the ribbon axis interchanges sublattice A and B. Hence, there is no symmetry that would exclude a spin Hall effect. Moreover, being localized towards one zigzag boundary, the edge state experiences a highly asymmetric scattering situation.

Second, by closer inspection we observe that the deflection already starts before the narrow GNR opens into the bulk. This is particularly obvious from the behavior of the current density close to the edge, as shown in the inset of Fig. 7.12(a), and is due to evanescent waves, that tend to be more important in graphene than in usual semiconductor electron gases (for an example see Ref. [216]; a further example are the (spin) currents in mesoscopic Hall bridges presented in the next section). Furthermore, from the current density we also observe that the edge state starts to expand uniformly (there is both current flowing towards the edge and

away from the edge). However, the part expanding towards the edge is immediately reflected, leading to an overall deflection away from the edge.

Hence, it seems that the deflection is simply a geometrical effect due to the localized nature of the edge state. This picture is supported by simulations, where we changed the position where the potential V_{gate} is switched on: In Fig. 7.12(b) the potential is switched on within already the narrow ribbon, and we observe again a deflection away from the edge. In contrast, when the potential is switched on after some distance within the wide ribbon, the edge state expands uniformly, as there is no edge at the point where the state starts to expand. Furthermore, numerical simulations show that even without a magnetization, the (then unpolarized) edge state deflects away from the edge, ruling out the valley effects discussed in Ref. [217].

In order to demonstrate that this deflection also leads to a spin-polarized current, we consider a three-terminal T-bar structure, as shown in the insets of Fig. 7.13(a) and (b), where we consider transport from lead 1 to the leads 2 and 3. Unlike before, we do not apply an additional potential V_{gate} .

The spin conductance from lead l' to lead l is then defined as

$$G_{ll'}^s = G_{ll'}^{\uparrow} - G_{ll'}^{\downarrow} \quad (7.15)$$

where $G_{mn}^{\uparrow(\downarrow)}$ are the corresponding charge conductances for spin up (down). Note that due to the symmetry reasons discussed above, $G_{21}^{\uparrow(\downarrow)} = G_{31}^{\downarrow(\uparrow)}$, and hence $G_{21}^s = -G_{31}^s$.

In Fig. 7.13(a) we plot G_{21}^s as a function of energy for a geometry where a narrow zigzag GNR contacts a very wide armchair ribbon. This situation is close to a zigzag GNR contacting bulk graphene. In the n-type regime, we observe an on average positive spin conductance, in agreement with the previous spin density pictures (Fig. 7.12). The spin conductance exhibits sharp peaks corresponding to energies, where a new channel opens in the wide nanoribbon; this is remanence of the still finite width. In the p-type regime, the spin conductance is negative, because spin up and down are then located on the opposite edge as compared to the n-type case.

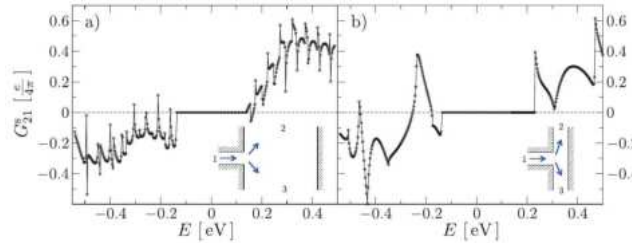


Figure 7.13 - Spin Hall effect: G_{21}^s as a function of energy E in a T-bar geometry with a zigzag GNR (lead 1) with width $W_1 = 75.5a/\sqrt{3} \approx 36a$ and (a) an armchair GNR (lead 2 and 3) with width $W_2 = W_3 = 152a$ and (b) an armchair GNR with width $W_2 = W_3 = 31a$.

In contrast to the case, where we inject spins into a wide armchair ribbon, Fig. 7.13(b) shows the spin conductance G_{21}^s for a geometry consisting of a narrow zigzag ribbon and an armchair ribbon of comparable width. In this case, the number of channels in the zigzag and armchair ribbon are comparable, and the arguments from above (in particular, the "expansion" of the edge state, based on bulk graphene with many channels) only apply to some extent here. However, the scattering situation is again asymmetric for an edge state localized to a boundary. Correspondingly, we again observe a finite spin Hall effect, but the spin conductance G_{21}^s exhibits larger fluctuations related to the number of open channels in the armchair ribbon, and also changes in sign. Still, the overall trend of the sign of G_{21}^s is compatible with the deflection picture.

19.6.10 7.4.2 Pure spin currents in graphene micro-bridges

After using the spin-dependent deflection of the edge state in order to generate spinpolarized currents in a T-bar geometry, we now turn to the problem of generating pure spin currents, i.e. spin currents without corresponding charge currents.

For this, we use a mesoscopic Hall-bar geometry as sketched in Fig. 7.14. Current is flowing through the leads 1 and 4, whereas the leads 2 and 3 act as voltage probes, such that $I_2 = I_3 = 0$. We choose lead 4 as the voltage reference, and set $V_4 = 0$. In linear response, the currents I_l and the voltages V_l are related through Eq. (2.118)

$$I_l = \sum_{l'} G_{ll'} (V_{l'} - V_l) \quad (7.16)$$

and for given values of the currents I_l , the voltages V_l are determined uniquely (see, for example, Ref. [59, chap. 4]). Whereas the charge currents $I_l = I_l^\uparrow + I_l^\downarrow$ are strictly

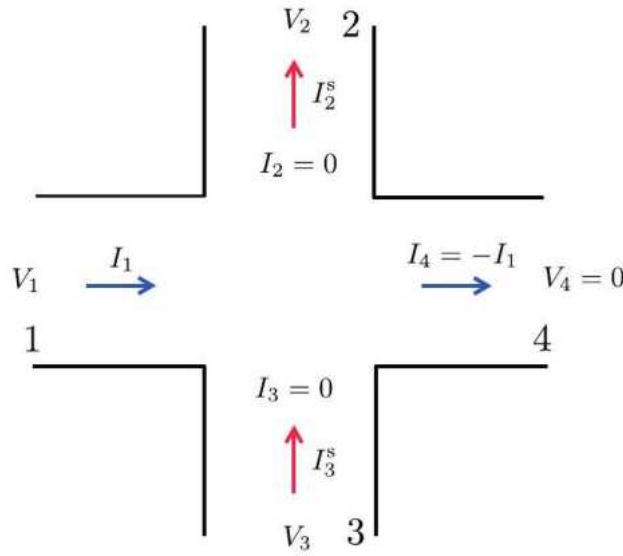


Figure 7.14 - Hall-bar geometry used to generate pure spin currents. Leads 1 and 4 are made of zigzag GNRs, whereas leads 2 and 3 are armchair GNRs. zero within the voltage probes, the spin currents

$$I_l^s = \frac{\hbar}{2e} (I_l^\uparrow - I_l^\downarrow) \quad (7.17)$$

can be very well finite. Here, $I_l^{\uparrow(\downarrow)}$ is defined as the current carried by spin up (down). Note that for the system under consideration the definition of the spin-current is not plagued by ambiguities, in contrast to systems with spin-orbit coupling [218], as spin is a good quantum number for a collinear magnetization. The spin Hall conductance of the micro-bridge is defined as [219]

$$G_{\text{SH}} = \frac{I_2^s}{V_1 - V_4} = \frac{I_2^s}{V_1}, \quad (7.18)$$

i.e. as the spin current response to a applied voltage. Since spin up and down form two independent subsystems, the spin currents are simply given as

$$I_l^s = \sum_{l'} G_{ll'}^s (V_{l'} - V_l) \quad (7.19)$$

where $G_{ll'}^s$ is defined as in Eq. (7.15). Hence, the spin Hall conductance, Eq. (7.18), is given as

$$G_{\text{sH}} = G_{21}^{\text{s}} + G_{23}^{\text{s}} \frac{V_3}{V_1} - (G_{21}^{\text{s}} + G_{23}^{\text{s}} + G_{24}^{\text{s}}) \frac{V_2}{V_1} \quad (7.20)$$

Before evaluating G_{sH} numerically, it is useful to discuss its behavior for symmetric systems. Consider a Hall-bar that has a reflection symmetry about the y -axis, such as

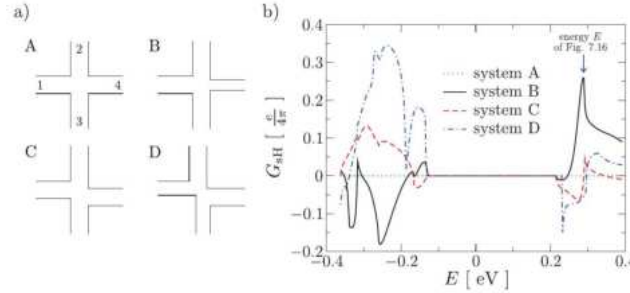


Figure 7.15 - (a) Different geometries considered for generating pure spin currents: system A is symmetric with respect to the axis spanned by leads 2 and 3, whereas systems B, C, and D have been slightly desymmetrized. A: $W_1 = W_4 = 49a/\sqrt{3} \approx 28a$, $W_2 = W_3 = 32a$. B: as system A, but $W_4 = 46a/\sqrt{3}$. C: as system A but lead 4 is shifted by $2a$. D: as in system A, but all leads are shifted and varied in width by $1 - 4a$. Note that for clarity of the presentation, the asymmetries are exaggerated in the picture. (b) Spin Hall conductance G_{sH} as a function of energy E for the Hall-bar geometries shown in (a).

the geometry shown in Fig. (7.14). This reflection maps lead 1 into lead 4 and vice versa, but keeps leads 2 and 3 unchanged. Then, $G_{24}^{\uparrow(\downarrow)} = G_{21}^{\uparrow(\downarrow)}$ and hence $G_{24}^{\text{s}} = G_{21}^{\text{s}}$. Eq. (7.20) then simplifies to

$$G_{\text{sH}} = G_{21}^{\text{s}} \left(1 - \frac{2V_2}{V_1}\right) + G_{23}^{\text{s}} \left(\frac{V_3}{V_1} - \frac{V_2}{V_1}\right). \quad (7.21)$$

Moreover, the reflection symmetry about the y -axis leads to $V_2 = V_3 = V_1/2$ [219], and therefore in a perfectly symmetric Hall-bar $G_{\text{sH}} = 0$. In such a symmetric system it is thus not possible to generate pure spin currents, but only spin-polarized currents.

This is fundamentally different from a system with spin-orbit coupling, where the Hamiltonian of the system is only invariant under a reflection and a spin flip, leading to $G_{24}^{\uparrow(\downarrow)} = G_{21}^{\downarrow(\uparrow)}$ (and, in addition, spin-flip terms) and a finite spin Hall conductance even for symmetric geometries, as seen for the case of Rashba spin-orbit coupling in Ref. [219]. In the graphene Hall-bar, spin is a good quantum number, and the spin Hall effect arises only due to the geometry, i.e. the boundaries of the system. In particular, the spin Hall conductance vanishes, when the hall voltages V_2 and V_3 of the full system (spin up and down) also lead to $I_2^{\uparrow} = I_2^{\downarrow} = 0$, i.e. if the spin subsystems exhibit the same hall voltages as the full system. This is the case for a symmetric Hall-bar. Nevertheless, already little deviations from perfect symmetry are enough to obtain a finite spin Hall

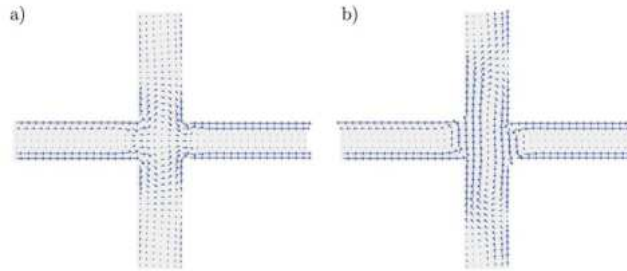


Figure 7.16 - (a) Charge and (b) spin current density in a desymmetrized Hall-bar, corresponding to system B in Fig. 7.15(a). The energy $E \approx 0.285\text{eV}$, corresponding to the peak indicated with an arrow in Fig. 7.15(b). In both (a) and (b), the current densities have been

averaged over several hexagons in order to bring out the characteristics of the current flow. conductance, as shown below.

Fig. 7.15(b) shows the spin Hall conductance G_{sH} as function of energy E for a symmetric Hall-bar (system A in Fig. 7.15(a)) and slightly desymmetrized Hall-bars (system B, C and D in Fig. 7.15(a)). Note that the asymmetry in the picture is exaggerated to show the difference in the systems more clearly. The real asymmetry involves only a few hexagons, as seen from the numbers in the figure caption.

As expected from the symmetry considerations above, the spin Hall conductance of the symmetric Hall bar vanishes. Nevertheless, the slightly desymmetrized Hall-bars show a finite spin Hall conductance $G_{\text{sH}} = 0.1 - 0.4 \frac{e}{4\pi}$. Thus, the spin Hall conductance has a magnitude similar to the mesoscopic spin Hall effect in semiconductor-based two-dimensional electron gases with spin-orbit coupling [219, 220]. Furthermore, we observe that the spin Hall conductance depends strongly on the geometrical details of the Hall-bar.

Finally, in Fig. 7.16 we show an example of both the charge (Fig. 7.16(a)) as well as the spin current density (Fig. 7.16(b)) in a desymmetrized Hall-bar. Whereas no net charge current is flowing in the voltage probes, we observe that the charge current from the current probes extends rather far into the voltage probes, and in addition, we observe current vortices within the voltage probes. These are due to evanescent modes, that tend to be of great importance in graphene (see, for example, Ref. [216]). In the zigzag ribbons, the spin current is flowing along the edges, and in opposite directions for opposite edges. This is due to the fact that opposite edges carry opposite spins. In addition, we also observe a finite, pure spin current flowing in the transverse direction, as expected from a finite G_{sH} .

19.6.11 7.5 Summary

In this chapter we have discussed various possibilities to generate spin-polarized charge and pure spin currents from the edge magnetism in zigzag graphene nanoribbons. These predictions were verified by numerical simulations for a mean-field model of the edge magnetism within the next-nearest neighbor approximation.

In particular, we have found that edge disorder in zigzag GNRs naturally leads to spin-polarized currents, as it breaks the symmetry between the two oppositely spinpolarized edges. If a zigzag nanoribbon is fabricated such that one of the edges is more clean than the other edge, conductance will be preferentially carried by the spinpolarized edge state at the cleaner edge, leading to a finite spin conductance. Even if both edges are equally rough, single zigzag GNRs exhibit a finite spin conductance due to spin conductance fluctuations. These spin conductance fluctuations exhibit a universal maximum value of $\text{rms}G_{\text{s}} \approx 0.4 \frac{e}{4\pi}$ and are independent of the type of disorder at the edges, and also independent of the details in the mean-field model for the edge magnetism.

In addition, the spin-polarized edge states can give rise to a spin Hall effect. This spin Hall effect can be used to generate spin-polarized currents in three-terminal geometries, and pure spin currents in four-terminal geometries. In contrast to the spin-polarized currents in rough graphene nanoribbons, the spin Hall effect requires ballistic systems, and does not exhibit a universal behavior, as the effect is due to the details of the underlying geometry.

The results of this chapter show that zigzag nanoribbons are an efficient alternative to ferromagnetic metals for generating spin-polarized and pure spin currents, thus paving the way to an all-graphene spintronics.

19.7 Chapter 8 Summary and perspectives

19.7.1 8.1 Summary

In the first part of this thesis we have developed algorithms for a generic approach to quantum transport calculations, valid for arbitrary tight-binding models.

From general considerations within the non-equilibrium Green's function theory we have found that for such a generic approach it is necessary to solve two problems: first, the calculation of the surface Green's function of a lead and second, the efficient calculation of the retarded Green's function of the scattering region. In order to be generally applicable, these problems had to be solved without restrictive assumptions on the underlying tight-binding model or the geometry of the system.

The problem of calculating the surface Green's function of the leads was approached in Chapter 3. The lead was modeled as a semi-infinite wire consisting of a periodic repetition of unit cells. Under that premise, we derived an expression for the lead Green's function in terms of Bloch eigenstates of the corresponding infinite wire, extending previously known results. A direct implementation of this expression was found to be numerically unstable in certain, important examples. Motivated by insights from an analysis of this failure, we developed a numerically stable algorithm to calculate the lead Green's function in terms of the Schur decomposition of a certain matrix, in contrast to the direct implementation that was based on the eigendecomposition.

Chapter 4 was then concerned with the efficient calculation of the retarded Green's function. To this end we analyzed several well-established quantum transport algorithms, such as the recursive Green's function method. We found that a wide class of transport algorithms depends on the Hamiltonian matrix being in a block-tridiagonal form; this is the only essential requirement. Based on graph partitioning techniques we developed a matrix reordering technique to bring an arbitrary matrix into a blocktridiagonal form suitable and optimal for transport. Through this reordering, well-established quantum transport algorithms, that were restricted to wire geometries in their original formulation, can be applied to any tight-binding model and any geometry, including multi-terminal systems. In addition we found that matrix reordering can also lead to a significant speed-up of numerical calculations. These advantages of using the matrix reordering technique were then demonstrated on representative examples.

The results of Chapters 3 and 4, together with the generalized Fisher-Lee relation of Appendix C, form the foundation of a generic computer program, applicable to any tight-binding system. In order to adapt this approach to a certain tight-binding model, the only remaining task is to set up the respective Hamiltonian matrix. This may still be laborious, but is straightforward. For the purpose of this work, we considered the tight-binding description arising from the finite differences approximation to the effective mass Schrödinger equation in two and three dimensions (Appendix F), as well as the next-nearest neighbor tight-binding model for graphene (Appendix G).

In the second part of this thesis we then applied the numerical methods developed in the first part to study spin-dependent effects in graphene and magnetic tunnel junctions.

The tunneling anisotropic magnetoresistance effect in an epitaxial Fe/GaAs/Au tunnel junction was the subject of Chapter 5. In particular, we investigated the magnetic field dependence of the TAMR effect in order to explain recent experimental findings. Previous work suggested that the origin of the TAMR effect in this magnetic tunnel junction is due to spin-orbit coupling in the GaAs barrier, and presented a corresponding microscopic model. In this work, we found that it is enough to include the orbital effects of a magnetic field into this model to explain the experimental observations. Both numerical results and a phenomenological model reproduce all characteristic features of the measured magnetic field dependence of the TAMR

effect. This magnetic field dependence could be explained by an interplay of orbital effects due to the magnetic field and the Dresselhaus spin-orbit coupling within the GaAs barrier.

Finally, Chapters 6 and 7 were concerned with charge and spin transport properties of the graphene edge state in zigzag graphene nanoribbons.

The results of previous work seemed to indicate very surprising transport properties of the edge state: Although the probability density is located close to the edges of the ribbon, current flows through the middle of the ribbon, rendering the edge state rather insensitive to edge disorder. In Chapter 6, we have found these counter-intuitive results to be an artifact of the nearest-neighbor tight-binding model used almost exclusively for transport calculations in graphene. In fact, exponentially small corrections to this model in the form of next-nearest neighbor hopping or a position-dependent potential or staggered potential were found to change the transport properties fundamentally, restoring the current flow to the edge, in accordance with the probability density. This then also leads to a more intuitive transport behavior: The edge state is strongly scattered by defects located on the same edge, whereas it remains virtually unaffected by defects located on the opposite edge. These effects could be understood by symmetry considerations within perturbation theory. From these findings we have concluded that

nearest-neighbor tight-binding does not apply to any realistic system where the edge state plays a role, and hence strongly advocate the inclusion of next-nearest neighbor hopping.

Based on the theoretically predicted edge magnetism in zigzag graphene nanoribbons, we have proposed several schemes for generating spin-polarized and pure spin currents within graphene nanoribbons in Chapter 7. For this we have employed a mean-field model of the edge magnetism within the next-nearest neighbor tight-binding approximation of graphene, which, as seen from Chapter 6, is suitable for studying edge state properties. We have found that edge disorder in zigzag graphene nanoribbons naturally leads to spin-polarized currents. In particular, rough graphene nanoribbons exhibit spin conductance fluctuations with a universal maximum value of $\text{rms}G_s \approx 0.4e/4\pi$, independent of the type of edge disorder and the details of the mean-field model for the edge magnetism. Based on the finite spin conductance of zigzag nanoribbons, we have suggested an all-electrical measurement to prove the existence of the edge magnetism that has been experimentally elusive so far. Moreover, we have found that the edge state can lead to a geometry-induced spin Hall effect in ballistic nanostructures. This spin Hall effect can be used to generate spin-polarized currents in three-terminal devices, and pure spin currents in four-terminal structures. These results show that zigzag nanoribbons are an efficient alternative to ferromagnetic metals for generating spin currents, paving the way to an all-graphene based spintronics.

19.7.2 8.2 Outlook

We have emphasized that the generic framework for numerical transport calculations developed in this thesis is applicable to arbitrary tight-binding systems, beyond the examples discussed in this work. For example, the code has also been used to study the topological Hall effect [57]. Moreover, Viktor Krükl [56] has extended the existing program code to study transport of heavy holes within the Luttinger-Kohn Hamiltonian [221]. Applying this transport framework to other systems is just as well possible, including tight-binding models originating from density functional theory calculations, as in Refs. [84-86].

Note that the discussion in this work was restricted to orthogonal tight-binding models. The extension to nonorthogonal tight-binding models is straightforward, but tedious, as the use of a nonorthogonal basis involves some subtleties: For example, the anticommutation relations for creation and annihilation operators $c_i, c_j^\dagger$ then involve the overlap matrix S instead of a Kronecker delta (for example, see Ref. [222]):

$$\{c_i, c_j^\dagger\} = \delta_{ij} \quad \rightarrow \quad \{c_i, c_j^\dagger\} = (S^{-1})_{ij} \quad (8.1)$$

Thus, extending the theory to nonorthogonal tight-binding models requires some care. The theoretical and experimental work on the TAMR effect in Fe/GaAs/Au tunnel junctions has always involved mutual stimulations between theorists and experimentalists here in Regensburg, both in the original work [47] as well as in the work presented here. From the model presented in Chapter 5 we have made some predictions on the magnetic field behavior of the conductance that can be tested experimentally. In turn, new experimental data may lead to further extensions of the existing theory.

The field of graphene physics has seen a tremendous evolution in the past four years and still contains many unsolved puzzles. In this work, we were concerned only with one aspect of graphene, namely the graphene edge state and its spin transport properties.

In this study, the role of the edges was of essential importance. In fact, edges are expected to influence many aspects of graphene: For example, there have been several studies on the influence of edge disorder on the (charge) conductance of graphene nanoribbons [162, 165, 166] in order to explain the yet ununderstood transport gap of graphene nanoribbons [158-160]. All of these studies are carried out using the nearestneighbor tight-binding model for graphene. As seen discussed in Chapter 6, this model leads to spurious transport results for the edge state. Especially in the low-energy regime, where the transport gap is experimentally observed, the edge state is expected to play an important role. Hence, the transport calculations in Refs. [162, 165, 166] should be tested against a model including next-nearest neighbors.

Edges have already been shown to be of crucial importance for the symmetry classes in ballistic graphene cavities [50], and may be of relevance in recent experiments on graphene quantum dots [223]. On the other hand, in diffusive graphene the chiral nature of the charge carriers is expected to show in a weak antilocalization effect, if intervalley scattering is suppressed [224]. However, the weak antilocalization observed in the experiment has often been more or less suppressed with respect to the theoretically predicted value (for examples, see Refs. [225-227]). Again, the edges may play an important role in this case. Obviously in narrow graphene nanoribbons, intervalley scattering from sharp edge disorder will dominate, suppressing weak antilocalization [224]. However, also for samples that have a width comparable to their length the influence of the edges may be significant [228]. These predictions can be conveniently tested by numerical simulations.

Interest in graphene has grown tremendously in the past few years, and will continue to yield fascinating physics in the future. Here, we have only mentioned a few examples, that are related to the topics discussed in this work. Especially the numerical techniques developed in the course of this thesis will be of great use for future studies on graphene.

19.8 Appendix A Observables in tight-binding

19.8.1 A. 1 Single-particle operators in tight-binding

Expressions for observables in envelope approximation, such as charge or current densities, can be derived quite generally for any tight-binding Hamiltonian. For that purpose, we consider a tight-binding Hamiltonian including spin with orbitals $\chi_{i,s}(\mathbf{x})$, where i is the index of the lattice site $\mathbf{x}_i$ and $s = 1, 2$ the spin index. Without loss of generality, the discussion is restricted to a single orbital per lattice site, and the orbitals are assumed to be eigenstates of σ^z . The Hamiltonian is given by

$$H = \sum_{i,j} \sum_{s,s'=1}^2 t_{ij,ss'} |is\rangle \langle js'| \quad (A.1)$$

Since H must be a hermitian operator, $t_{ij,ss'} = t_{ji,s's}^*$. The density operator in tight-binding representation is given by

$$\begin{aligned} n(\mathbf{x}) &= \sum_{i,j,s,s'} \int d\mathbf{x}' \chi_{i,s}(\mathbf{x}')^* \delta(\mathbf{x} - \mathbf{x}') \chi_{j,s'}(\mathbf{x}') |is\rangle \langle js'| \\ &= \sum_{i,j,s} \chi_{i,s}(\mathbf{x})^* \chi_{j,s}(\mathbf{x}) |is\rangle \langle js| \end{aligned} \quad (\text{A.2})$$

This expression contains an overlap between orbitals of different lattice sites. If the orbitals are sufficiently localized, this overlap can be neglected and only the orbital i closest to $\mathbf{x}$ will contribute to the density:

$$n(\mathbf{x}) = \sum_s |\chi_{i,s}(\mathbf{x})|^2 |is\rangle \langle is| \quad (\text{A.3})$$

Furthermore, if we are only interested in the variation of the wave function's envelope, we can neglect the variations of the orbitals. For that purpose we approximate the atomic orbitals by the characteristic function of a volume (or area) V around the lattice site $\mathbf{x}_i$:

$$\chi_{i,s}(\mathbf{x}) = \begin{cases} \frac{1}{V} & \text{if } \mathbf{x} \text{ is in the vicinity of } \mathbf{x}_i \\ 0 & \text{else} \end{cases} \quad (\text{A.4})$$

To keep the notation simple, we assume that V is the same for each lattice point¹. With that, we arrive at the final expression for the electron density operator

$$n(\mathbf{x}_i) = \frac{1}{V} \sum_s |is\rangle \langle is| \quad (\text{A.5})$$

where $\mathbf{x}_i$ should be understood to represent any point $\mathbf{x}$ in the vicinity of $\mathbf{x}_i$. In a similar manner we can derive the tight-binding representation of the spin density operator

$$n^\mu(\mathbf{x}_i) = \frac{1}{V} \sum_{s,s'} \sigma_{ss'}^\mu |is\rangle \langle is'| \quad \text{for } \mu = 0, x, y, z \quad (\text{A.6})$$

where $\sigma^0 = \mathbb{1}$, so that Eq. A. 6 also contains the previously derived $n(\mathbf{x}_i)$. For the current density, it is necessary to derive an expression for the velocity operator first. The position operator is given by

$$\begin{aligned} \mathbf{x} &= \sum_{i,j,s,s'} \int d\mathbf{x}' \chi_{i,s}(\mathbf{x}')^* \mathbf{x}' \chi_{j,s'}(\mathbf{x}') |is\rangle \langle js'| \\ &= \sum_{i,s} \mathbf{x}_i |is\rangle \langle is| \end{aligned} \quad (\text{A.7})$$

where we assume again that the atomic orbitals are sufficiently localized. The velocity operator is related to the time derivative of the position operator:

$$\begin{aligned}
\mathbf{v} &= \frac{d}{dt} \mathbf{x} = \frac{i}{\hbar} [H, \mathbf{x}] \\
&= \frac{i}{\hbar} \sum_{i,j,s,s'} \sum_{n,t} t_{ij,ss'} \mathbf{x}_n [|is\rangle \langle js'|, |nt\rangle \langle nt|] \\
&= \frac{i}{\hbar} \sum_{i,j,s,s'} (\mathbf{x}_j - \mathbf{x}_i) t_{ij,ss'} |is\rangle \langle js'|
\end{aligned} \tag{A.8}$$

From the density and the velocity operator, we can define the current operator as $\mathbf{j}(\mathbf{x}) = \frac{1}{2} \{n(\mathbf{x}), \mathbf{v}\}$ [229]. Note that in the transition from classical to quantum mechanics, the classical expression for the current, $\mathbf{j} = n\mathbf{v}$, must be symmetrized by the anti-commutator $\{\cdot, \cdot\}$ to obtain a Hermitian operator [153, Chap. 3]. Inserting Eqs. (A.2) and (A.8), we obtain

$$\mathbf{j}(\mathbf{x}_i) = \frac{i}{2\hbar V} \sum_{j,s,s'} (\mathbf{x}_j - \mathbf{x}_i) (t_{ij,ss'} |is\rangle \langle js'| - t_{ji,s's} |js'\rangle \langle is|). \tag{A.9}$$

When local currents are considered in tight-binding models, this is often done in terms of bond currents I_{ij} between two sites i and j [179, 230]. The current density operator (A.9) is equivalent to these bond currents, but contains in addition the distance vector between two lattice points $(\mathbf{x}_j - \mathbf{x}_i)$ (leading to the vector character of $\mathbf{j}(\mathbf{x})$, these are usually put in by hand for the bond currents) and a constant prefactor (giving the correct units for a current density).

Analogous to the current density operator, we also define the spin current density operator as $\mathbf{j}^\mu(\mathbf{x}) = \frac{1}{2} \{n^\mu(\mathbf{x}), \mathbf{v}\}$, where again $\mu = 0$ contains the previous expression for the current density, Eq. (A.9). Then we find

$$j^\mu(\mathbf{x}_i) = \frac{i}{2\hbar V} \sum_{j,s_1,s_2,s_3} (\mathbf{x}_j - \mathbf{x}_i) (\sigma_{s_1 s_3}^\mu t_{ij,s_3 s_2} |is_1\rangle \langle js_2| - \sigma_{s_3 s_1}^\mu t_{ji,s_2 s_3} |js_2\rangle \langle is_1|) \tag{A.10}$$

Note that there are several inequivalent definitions for the spin current. In addition, if the spin operator does not commute with the Hamiltonian, as e.g. for systems with spin-orbit interaction, these spin currents do not obey current conservation laws [218]. However, in this work we need not worry about such subtleties, as we will consider the spin current only in systems where spin is a good quantum number. In particular, when $t_{ij,ss'} = t_{ij,s} \delta_{ss'}$, the spin current density simplifies to

$$\mathbf{j}^\mu(\mathbf{x}_i) = \frac{i}{2\hbar V} \sum_{j,s,s'} (\mathbf{x}_j - \mathbf{x}_i) (\sigma_{ss'}^\mu t_{ij,s'} |is\rangle \langle js'| - \sigma_{s's}^\mu t_{ji,s} |js'\rangle \langle is|) \tag{A.11}$$

19.8.2 A. 2 Observables in terms of Green's functions

The transition from a single-particle operator A to the corresponding many-body operator $\mathcal{A}$ is particularly simple in the case of a tight-binding Hamiltonian: Since $\mathcal{A} = \sum_{i,j,s,s'} \langle is|A|js'\rangle c_{is}^\dagger c_{js'}$ (for an introduction to many-body operators see Ref. [70]), the many-body operator is simply obtained by replacing $|is\rangle \rightarrow c_{is}^\dagger$ and $\langle is| \rightarrow c_{is}$. Observables, i.e. expectation values of these operators, are then given in terms of lesser Green's functions at equal times $t' = t$,

$$G^<(js', t, is, t) = G_{js', is}^<(t, t) = \frac{i}{\hbar} \langle c_{is}^\dagger(t) c_{js'}(t) \rangle \tag{A.12}$$

¹²¹ For example, this is true for any grid based on a Bravais lattice.

as defined in Section 2.5.1. In particular, we find the spin density

$$n^\mu(\mathbf{x}_i, t) = -\frac{i\hbar}{V} \sum_{ss'} \sigma_{ss'}^\mu G_{is',is}^<(t, t) \quad (\text{A.13})$$

and the spin current density

$$\begin{aligned} j^\mu(\mathbf{x}_i, t) &= \frac{1}{2V} \sum_{j,s_1,s_2,s_3} (\mathbf{x}_j - \mathbf{x}_i) (\sigma_{s_1s_3}^\mu t_{ij,s_3s_2} G_{js_2,s_1}^<(t, t) - \sigma_{s_3s_1}^\mu t_{ji,s_2s_3} G_{is_1,s_2}^<(t, t)) \\ &= \frac{1}{V} \sum_{j,s_1,s_2,s_3} (\mathbf{x}_j - \mathbf{x}_i) \Re \left(\sigma_{s_1s_3}^\mu t_{ij,s_3s_2} G_{js_2,is_1}^<(t, t) \right) \end{aligned} \quad (\text{A.14})$$

In the expression for the spin current density we made use of the fact that $\sigma_{ss'}^\mu = (\sigma_{s's}^\mu)^*$, $t_{ij,ss'} = t_{ji,s's}^*$ and Eq. (2.22), $G_{is_1,js_2}^<(t, t) = -(G_{js_2,is_1}^<(t, t))^*$.

In steady-state, the Green's functions only depend on time-differences, and hence the observables are time-independent, as discussed in Chapter 2. In this case it is advantageous to work in the energy domain: From the Fourier transformation of the lesser Green's function, Eq. (2.24), we then find

$$n^\mu(\mathbf{x}_i) = \int dE n^\mu(\mathbf{x}_i, E) \quad (\text{A.15})$$

where

$$n^\mu(\mathbf{x}_i, E) = -\frac{1}{2\pi V} \sum_{ss'} \sigma_{ss'}^\mu G_{is',is}^<(E) \quad (\text{A.16})$$

and

$$\mathbf{j}^\mu(\mathbf{x}_i) = \int dE \mathbf{j}^\mu(\mathbf{x}_i, E) \quad (\text{A.17})$$

where

$$j^\mu(\mathbf{x}_i, E) = \frac{1}{2\pi\hbar V} \sum_{j,s_1,s_2,s_3} (\mathbf{x}_j - \mathbf{x}_i) \Re \left(\sigma_{s_1s_3}^\mu t_{ij,s_3s_2} G_{js_2,is_1}^<(E) \right) \quad (\text{A.18})$$

The lesser Green's function necessary for the evaluation of these observables are conveniently obtained from the Keldysh equation (2.112).

Note that the energy-dependent observables $n^\mu(\mathbf{x}_i, E)$ and $\mathbf{j}^\mu(\mathbf{x}_i, E)$ are the expectation values of the respective operators in terms of eigenstates of the system at energy E , weighted with the density of states, as seen from Eq. (2.123). As such, they are not directly observable in an experiment.

In linear response however, we can interpret, for example, $n(\mathbf{x}_i, E)$ as a non-equilibrium density. To this end, consider a system with two leads, left (L) and right (R), with $\mu_L > \mu_R$ as in Fig. 2.3. As discussed in Section 2.5.2, in linear response the electrochemical potentials of the leads are close to the Fermi energy, $\mu_R \approx \mu_L \approx E_F$, and the voltage $V = (\mu_L - \mu_R)/e$ is small. We split the density $n(\mathbf{x})$ in two parts,

$$n(\mathbf{x}) = n_{\text{eq}}(\mathbf{x}) + n_{\text{non-eq}}(\mathbf{x}) \quad (\text{A.19})$$

where

$$n_{\text{eq}}(\mathbf{x}) = \int_{-\infty}^{E_F} dE n(\mathbf{x}, E) \quad (\text{A.20})$$

is the equilibrium density, and

$$n_{\text{non-eq}}(\mathbf{x}) = \int_{E_F}^{E_F+eV} n(\mathbf{x}, E) \quad (\text{A.21})$$

is the non-equilibrium density induced by the voltage V . When $n(\mathbf{x}, E)$ does not vary much in the energy set by the voltage, we can then approximate

$$n_{\text{non-eq}}(\mathbf{x}) \approx n\left(\mathbf{x}, E_F + \frac{eV}{2}\right) \times eV \quad (\text{A.22})$$

We can thus interpret the energy-dependent observables evaluated at some energy above E_F as voltage-induced non-equilibrium densities.

19.9 Appendix B The recursive Green's function method

19.9.1 B. 1 Basic principles

The recursive Green's function technique (RGF) was pioneered by Thouless and Kirkpatrick [118] and by Lee and Fisher [119] in 1981. A few years later, MacKinnon presented a formulation of the RGF technique that is commonly used today [103]. Since then, the RGF method has been applied to various systems numerous times. There are also a number of extensions, for example to deal with non-equilibrium systems [121, 122]

Instead of calculating the retarded Green's function G_S^r of the scattering region directly from Eq. (2.111), the recursive Green's function method constructs G_S^r by building up the system recursively. To this end, the system is divided into slices [103]: In the example of a square grid, lattice points with the same longitudinal coordinate are grouped into vertical slices, as depicted in Fig. B.1. The RGF method starts from a system without any slices in the scattering region, and then successively adds layer by layer, until the full system has been constructed. Obviously, the system can only be partitioned into slices as shown in Fig. B.1, if there are only two, parallel leads, i.e. a geometry similar to a wire. Thus, the RGF method is in its common formulation restricted to geometries with only two collinear leads.

The recursive Green's function method is very often only formulated for the particular case of a square grid. It is however much more generic, as a closer inspection of the prerequisites of the RGF technique shows. In mathematical terms, the RGF method is based on the fact that the tight-binding Hamiltonian H of a two-terminal system can

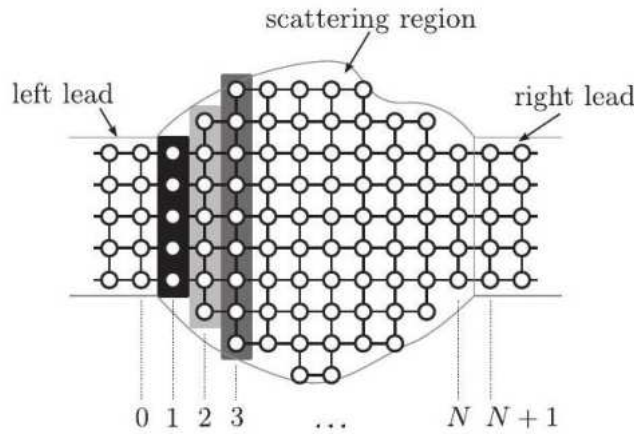


Figure B. 1 - A system defined on a square grid can be partitioned into slices i , by grouping lattice points with the same longitudinal coordinate. This kind of partitioning leads to a block-tridiagonal tight-binding matrix, where every slice corresponds to a matrix block. The different slices are shown in alternating shades of grey.
be written in block-tridiagonal form:

$$H_0 = \begin{pmatrix} H_{i-1,i-1} & 0 & & & \\ 0 & H_{i,i} & H_{i,i+1} & & \\ & H_{i+1,i} & H_{i+1,i+1} & H_{i+1,i+2} & \\ & & H_{i+2,i+1} & \ddots & \end{pmatrix} \quad (\text{B.2})$$

From this form of the Hamiltonian, we have

$$G_{0,j,k}^r = \begin{cases} G_{j,k}^{r,R(i)} & \text{for } j, k \geq i \\ (E - H_{i-1,i-1} + i\eta)^{-1} & \text{for } j = k = i - 1 \\ 0 & \text{otherwise} \end{cases} \quad (\text{B.3})$$

The isolated block is attached to the system by means of the "perturbation"

$$V = \begin{pmatrix} 0 & H_{i-1,i} & & \\ H_{i,i-1} & 0 & & \\ & & \ddots & \end{pmatrix} \quad (\text{B.4})$$

Note that V only has matrix elements between the slices $i - 1$ and i .

From the Dyson equation, we find, in a calculation similar to that in Section 2.5.2,

$$G_{i-1,i-1}^{r,R(i-1)} = G_{0,i-1,i-1}^r + G_{0,i-1,i-1}^r H_{i-1,i} G_{i,i-1}^{r,R(i-1)} \quad (\text{B.5})$$

and

$$G_{i,i-1}^{r,R(i-1)} = G_{0,i,i}^r H_{i,i-1} G_{i-1,i-1}^{r,R(i-1)} \quad (\text{B.6})$$

Inserting (B.6) into (B.5) and considering the individual contributions of G_0^r , Eq. (B.3), we arrive at

$$G_{i-1,i-1}^{r,R(i-1)} = \left(E - H_{i-1,i-1} - H_{i-1,i} G_{i,i}^{r,R(i)} H_{i,i-1} \right)^{-1} \quad (\text{B.7})$$

where we neglected the infinitesimal $i\eta$, as this does not play a role in the numerics, as long as $H_{i-1,i} G_{i,i}^{r,R(i)} H_{i,i-1}$ has a finite imaginary part. Usually this is the case, for instance when there are propagating modes in the structure. In a similar fashion, we also obtain

$$G_{N+1,i-1}^{r,R(i-1)} = G_{N+1,i}^{r,R(i)} H_{i,i-1} G_{i-1,i-1}^{r,R(i-1)} \quad (\text{B.8})$$

and

$$G_{i-1,N+1}^{r,R(i-1)} = G_{i-1,i-1}^{r,R(i-1)} H_{i-1,i} G_{i,N+1}^{r,R(i)}. \quad (\text{B.9})$$

In a similar fashion, we can define the Green's function $G^{r,L(i)}$ that contains all blocks $\leq i$ (see Fig. B.2(c)).

The recursive Green's function calculation then starts from $G_{N+1,N+1}^{r,R(N+1)} = g_R^r$ and $G_{0,0}^{r,L(0)} = g_L^r$, where g_R^r and g_L^r are the surface Green's functions of the right and left lead respectively (on the topic of calculating these surface Green's functions, see Chapter 3). Applying Eqs. (B.7)-(B.9) recursively allows then to calculate $G^{r,R(i)}$ for any i . Note however that, for example, $G^{r,R(1)}$ is not yet the desired Green's function G^r of the full system - it is still disconnect from the left lead. Thus, in order to obtain the full Green's function of the system, the Green's functions $G^{r,R(i+1)}$ and $G^{r,L(i)}$ must be connected via the Dyson equation, as depicted in Fig. B.2(d). Depending on the desired values of the Green's function, this is done in various ways, as summarized in the next section.

In the case of non-interacting systems, or interacting systems in mean-field approximation, the lesser self-energy $\Sigma^<$ only contains contributions from the leads. In this case, the Keldysh equation (2.112) can be written as

$$G_{i,j}^< = G_{i,0}^r \Sigma_L^< (G_{j,0}^r)^\dagger + G_{i,N+1}^r \Sigma_R^< (G_{j,N+1}^r)^\dagger \quad (\text{B.10})$$

where we made use of $G^a = (G^r)^\dagger$, Eq. (2.21). Here, $\Sigma_L^<$ ($\Sigma_R^<$) is the self-energy contribution from the left (right) lead. It is thus sufficient to compute $G_{i,0}^r$ and $G_{i,N+1}^r$ for $0 \leq i \leq N+1$ in order to obtain all desired values of $G^<$.¹ This is the basic idea behind the extensions of the RGF algorithm to non-equilibrium systems [121, 122].

19.9.2 B. 2 Summary of algorithms

Eqs. (B.7) - (B.9) are the building blocks of the various forms of the RGF method. The exact procedure of a calculation depends on which values of the Green's function G^r should be calculated. Here we summarize the computational steps distinguishing a few, prominent examples.

Algorithm B.1. Compute $G_{N+1,0}^r$ and $G_{0,0}^r$.

A Use Eqs. (B.7) and (B.8) to compute $G_{i,i}^{r,R(i)}$ and $G_{N+1,i}^{r,R(i)}$ successively for $i = N \dots 1$.

B Compute $G_{N+1,0}^r$ and $G_{0,0}^r$ as

$$G_{0,0}^r = \left((g_L^r)^{-1} - H_{0,1} G_{1,1}^{r,(1)} H_{1,0} \right)^{-1} \quad (\text{B.11})$$

and

$$G_{N+1,0}^r = G_{N+1,1}^{r,(1)} H_{1,0} G_{0,0}^r, \quad (\text{B.12})$$

Applications: From $G_{N+1,0}^r$ and $G_{0,0}^r$ it is possible to calculate the transmission and reflection amplitudes t and r from left to right from the Fisher-Lee relation, Eq. (C.31) and (C.32).

Algorithm B.2. Compute $G_{0,0}^r, G_{N+1,0}^r, G_{0,N+1}^r, G_{N+1,N+1}^r$

A Use Eqs. (B.7), (B.8), and (B.9) and the recursion relation

$$G_{N+1,N+1}^{r,R(i-1)} = G_{N+1,N+1}^{r,R(i)} + G_{N+1,i}^{r,R(i)} H_{i,i-1} G_{i-1,N+1}^{r,R(i-1)} \quad (\text{B.13})$$

to compute $G_{i,i}^{r,R(i)}, G_{N+1,i}^{r,R(i)}, G_{i,N+1}^{r,R(i)}$ and $G_{N+1,N+1}^{r,R(i)}$ successively for $i = N \dots 1$.

B Compute $G_{0,0}^r, G_{N+1,0}^r, G_{0,N+1}^r$ and $G_{N+1,N+1}^r$ from Eqs. (B.11), (B.12) and the relations

$$G_{0,N+1}^r = G_{0,0}^r H_{0,1} G_{1,N+1}^{r,(1)} \quad (\text{B.14})$$

and

$$G_{N+1,N+1}^r = G_{N+1,N+1}^{r,R(1)} + G_{N+1,1}^{r,R(1)} H_{1,0} G_{0,0}^r. \quad (\text{B.15})$$

Applications: These four Green's functions are needed in order to calculate transmission both from left to right and, vice versa, from right to left. This is usually not necessary for two-terminal systems, where the transmission probabilities in both directions are connected by symmetries [59]. However, the computation of all four Green's functions is necessary for systems with multiple leads, where the left and right lead used in the RGF calculation may be composed of several real leads (see Section 4.4.2).

Algorithm B.3. Calculation of $G_{N+1,0}^r$ and $G_{i,i}^r$ for $i = 0 \dots N+1$.

A Use Eqs. (B.7) and (B.8) to compute $G_{i,i}^{r,R(i)}$ and $G_{N+1,i}^{r,R(i)}$ successively for $i = N \dots 1$.

¹²¹ For the calculation of densities it is enough to compute $G_{i,i}^<$, whereas for current densities additionally $G_{i,i\pm 1}^<$ is needed.

B Compute $G_{N+1,0}^r$ and $G_{0,0}^r$ from Eqs. (B.11) and (B.12).

C Compute $G_{i,i}^r$ successively for $i = 1 \dots N + 1$ from the relation

$$G_{i+1,i+1}^r = G_{i+1,i+1}^{r,R(i+1)} + G_{i+1,i+1}^{r,R(i+1)} H_{i+1,i} G_{i,i}^r H_{i,i+1} G_{i+1,i+1}^{r,R(i+1)} \quad (\text{B.16})$$

Applications: Computing these Green's functions allows at the same time a calculation of transport properties (transmission and reflection amplitudes from $G_{N+1,0}^r$ and $G_{0,0}^r$) and the local density of states, Eq. (2.55) (from $G_{i,i}^r$).

Algorithm B.4. Compute $G_{i,0}^r$ and $G_{i,N+1}^r$ for $i = 1 \dots N + 1$.

A Compute $G_{i,i}^{r,R(i)}$ and $G_{i,N+1}^{r,R(i)}$ from Eqs. (B.7) and (B.9), successively from $i = N \dots 1$.

B Compute $G_{0,0}^r$ and $G_{0,N+1}^r$ from Eqs. (B.11) and (B.14).

C Compute $G_{i,0}^r$ and $G_{i,N+1}^r$ successively for $i = 1 \dots N + 1$ from the relations

$$G_{i+1,0}^r = G_{i+1,i+1}^{r,R(i+1)} H_{i+1,i} G_{i,0}^r \quad (\text{B.17})$$

and

$$G_{i+1,N+1}^r = G_{i+1,i+1}^{r,R(i+1)} + G_{i+1,i+1}^{r,R(i+1)} H_{i+1,i} G_{i,N+1}^r \quad (\text{B.18})$$

Applications: $G_{i,0}^r$ and $G_{i,N+1}^r$ can be used for calculating transport (since they include $G_{0,0}^r$ and $G_{0,N+1}^r$) and for evaluating the lesser Green's function $G^<$ from Eq. (B.10).

19.9.3 B. 3 Computational complexity and implementation

The recursive Green's function algorithms B. 1 - B. 4 involve multiplications and inversions of $M_i \times M_i$ matrices, where M_i is the size of the matrix block $H_{i,i}$. Both multiplication and inversion scale cubically with the matrix size, i.e. as M_i^3 . Each of the algorithms performs a constant number of operations for block i , and hence the computational complexity of all these algorithms is $\propto \sum_{i=0}^{N+1} M_i^3$. Thus, all algorithms scale equally with the system size, although of course the absolute computational time depends on the number of operations per block.

In contrast to the computational complexity that is the same for all algorithms, the storage requirements of these algorithms is quite different. In the algorithms B. 1 and B. 2 we need the values of block i only to calculate block $i - 1$. As this is done successively, i.e. the algorithms pass through the structure once from right to left, they require $\propto \max(M_i)^2$ memory. In contrast, the algorithms B. 3 and B. 4 additionally require the value of block i in order to calculate the full Green's function at i . In other words, these algorithms first pass through the structure from right to left and then back again from left to right. As a consequence, we need to store the intermediate results of the first pass for the second pass. Thus, algorithms B. 3 and B. 4 require $\propto \sum_{i=0}^{N+1} M_i^2$ memory. If the lack of memory becomes an issue, it would be possible to reduce the storage requirements using a technique similar to the "unknitting" step in Ref. [146].

Finally, a few words about the implementation. Once the Hamiltonian is in blocktridiagonal form, the RGF algorithms only involve straightforward matrix algebra. In order to implement the RGF method, we can therefore employ highly optimized linear algebra libraries. Nowadays, the LAPACK and BLAS interfaces [107] have become the standard in terms of linear algebra routines. Using these interfaces allows to use the most optimized linear algebra routines available for a given machine. Since the computational time is almost exclusively spent in linear algebra calculations, the performance gain through the use of an optimized library can be huge [155]. On standard workstations we therefore employ the optimized ATLAS libraries [145].

19.10 Appendix C Generalized Fisher-Lee relation

19.10.1 C. 1 Landauer-Büttiker formalism

A very intuitive approach to the transport problem, pioneered by Landauer, connects the conductance of a device to a scattering problem [88, 89, 231]. Here we will formulate the scattering problem in the language of a tight binding model.

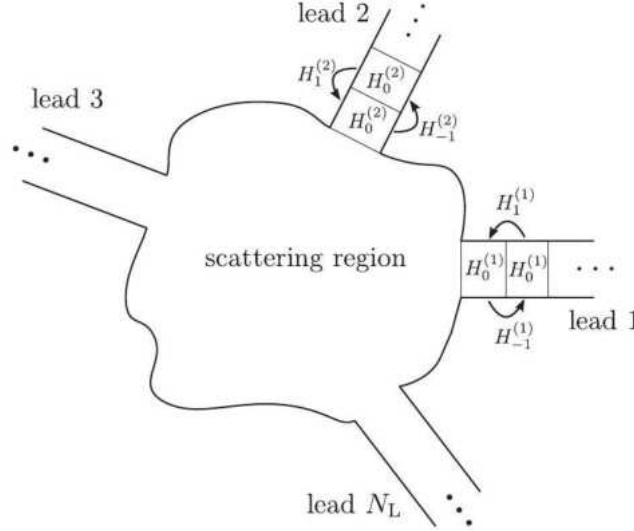


Figure C. 1 - Schematic view of a scattering region connected to N_L leads. Note that the lead unit cells are labeled such that index 0 corresponds to the unit cell adjacent to the scattering region and increases when moving further into the lead.

Suppose a scattering region connected to N_L leads, as depicted in Fig. C.1. As in Chapter 3, we describe lead l as a semi-infinite repetition of unit cells, such that the Hamiltonian of the isolated lead is given by

$$H_{\text{semi-infinite}}^{(l)} = \sum_{j=0}^{\infty} H_0^{(l)} |j\rangle^{(l)} \langle j|^{(l)} + H_1^{(l)} |j\rangle^{(l)} \langle j+1|^{(l)} + H_{-1}^{(l)} |j\rangle^{(l)} \langle j-1|^{(l)}, \quad (\text{C.1})$$

where $H_0^{(l)}$ is the unit cell matrix and $H_{1,-1}^{(l)}$ are the hopping matrices. Note that we label the unit cells such that 0 denotes the unit cell adjacent to the scattering region, and the unit cell index increases as we move away from the scattering region (see Fig. C.1). The asymptotic eigenstates $\varphi_n^{(l)}$ for a given energy E and far away from the scattering region are given by the eigenstates of the corresponding infinite wire. They can be written in the form

$$\varphi_n^{(l)}(j) = \frac{1}{\sqrt{|v_n|}} \phi_n^{(l)} e^{ik_n j} \quad (\text{C.2})$$

where $\phi_n^{(l)}$ is normalized to unity and obeys the equation

$$(E - H_0^{(l)} - H_1^{(l)} e^{ik_n} - H_{-1}^{(l)} e^{-ik_n}) \phi_n^{(l)} = 0 \quad (\text{C.3})$$

and v_n is the velocity of the eigenstate, Eq. 3.19. The asymptotic eigenstates $\varphi_n^{(l)}$ are called the channels of lead l . As in Chapter 3, we sort the channels according to their velocity v_n . For $v_n > 0$ we have a channel leaving the scattering region into the lead (outgoing, " $<$ "), whereas for $v_n < 0$ the channel enters the scattering region from the lead (incoming, " $>$ "). Note that we choose the notation " $>$ " and " $<$ " just as in Chapter 3.

A scattering-wave state $\Psi_n^{(l)}$ for a given energy E originates from an incoming channel $\varphi_{n,>}^{(l)}$ and is scattered into outgoing channels in the different leads attached to the scattering region. In

the asymptotic region, where all evanescent waves have decayed, the scattering state can be written as

$$\Psi_n^{(l)}(j) = \begin{cases} \varphi_n^{(l)}(j) + \sum_m r_{ll,mn} \varphi_{m,<}^{(l)}(j) & \text{if } j \text{ in lead } l \\ \sum_m t_{l',mn} \varphi_{m,<}^{(l')}(j) & \text{if } j \text{ in lead } l' \end{cases} \quad (\text{C.4})$$

With the choice of wave function normalization for the channels in Eq. C.2, the transmission and reflection coefficients $t_{l',mn}$ and $r_{ll,mn}$ form a unitary scattering matrix [231]. The conductance coefficients for leads l' and l are then given by the Landauer-Büttiker formula (for further introductory reading, see Refs. [59, 60]):

$$G_{l'l'} = \frac{e^2}{h} \sum_{n,m} |t_{l',mn}|^2. \quad (\text{C.5})$$

The problem of transport is thus reduced to calculating the scattering wave functions. However, very often it is more convenient to calculate the Green's function instead of the scattering wave functions itself. We are then left with the problem to obtain the transmission coefficients from the Green's function alone. Relations for the transmission or reflection coefficients and the Green's function of the system are called Fisher-Lee relations [90]. Here we will derive a generalized Fisher-Lee relation valid for any tight-binding Hamiltonian.

To this end we will closely follow Ref. [154], that derived Fisher-Lee relations for electrons in a magnetic field. The steps there carry over to arbitrary tight-binding Hamiltonians. Note that the problem of extracting the transmission coefficients from the Green's function has also been addressed in Ref. [82], utilizing the inverse matrix of eigenvectors ($U_{<}^{-1}$ in Chapter 3). The relations derived in this chapter avoid inverting the eigenvector matrix which can be problematic, as shown in Chapter 3. Furthermore, the expressions are more along the lines of the well-known Fisher-Lee relations for simple systems [59, 60, 90, 231], and will provide some additional physical insight.

The derivation of the generalized Fisher-Lee relation is complicated by the fact that the transverse wave functions of different channels need not be orthogonal any more, as discussed in Section 3.1. If the transverse wave functions were orthogonal, the transmission and reflection coefficients could be simply extracted from the scattering wave function by projection. To remedy the absence of a simple orthogonality relation, we derive a generalized orthogonality relation for the transverse wave functions ϕ_n in the next section. Then, we will construct the scattering wave function from the Green's function and derive expressions for the transmission and reflection coefficients. Finally, we emphasize the equivalence between the scattering formalism and the non-equilibrium Green's function formalism by showing how central results of both formalisms are connected.

19.10.2 C. 2 Generalized orthogonality relations

We now proceed to derive a generalized orthogonality relation for the transverse wave functions of propagating and decaying asymptotic eigenstates $\mathbf{u}_n$.¹ Note that the expressions derived here are valid within a single lead l , and we will therefore drop the lead index " l " in this section in order to simplify notation.

By combining Eq. (3.14a) and its complex conjugate,

$$\begin{aligned} (E - H_0) \lambda_n \mathbf{u}_n &= (H_1 \lambda_n^2 + H_{-1}) \mathbf{u}_n \text{ and} \\ \mathbf{u}_n^\dagger (E - H_0) \lambda_n^* &= \mathbf{u}_n^\dagger (H_{-1} (\lambda_n^*)^2 + H_1) \end{aligned} \quad (\text{C.6})$$

we find that

$$\mathbf{u}_{n'}^\dagger (H_1 \lambda_n^2 + H_{-1}) \mathbf{u}_n \lambda_{n'}^* = \mathbf{u}_{n'}^\dagger (H_{-1} (\lambda_{n'}^*)^2 + H_1) \mathbf{u}_n \lambda_n \quad (\text{C.7})$$

and thus

$$\mathbf{u}_{n'}^\dagger (H_1 \lambda_n - H_{-1} \lambda_{n'}^*) \mathbf{u}_n (1 - \lambda_n \lambda_{n'}^*) = 0 \quad (\text{C.8})$$

This equation is trivially fulfilled if $\lambda_n \lambda_{n'}^* = 1$. This is the case for propagating modes with $\lambda_n = \lambda_{n'} = e^{ik_n}$. If $n = n'$, the expression in the first bracket is then up to a factor the velocity v_n of the propagating mode. For $n \neq n'$, i.e. for degenerate propagating modes, we have shown in Section 3.2.3, that the eigenmodes ϕ_n can be chosen such that they diagonalize the expression in the first bracket. In the case of decaying modes, the equation is trivially fulfilled, if $\lambda_n = 1/\lambda_{n'}^*$, for pairs of right- (" < ") and left-decaying (" > ") modes. In all other cases, the expression in the first bracket must be zero. In summary, we find

$$i \mathbf{u}_{n'}^\dagger (H_1 \lambda_n - H_{-1} \lambda_{n'}^*) \mathbf{u}_n = \begin{cases} \hbar v_n & \text{for propagating modes with } n = n' \\ \text{const.} & \text{for pairs of decaying modes with } \lambda_n = 1/\lambda_{n'}^* \\ 0 & \text{else} \end{cases} \quad (\text{C.9})$$

where we have introduced a factor i to make the expression Hermitian. First, this proves that the propagating modes ϕ_n diagonalize the velocity operator, as stated in Section 3.2.3.² In addition, we are also provided with relations between propagating and decaying modes. In order to utilize these relations, we introduce the operator

$$K(j) = \frac{i}{\hbar} (H_1 |j\rangle \langle j+1| - H_{-1} |j+1\rangle \langle j|). \quad (\text{C.10})$$

The full wave function of a decaying or propagating asymptotic state can be written in a compact form as $\xi_n(j) = c \mathbf{u}_n \lambda_n^j$, where c is an appropriate normalization constant ($c = 1/\sqrt{|v_n|}$ for propagating modes). Note that we reserve the notation $\varphi_n(j)$ for propagating modes. Then we find the generalized orthogonality relations:

$$\langle \xi_{n,<(>)} | K(j) | \xi_{n',<(>)} \rangle = \begin{cases} \delta_{nn'} & \text{for a propagating mode} \\ 0 & \text{else} \end{cases} \quad (\text{C.11a})$$

and in particular for the propagating modes:

$$\langle \varphi_{n,<(>)} | K(j) | \varphi_{n',<(>)} \rangle = \delta_{nn'} \quad (\text{C.11b})$$

$$\langle \varphi_{n,>(<)} | K(j) | \xi_{n',<(>)} \rangle = 0 \quad (\text{C.11c})$$

regardless of j .

For the purposes of implementing these operator expressions in a computer program, it is useful to rewrite them in compact matrix form. For this, note that from the expressions for the surface Green's function in Section 3.2.4 we obtain:

¹²¹ Remember that $\mathbf{u}_n$ denotes all eigenstates, decaying and propagating, whereas ϕ_n denotes only propagating eigenstates.

¹²² Note furthermore, that if the transverse wave functions ϕ_n of the propagating modes would not diagonalize the velocity operator, the Landauer-Büttiker formula (C.5) could also contain terms of the form $t_{m_1 n_1} t_{m_2 n_2}^*$ with $m_1 \neq m_2$ and $n_1 \neq n_2$. This is an additional reason why in a numerical calculation degenerate propagating modes must be rotated so that the velocity operator is diagonal in this basis, as described in Section 3.2.3.

$$g_{\text{R}}^{\text{r}} H_{-1} \mathbf{u}_{n,<} = \lambda_{n,<} \mathbf{u}_{n,<} \quad \text{and} \quad (\text{C.12a})$$

$$\mathbf{u}_{n,<}^{\dagger} H_1 g_{\text{R}}^{\text{a}} = \lambda_{n,<}^* \mathbf{u}_{n,<}^{\dagger} \quad (\text{C.12b})$$

Expressions of the type $g_{\text{R}}^{\text{r}} H_{-1}$ essentially propagate a mode from one unit cell to the other and are thus commonly called transfer matrices [97].

Using the definition $\Gamma = i(\Sigma^{\text{r}} - \Sigma^{\text{a}})$ with the lead self-energies $\Sigma^{\text{r(a)}} = H_1 g_{\text{R}}^{\text{r(a)}} H_{-1}$, we find

$$\begin{aligned} \mathbf{u}_{n',<} \Gamma \mathbf{u}_{n,<} &= i \mathbf{u}_{n',<} (H_1 g_{\text{R}}^{\text{r}} H_{-1} - H_1 g_{\text{R}}^{\text{a}} H_{-1}) \mathbf{u}_{n,<} \\ &= i \mathbf{u}_{n',<} (H_1 \lambda_{n,<} - H_{-1} \lambda_{n',<}^*) \mathbf{u}_{n,<} \\ &= \begin{cases} \hbar |v_n| \delta_{nn'} & \text{for propagating modes,} \\ 0 & \text{else.} \end{cases} \end{aligned} \quad (\text{C.13})$$

The matrix Γ thus plays the role of a velocity or current operator within the lead. In an analogous fashion, by utilizing the expressions for g_{R}^{a} , we can derive a similar expression for modes decaying to the right and propagating in the opposite direction.

The matrix Γ is a central object that appears in many expressions of the nonequilibrium Green's function formalism (see Section 2.5.2). In order to make connections between this formalism and the scattering theory, we collect some useful identities for Γ .

First, we rewrite the expression for the transfer matrix $g_{\text{R}}^{\text{r}} H_{-1}$ as

$$g_{\text{R}}^{\text{r}} H_{-1} = U_{<} \Lambda_{<} U_{<}^{-1} = \sum_n \mathbf{u}_{n,<} \lambda_{n,<} \tilde{\mathbf{u}}_{n,<}^{\dagger}, \quad (\text{C.14})$$

where $\tilde{\mathbf{u}}_{n,<}^{\dagger}$ is the n -th row of the inverse eigenvector matrix $U_{<}^{-1}$. Inserting this expression in the definition of the matrix Γ we obtain

$$\begin{aligned} \Gamma &= i (H_1 g_{\text{R}}^{\text{r}} H_{-1} - H_1 g_{\text{R}}^{\text{a}} H_{-1}) \\ &= i \left(H_1 U_{<} \Lambda_{<} U_{<}^{-1} - (U_{<}^{\dagger})^{-1} \Lambda_{<}^{\dagger} U_{<}^{\dagger} H_{-1} \right) \\ &= i \sum_n \left(H_1 \mathbf{u}_{n,<} \lambda_{n,<} \tilde{\mathbf{u}}_{n,<}^{\dagger} - \tilde{\mathbf{u}}_{n,<} \lambda_{n,<}^* \mathbf{u}_{n,<}^{\dagger} H_{-1} \right) \end{aligned} \quad (\text{C.15})$$

where we utilized $(g_{\text{R}}^{\text{r}} H_{-1})^{\dagger} = H_1 g_{\text{R}}^{\text{a}}$. Now we insert unity at the appropriate places, $\mathbb{1} = U_{<} U_{<}^{-1} = \sum_m \mathbf{u}_{m,<} \tilde{\mathbf{u}}_{m,<}^{\dagger}$ and $\mathbb{1} = (U_{<}^{\dagger})^{-1} U_{<}^{\dagger} = \sum_m \tilde{\mathbf{u}}_{m,<} \mathbf{u}_{m,<}^{\dagger}$. Then we find

$$\begin{aligned} \Gamma &= i \sum_n \left(\mathbb{1} H_1 \mathbf{u}_{n,<} \lambda_{n,<} \tilde{\mathbf{u}}_{n,<}^{\dagger} - \tilde{\mathbf{u}}_{n,<} \lambda_{n,<}^* \mathbf{u}_{n,<}^{\dagger} H_{-1} \mathbb{1} \right) \\ &= i \sum_{n,m} \left(\tilde{\mathbf{u}}_{m,<} \mathbf{u}_{m,<}^{\dagger} H_1 \mathbf{u}_{n,<} \lambda_{n,<} \tilde{\mathbf{u}}_{n,<}^{\dagger} - \tilde{\mathbf{u}}_{m,<} \lambda_{m,<}^* \mathbf{u}_{m,<}^{\dagger} H_{-1} \mathbf{u}_{n,<} \tilde{\mathbf{u}}_{n,<}^{\dagger} \right) \\ &= \sum_n \tilde{\phi}_{n,<} \hbar |v_{n,<}| \phi_{n,<}^{\dagger}, \end{aligned} \quad (\text{C.16})$$

where we used Eq. (C.9) and $|v_{n,<}| = v_{n,<}$. Furthermore, according to Eq. (C.9)

$$\tilde{\phi}_{n,<}^{\dagger} = \phi_{n,<}^{\dagger} \Gamma / (\hbar |v_{n,<}|) \quad (\text{C.17})$$

and we obtain the final expression

$$\Gamma = \sum_n \Gamma \phi_{n,<} \frac{1}{\hbar |v_{n,<}|} \phi_{n,<}^\dagger \Gamma. \quad (\text{C.18})$$

Only outgoing propagating modes $\phi_{n,<}$ contribute to the velocity/current operator Γ which comes at no surprise, since only the propagating modes carry current. From Eq. (C.17) and the eigendecomposition expression for the transfer matrix, we immediately obtain the useful relation

$$\phi_{n,<}^\dagger \Gamma g_R^r H_{-1} = \lambda_{n,<} \phi_{n,<}^\dagger \Gamma \quad (\text{C.19})$$

which is the complement to Eq. (C.12a).

In an analogous way, utilizing the eigenvector expression for the advanced surface Green's function, we can derive an expression for Γ only in terms of incoming propagating modes,

$$\Gamma = \sum_n \Gamma \phi_{n,>} \frac{1}{\hbar |v_{n,>}|} \phi_{n,>}^\dagger \Gamma. \quad (\text{C.20})$$

Note that $v_{n,>} < 0$ here. For the transfer matrices, we obtain

$$g_R^a H_{-1} \phi_{n,>} = \lambda_{n,>} \phi_{n,>} \quad \text{and} \quad (\text{C.21a})$$

$$H_1 g_R^r \Gamma \phi_{n,>} = \lambda_{n,>}^* \Gamma \phi_{n,>} \quad (\text{C.21b})$$

19.10.3 C. 3 Fisher-Lee relation

We now relate the scattering wave function $\Psi_n^{(l)}$ to the Green's function G^r . In order to keep the notation simple, we assume that the Hamiltonian of the full system can be written in a quasi-linear form,

$$H = \sum_{j=-\infty}^{\infty} H_{j,j} |j\rangle \langle j| + H_{j,j+1} |j\rangle \langle j+1| + H_{j,j-1} |j\rangle \langle j-1|. \quad (\text{C.22})$$

All blocks $j < 1$ are supposed to correspond to lead l , $|0\rangle = |0\rangle^{(l)}$, $|-1\rangle = |1\rangle^{(l)}$, $|-2\rangle = |2\rangle^{(l)}$, $\dots$, and all blocks $j > L$ are supposed to contain all other leads, $|L+1\rangle = \bigotimes_{l' \neq l} |0\rangle^{(l')}$, $|L+2\rangle = \bigotimes_{l' \neq l} |1\rangle^{(l')}$, $\dots$. The Hamiltonian H can always be written in such a form, and this form can be explicitly found by the algorithm presented in Chapter 4.

We now choose a finite region of integration $\mathcal{A}$ containing the scattering region and a sizeable part of the leads, such that the boundary of $\mathcal{A}$ is in the asymptotic region of the

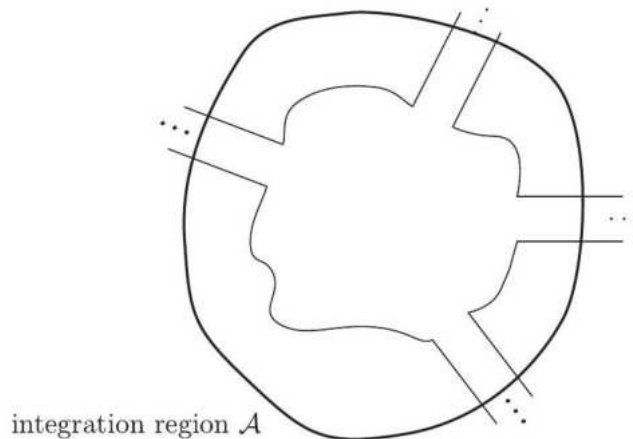


Figure C. 2 - The integration region $\mathcal{A}$ includes the scattering region and parts of the leads up to the asymptotic region.

leads, as schematically depicted in Fig. C.2. Using $G^r(E - H) = \mathbb{K}$ and $(E - H)\Psi_n^{(l)} = 0$ we can write

$$\Psi_n^{(l)}(\mathbf{x}) = \sum_{j \in \mathcal{A}} \langle \mathbf{x} | G^r(E - H) | j \rangle \langle j | \Psi_n^{(l)} \rangle - \langle \mathbf{x} | G^r | j \rangle \langle j | (E - H) | \Psi_n^{(l)} \rangle, \text{ for } \mathbf{x} \in \mathcal{A}. \quad (\text{C.23})$$

Using the quasi-linear form of H , Eq. (C.22), and writing accordingly $\sum_{j \in \mathcal{A}}$ as $\sum_{j=z_1}^{z_2}$, where $z_1 < 0$ and $z_2 > L + 1$ are blocks corresponding to the boundaries of $\mathcal{A}$, we find that

$$\begin{aligned} \Psi_n^{(l)}(\mathbf{x}) = & \langle \mathbf{x} | G^r | z_2 \rangle H_{z_2, z_2+1} \langle z_2 + 1 | \Psi_n^{(l)} \rangle - \\ & \langle \mathbf{x} | G^r | z_2 + 1 \rangle H_{z_2+1, z_2} \langle z_2 | \Psi_n^{(l)} \rangle + \\ & \langle \mathbf{x} | G^r | z_1 \rangle H_{z_1, z_1-1} \langle z_1 - 1 | \Psi_n^{(l)} \rangle - \\ & \langle \mathbf{x} | G^r | z_1 - 1 \rangle H_{z_1-1, z_1} \langle z_1 | \Psi_n^{(l)} \rangle. \end{aligned} \quad (\text{C.24})$$

The sum over all lattice points within the area $\mathcal{A}$ has thus been reduced to a sum involving only contributions from the surface. Hence, this can be understood as a discrete version of Green's theorem. Using the definition from Eq. (C.10) we finally arrive at

$$\Psi_n^{(l)}(\mathbf{x}) = \sum_{l' \neq l} \frac{\hbar}{i} \langle \mathbf{x} | G^r K^{(l')} (z_2) | \Psi_n^{(l)} \rangle + \frac{\hbar}{i} \langle \mathbf{x} | G^r K^{(l)} (z_1) | \Psi_n^{(l)} \rangle \quad (\text{C.25})$$

The expression for the scattering wave function $\Psi_n^{(l)}$ thus involves the operator $K^{(l)}$ for which we derived generalized orthogonality relations in the previous section. To simplify this expression further, we consider the behavior of the Green's function in the asymptotic region of a lead, $\lim_{z \rightarrow \pm\infty} \langle \mathbf{x} | G^r | z \rangle$. Here, $\lim_{z \rightarrow \pm\infty}$ should be understood as taking z into the asymptotic region of lead l ($z \rightarrow -\infty$) and $l' \neq l$ ($z \rightarrow +\infty$). The Green's function $\lim_{z \rightarrow \pm\infty} \langle \mathbf{x} | G^r | z \rangle$ contains only incoming waves $\varphi >$ (for a detailed discussion see Ref. [154]). Considering further that the scattering wave function $\Psi_n^{(l)}(z)$ contains incoming waves only in lead l , by virtue of the generalized orthogonality relations all surface terms except for the source term in lead l vanish: ³

$$\begin{aligned} \Psi_n^{(l)}(\mathbf{x}) = & \lim_{z_1 \rightarrow -\infty} \langle \mathbf{x} | G^r | z_1 \rangle^{(l)} H_1^{(l)} \langle z_1 - 1 |^{(l)} | \Psi_n^{(l)} \rangle - \\ & \langle \mathbf{x} | G^r | z_1 - 1 \rangle^{(l)} H_{-1}^{(l)} \langle z_1 |^{(l)} | \Psi_n^{(l)} \rangle \\ = & \lim_{z_1 \rightarrow -\infty} \left(G_{\mathbf{x}, z_1}^r H_1^{(l)} \frac{1}{\sqrt{|v_{n, >}|}} \phi_{n, >}^{(l)} e^{-ik_{n, >}(z_1-1)} - \right. \\ & \left. G_{\mathbf{x}, z_1+1}^r H_{-1}^{(l)} \frac{1}{\sqrt{|v_{n, >}|}} \phi_{n, >}^{(l)} e^{-ik_{n, >}} \right) \\ = & \lim_{z_1 \rightarrow -\infty} G_{\mathbf{x}, z_1}^r \left(H_1 g_R^{a, (l)} H_{-1} - H_1 g_R^{r, (l)} H_{-1} \right) \frac{1}{\sqrt{|v_{n, >}|}} \phi_{n, >}^{(l)} e^{-ik_{n, >, z_1}}, \end{aligned} \quad (\text{C.26})$$

where we used Eqs. (C.21a) and (3.8). Using the definition of Γ_l we thus arrive finally at

$$\Psi_n^{(l)}(\mathbf{x}) = \lim_{z_1 \rightarrow -\infty} \frac{i}{\sqrt{|v_{n, >}|}} e^{-ik_{n, >, z_1}} G_{\mathbf{x}, 0}^r \Gamma_l \phi_{n, >}^{(l)} \quad (\text{C.27})$$

valid for any point $\mathbf{x} \in \mathcal{A}$. This expression is thus a viable way to calculate the full scattering wave function in the scattering region from the Green's function G^r .

As usual [90, 154, 231], the relation between the scattering wave function and the Green's function involves values of G^r in the asymptotic region, i.e. far from the scattering region. For numerics, this is rather inconvenient, as it implies including large parts of the leads in the calculation, increasing computing time. It would therefore be desirable to obtain an expression involving only the Green's function within the first few layers of the lead. For this, note that from Eq. (3.8) we can write

$$G_{z_2, z_1}^r = G_{z_2, z_1'}^r \left(H_1 g_R^{r, (l)} \right)^{z_1' - z_1} \quad (\text{C.28})$$

whenever z_2 is not further in lead l than z_1' . In particular, we find restricting x to be contained in a block $j \geq 0$

$$\Psi_n^{(l)}(\mathbf{x}) = \lim_{z_1 \rightarrow -\infty} \frac{i}{\sqrt{|v_{n, >}|}} e^{-ik_{n, >}} G_{\mathbf{x}, 0}^r \left(H_1 g_R^{r, (l)} \right)^{-z_1} \Gamma_l \phi_{n, >}^{(l)} \quad (\text{C.29})$$

Since according to Eq. (C.21b) $\Gamma_l \phi_{n, >}^{(l)}$ is an eigenvector of the transfer matrix $H_1 g_R^{r, (1)}$ with eigenvalue $\lambda_{n, >}^* = e^{-ik_{n, >}}$, we finally arrive at

$$\Psi_n^{(l)}(\mathbf{x}) = \frac{i}{\sqrt{|v_{n, >}|}} G_{\mathbf{x}, 0}^r \Gamma_l \phi_{n, >}^{(l)}. \quad (\text{C.30})$$

This expression can be used directly for numerical computations.

In order to calculate the transmission and reflection coefficients we do not need the scattering wave function in the scattering region, but only in the respective leads. In the asymptotic region, the scattering wave function $\Psi_n^{(l)}$ contains only propagating modes, as indicated by Eq. (C.4). However, as argued above, for numerical purposes it is rather inconvenient to use the asymptotic region.

Closer to the scattering region, $\Psi_n^{(l)}$ also contains evanescent waves, decaying away in the direction of the lead. In the previous section we have shown that $\phi_{m, >}^{(l')}$ is orthogonal to both propagating and decaying modes with respect to the current operator $\Gamma_{l'}$. Therefore, we can use this generalized orthogonality relation to project out the desired scattering coefficients from the scattering wave function $\Psi_n^{(l)}$ at any point of the leads.

We first consider the case that $\mathbf{x}$ is in lead $l' \neq l$. From $\left(\phi_{m, >}^{(l')} \right)^\dagger \Gamma_{l'} \Psi_n^{(l)}$ we find

$$t_{l', mn} = \frac{i}{\hbar} \frac{1}{\sqrt{|v_{m, <}| |v_{n, >}|}} \left(\phi_{m, <}^{(l')} \right)^\dagger \Gamma_{l'} G_{L+1, 0}^r \Gamma_l \phi_{n, >}^{(l)}. \quad (\text{C.31})$$

In the case of $l' = l$, we obtain the reflection coefficients. In this case it is necessary to keep in mind that incoming and outgoing waves need not be orthogonal with respect to Γ_l , and we obtain

$$r_{ll, mn} = \frac{1}{\hbar \sqrt{|v_{m, <}| |v_{n, >}|}} \left(i \left(\phi_{m, <}^{(l)} \right)^\dagger \Gamma_l G_{0, 0}^r \Gamma_l \phi_{n, >}^{(l)} - \left(\phi_{m, <}^{(l)} \right)^\dagger \Gamma_l \phi_{n, >}^{(l)} \right). \quad (\text{C.32})$$

These general expressions for the transmission and reflection coefficients are very similar to the Fisher-Lee relations derived for the particular case of electrons in a twodimensional electron gas in textbooks [59, 60]. There, in- and outgoing modes have the same transverse wave function, $\phi_{n, <} = \phi_{n, >} = \phi_n$, and the transverse wave functions are eigenstates to the current operator, $\Gamma \phi_n = \hbar v_n \phi_n$. In general, this is not the case, and one has to resort to the formulas derived here. Examples where this is the case are electrons in a magnetic field, dealt with in Chapter 3 of this work, and zigzag graphene nanoribbons, Chapters 6 and 7.

¹²³ When reading the formulas remember that $|z_1\rangle = |-z_1\rangle^{(l)}$ in our notation. Therefore, additional "-" signs appear with respect to the definition of the scattering wave-function $\Psi_n^{(l)}$ in Eq. (C.4)

19.10.4 C. 4 Equivalence of scattering and non-equilibrium Green's function formalism

The generalized Fisher-Lee relation allows to relate central results from the nonequilibrium Green's function (NEGF) formalism with concepts from the Landauer-Büttiker formalism based on a scattering problem.

In the NEGF formalism, Eq. (2.119), we obtained the conductance between two leads l and l' in linear response as

$$G_{ll'} = \frac{e^2}{h} \text{Tr} (\Gamma_l G^r \Gamma_{l'} G^a). \quad (\text{C.33})$$

Expressing the current operators $\Gamma^{(l),(l')}$ in terms of in- and outgoing modes, Eqs. (C.18) and (C.20) we find

$$\begin{aligned} G_{ll'} &= \frac{e^2}{h} \text{Tr} (\Gamma_l G^r \Gamma_{l'} G^a) \\ &= \frac{e^2}{h} \text{Tr} \left(\sum_n \Gamma_l \phi_{n,<}^{(l)} \frac{1}{\hbar |v_{n,<}|} \left(\phi_{n,<}^{(l)} \right)^\dagger \Gamma_l G^r \sum_m \Gamma_{l'} \phi_{m,>}^{(l')} \frac{1}{\hbar |v_{m,>}|} \left(\phi_{m,>}^{(l')} \right)^\dagger \Gamma_{l'} G^a \right) \\ &= \frac{e^2}{h} \sum_{m,n} \frac{1}{\hbar^2 |v_{n,<}| |v_{m,>}|} \left(\left(\phi_{n,<}^{(l)} \right)^\dagger \Gamma_l G^r \Gamma_{l'} \phi_{m,>}^{(l')} \right) \left(\left(\phi_{m,>}^{(l')} \right)^\dagger \Gamma_{l'} G^a \Gamma_l \phi_{n,<}^{(l)} \right) \\ &= \frac{e^2}{h} \sum_{m,n} |t_{ll',nm}|^2, \end{aligned} \quad (\text{C.34})$$

where we made use of the Fisher-Lee relation (C.31). Thus, the NEGF expression for the conductance in linear response is equivalent to the Landauer-Büttiker formula.

Another important relation from the NEGF formalism is the expression for the lesser Green's function from the Keldysh equation, Eq. (2.95),

$$G^< = G^r \Sigma^< G^a. \quad (\text{C.35})$$

In the case of non-interacting electrons, or for Coulomb interaction in Hartree-Fock approximation, $\Sigma^<$ only contains contributions from the leads and can be written as

$$\Sigma^<(E) = \sum_l i \Gamma_l f(E, \mu_l) \quad (\text{C.36})$$

where μ_l is the electrochemical potential in lead l . The contribution of a single lead l to $G^<$ is given by

$$\begin{aligned} G^{<,(l)}(\mathbf{x}, \mathbf{x}') &= G^r i \Gamma_l f(E, \mu_l) G^a \\ &= f(E, \mu_l) \sum_n i G^r \Gamma_l \phi_{n,>}^{(l)} \frac{1}{\hbar |v_{n,>}|} \left(\phi_{n,>}^{(l)} \right)^\dagger \Gamma_l G^a \\ &\propto f(E, \mu_l) \sum_n i \Psi_n^{(l)}(\mathbf{x}) \left(\Psi_n^{(l)}(\mathbf{x}') \right)^* \frac{1}{\hbar |v_{n,>}|}. \end{aligned} \quad (\text{C.37})$$

Note that we neglected possible constant terms in the last line, in order to avoid discussing the appropriate normalization of the scattering wave functions. The term $1/|v_{n,>}|$ is up to a constant the one-dimensional density of states of the incoming channel n in lead l [89], and we denote this density of states as $D_n^{1d,(l)}(E)$. The full lesser Green's function thus reads

$$G^<(\mathbf{x}, \mathbf{x}', E) \propto \sum_l f(E, \mu_l) \sum_n i \Psi_n^{(l)}(\mathbf{x}) \left(\Psi_n^{(l)}(\mathbf{x}') \right)^* D_n^{1d, (l)}(E) \quad (\text{C.38})$$

Hence, the lesser Green's function, and thus in particular the electron density for a given energy E , is composed of all scattering states, weighted with the respective density of states of the incoming channel and the Fermi function of the respective lead.

19.11 Appendix D Details of the derivation of the lead Green's function

19.11.1 D. $1\lambda(E + i\eta)$ for propagating modes

In the derivation of the general expression for the surface Green's function of a lead it is necessary to determine which poles contribute to the contour integral in Eq. (3.12). That is, we need to determine whether the eigenvalues λ_n of the quadratic eigenvalue equation

$$((E + i\eta - H_0) \lambda_n - H_1 \lambda_n^2 - H_{-1}) \mathbf{u}_n = 0 \quad (\text{D.1})$$

are located inside or outside the unit circle. For any eigenvalue with $|\lambda_n| \neq 1$ this can be done without considering the infinitesimal shift $i\eta$. However, in the case of a propagating mode $\mathbf{u}_n = \boldsymbol{\phi}_n$, we find that the eigenvalue $\lambda_n = e^{ik}$ and thus lies directly on the unit circle, if we neglect the infinitesimal shift $i\eta$. Therefore, it is vital to include $i\eta$ in the calculation for this case.

Numerically, this could be done by explicitly including some small but final imaginary part of the energy E in the calculation [102]. However, this just introduces a parameter into the numerical calculation that might need manual adjustment in order to achieve convergence. Therefore we will show in the following, how we can determine $\lambda_n(E + i\eta)$ for the solution including the infinitesimal shift $i\eta$ from the properties of the solution $\lambda_n(E), \boldsymbol{\phi}_n(E)$ of the quadratic eigenvalue problem without the infinitesimal $i\eta$:

$$((E - H_0) \lambda_n(E) - H_1 \lambda_n(E)^2 - H_{-1}) \boldsymbol{\phi}_n(E) = 0. \quad (\text{D.2})$$

In the course of this derivation we will also gain some insight into the arguments used in the derivation of the lead Green's function in Refs. [59, 82, 86] and justify these arguments mathematically.

In order to determine the solution including the infinitesimal shift $i\eta$ we use the Taylor expansion

$$\lambda_n(E + i\eta) = \lambda_n(E) + \left. \frac{d\lambda_n}{dE} \right|_E i\eta + \mathcal{O}(\eta^2). \quad (\text{D.3})$$

Note that a similar approaches have been used in Refs. [94, 100, 103]. In order to calculate $\frac{d\lambda_n}{dE}$ we take the derivative of Eq. (D.2) with respect to the energy:

$$\begin{aligned} & \left(\lambda_n(E) + (E - H_0) \frac{d\lambda_n}{dE} - H_1 2\lambda_n(E) \frac{d\lambda_n}{dE} \right) \boldsymbol{\phi}_n(E) + \\ & ((E - H_0) \lambda_n(E) - H_1 \lambda_n(E)^2 - H_{-1}) \frac{d\boldsymbol{\phi}_n}{dE} = 0 \end{aligned} \quad (\text{D.4})$$

Using the fact that, as shown in Section 3.2.2, $\boldsymbol{\phi}_n^\dagger$ is also a right eigenvector of the quadratic eigenvalue problem Eq. (D.2), we can get rid of the second term of Eq. (D.4) and find

$$\phi_n^\dagger(E) \left(\lambda_n(E) + (E - H_0) \frac{d\lambda_n}{dE} - H_1 2\lambda(E) \frac{d\lambda_n}{dE} \right) \phi(E) = 0. \quad (\text{D.5})$$

Further exploiting the fact that the ϕ_n are normalized, $\phi_n^\dagger \phi_n = 1$, we arrive at the expression

$$\left. \frac{d\lambda_n}{dE} \right|_E = \frac{\lambda_n(E)}{2\phi_n^\dagger H_1 \phi_n \lambda_n(E) - \phi_n^\dagger (E - H_0) \phi_n} \quad (\text{D.6})$$

To determine whether $\lambda_n(E + i\eta)$ lies inside or outside the unit circle, we need to calculate $|\lambda_n(E + i\eta)|$. Using the Taylor expansion (D.3) we find

$$\begin{aligned} |\lambda_n(E + i\eta)| &= \left| \lambda_n(E) + \frac{d\lambda_n}{dE} i\eta + \mathcal{O}(\eta^2) \right| \\ &= \left(\lambda_n(E) + \frac{d\lambda_n}{dE} i\eta + \mathcal{O}(\eta^2) \right) \left(\lambda_n^*(E) - \left(\frac{d\lambda_n}{dE} \right)^* i\eta + \mathcal{O}(\eta^2) \right) \\ &= |\lambda_n(E)|^2 + i\eta \left(\lambda_n^*(E) \frac{d\lambda_n}{dE} - \lambda_n(E) \left(\frac{d\lambda_n}{dE} \right)^* \right) + \mathcal{O}(\eta^2) \\ &= 1 - 2\eta \Im \left(\lambda_n^*(E) \frac{d\lambda_n}{dE} \right) + \mathcal{O}(\eta^2) \end{aligned} \quad (\text{D.7})$$

Using Eq. (D.6) we find

$$\begin{aligned} \lambda_n^*(E) \frac{d\lambda_n}{dE} &= \frac{\lambda_n^*(E) \lambda_n(E)}{2\phi_n^\dagger H_1 \phi_n \lambda_n(E) - \phi_n^\dagger (E - H_0) \phi_n} \\ &= \underbrace{\frac{1}{2\phi_n^\dagger H_1 \phi_n \lambda_n(E)}}_{(1)} - \underbrace{\frac{\phi_n^\dagger (E - H_0) \phi_n}{2\phi_n^\dagger H_1 \phi_n \lambda_n(E)}}_n. \end{aligned} \quad (\text{D.8})$$

Since for real energy E and only for real energies we can have propagating modes the matrix $E - H_0$ is hermitian, term (2) in the equation above is real, and only term (1) contributes to the imaginary part. Furthermore, note that the imaginary part of term (1) is related to the velocity of the propagating mode according to Eq. (3.19). Combining this information we finally arrive at

$$\Im \left(\lambda_n^*(E) \frac{d\lambda_n}{dE} \right) = - \frac{\Im (2\phi_n^\dagger H_1 \phi_n \lambda_n(E))}{\underbrace{|\dots|}_{>0}} = \frac{\hbar v_n(E)}{|\dots|}, \quad (\text{D.9})$$

and thus

$$|\lambda_n(E + i\eta)| = 1 - 2\hbar \underbrace{\frac{\eta}{|\dots|}}_{>0} v_n(E) + \mathcal{O}(\eta^2). \quad (\text{D.10})$$

In summary we find that for a propagating mode $\phi_n(E) \lambda_n(E + i\eta)$ lies

- inside the unit circle, if $v_n(E) > 0$, i.e. for right-moving modes,
- inside the unit circle, if $v_n(E) < 0$, i.e. for left-moving modes.

This criterion only fails for $v_n = 0$, i.e. directly at a band edge. However, $v_n = 0$ also implies a diverging density of states and thus impedes a numerical calculation of the Green's function anyways. In such a situation it is advisable to move away a little bit from the band edge.

Refs. [59, 82, 86] directly invoke the concept of left- and right-going modes in order to fulfill the boundary conditions of retarded and advanced Green's function, i.e. inand outgoing waves. The calculations of this section justify these assumptions mathematically.

19.11.2 D. 2 Derivation of Eq. (3.27)

We shall now show how to derive Eq. (3.27). In Eq. (3.26) we have

$$G_{j,j'}^r = \frac{1}{2\pi i} \oint dz U_{<} z^{j-j'} (h(z))^{-1} \quad \text{for } j' \leq j \quad (\text{D.11})$$

where we defined $h(z) = \tilde{H}(E + i\eta, z) U_{<}$.

In order to solve the contour integral of Eq. (D.11) we need to consider each component of the inverse matrix $h(z)^{-1}$ separately. The components of the inverse matrix are given by

$$(h(z)^{-1})_{ij} = \frac{1}{\det(h(z))} \text{adj}(h(z))_{ij} \quad (\text{D.12})$$

where $\text{adj}(h(z))$ is the adjugate of the matrix $h(z)$. $(h(z)^{-1})_{ij}$ is thus a rational function and we need to determine the zeroes of denominator and numerator and their multiplicities in order to solve the contour integral.

First, we consider the denominator. We have $\det(h(z)) = \det(\tilde{H}(E + i\eta, z)) \det(U_{<})$. $\det(U_{<})$ is simply a constant. Therefore we find

$$\det(h(z)) \propto \prod_n (z - \lambda_{n,<}) \prod_n (z - \lambda_{n,>}) \quad (\text{D.13})$$

as $\det(\tilde{H}(E + i\eta, z)) = 0$ for $z = \lambda_{n,<}/>$. Note that for a rank-deficient hopping matrix H_1 , some of the eigenvalues are infinite. The expression above contains finite eigenvalues, and $\det(h(z))$ is a polynomial of degree $N + \text{rank}(H_1)$.

Next, we examine the numerator further. The adjugate of the matrix $h(z)$ is given by $\text{adj}(h(z))_{ij} = (-1)^{i+j} M_{ji}$ where M_{ji} is given by the determinant of the matrix resulting from deleting row j and column i from $h(z)$. Let $S_{N \setminus i}$ denote the set of permutations of the numbers $\{1, \dots, i-1, i+1, \dots, N\}$. Using the Leibniz formula for determinants we can write

$$\text{adj}(h(z))_{ij} = (-1)^{i+j} \sum_{\sigma \in S_{N \setminus i}} \text{sgn}(\sigma) \prod_{\substack{n=1 \\ n \neq j}}^N h(z)_{n\sigma(n)}. \quad (\text{D.14})$$

The matrix $h(z)$ has columns that have definite zeroes at $\lambda_{n,<}$: We can write

$$h(z) = \tilde{H}(E + i\eta, z) U_{<} = \left(\tilde{H}(E + i\eta, z) \mathbf{u}_{1,<} | \dots | \tilde{H}(E + i\eta, z) \mathbf{u}_{N,<} \right) \quad (\text{D.15})$$

Since $\tilde{H}(E + i\eta, z = \lambda_{n,<}) \mathbf{u}_{n,<} = 0$, the entries of the matrix $h_{ij}(z) = \left(\tilde{H}(E + i\eta, z) \mathbf{u}_{j,<} \right)_i$ have a zero of at least multiplicity one at $\lambda_{j,<}$. Thus we find that

$$\text{adj}(h(z))_{ij} = C(z) \prod_{\substack{n=1 \\ n \neq i}}^N (z - \lambda_{n,<}) \quad (\text{D.16})$$

where $C(z)$ is some polynomial.

Combining the expressions for the numerator and the denominator we find

$$(h(z)^{-1})_{ij} = \frac{1}{z - \lambda_{i,<}} D(z) \prod_n \frac{1}{z - \lambda_{n,>}} \quad (\text{D.17})$$

where $D(z)$ is some polynomial. Only the eigenvalues inside the unit circle, $\lambda_{n,<}$, contribute to the contour integral (D.11). Thus we can ignore the poles at $\lambda_{n,>}$ and absorb them in the definition of $D(z)$. In order to solve the integral we need to distinguish two cases:

- $D(\lambda_{i,<}) \neq 0$. In this case, $(h(z)^{-1})_{ij}$ has a simple pole at $\lambda_{i,<}$ and we find

$$\begin{aligned} \frac{1}{2\pi i} \oint dz z^{j-j'} (h(z)^{-1})_{ij} &= \frac{1}{2\pi i} \oint dz \frac{z^{j-j'} D(z)}{z - \lambda_{i,<}} \\ &= (\lambda_{i,<})^{j-j'} D(\lambda_{i,<}) \\ &= (\lambda_{i,<})^{j-j'} \frac{1}{2\pi i} \oint dz (h(z)^{-1})_{ij}. \end{aligned} \quad (\text{D.18})$$

Note that this equation also holds for $\lambda_{i,<} = 0$, when $z^{j-j'} (h(z)^{-1})_{ij}$ is an analytic function within the unit circle.

- $D(\lambda_{i,<}) = 0$. In this case, $(h(z)^{-1})_{ij}$ is an analytic function within the unit circle and

$$\frac{1}{2\pi i} \oint dz z^{j-j'} (h(z)^{-1})_{ij} = \frac{1}{2\pi i} \oint dz (h(z)^{-1})_{ij} = 0 \quad (\text{D.19})$$

Hence, we can still write

$$\frac{1}{2\pi i} \oint dz z^{j-j'} (h(z)^{-1})_{ij} = (\lambda_{i,<})^{j-j'} \frac{1}{2\pi i} \oint dz (h(z)^{-1})_{ij}. \quad (\text{D.20})$$

Finally, from Eqs. (D.18) and (D.20) we find

$$G_{j,j'}^r = \frac{1}{2\pi i} \oint dz U_{<} z^{j-j'} (h(z))^{-1} = \frac{1}{2\pi i} \oint dz U_{<} \Lambda_{<}^{j-j'} (h(z))^{-1} \quad \text{for } j' \leq j, \quad (\text{D.21})$$

where $\Lambda_{<}$ is a diagonal matrix with entries $(\Lambda_{<})_{nn} = \lambda_{n,<}$. This concludes the proof of Eq. (3.27).

It should be noted that during the derivation it was assumed that $j' \leq j$. If $j' > j$, $z^{j-j'} = 1/z^{j'-j}$ introduces additional poles in the contour integral. However, by substituting $y = 1/z$ we can perform an analogous calculation, but in this case only the eigenvalues outside the unit circle contribute.

19.11.3 D. 3 Summary of the numerical algorithms

19.11.4 D.3.1 Eigendecomposition based algorithms

The results from Section 3.2 allow for the computation of the surface Green's function for arbitrary tight-binding models. Here, we summarize the numerical algorithms, distinguishing the two cases of invertible and non-invertible hopping matrix H_1 .

Algorithm D.1. Invertible hopping matrix H_1

A From H_0 , H_1 , and H_{-1} build up the matrix

$$A = \begin{pmatrix} 0 & 1 \\ -(H_1)^{-1} H_{-1} & (H_1)^{-1} (E - H_0) \end{pmatrix}. \quad (\text{D.22})$$

B Solve the eigenproblem

$$A \begin{pmatrix} \mathbf{u}_n \\ \lambda_n \mathbf{u}_n \end{pmatrix} = \lambda_n \begin{pmatrix} \mathbf{u}_n \\ \lambda_n \mathbf{u}_n \end{pmatrix}, \quad (\text{D.23})$$

and extract the $2N$ eigenvectors $\mathbf{u}_n$ and eigenvalues λ_n of the quadratic eigenproblem.

C Identify degenerate eigenvalues of the form $\lambda = e^{ik}$ and rotate the corresponding eigenvectors by diagonalizing the velocity operator.

D Build up the matrices $U_<$ and $\Lambda_<$ from the right-going eigenstates.

E Calculate the surface Green's function as

$$g_R^r = U_< \Lambda_< U_<^{-1} (H_{-1})^{-1}. \quad (\text{D.24})$$

Implementation: The eigenvalue problem can be solved numerically by the LAPACK routine ZGEEV. LAPACK also provides standard routines for matrix multiplication and inversion [107].

Algorithm D.2. Singular hopping matrix H_1

A From H_0 , H_1 , and H_{-1} build up the matrices

$$A = \begin{pmatrix} 0 & 1 \\ -H_{-1} & E - H_0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & H_1 \end{pmatrix} \quad (\text{D.25})$$

B Solve the generalized eigenproblem

$$A \begin{pmatrix} \mathbf{u}_n \\ \lambda_n \mathbf{u}_n \end{pmatrix} = \lambda_n B \begin{pmatrix} \mathbf{u}_n \\ \lambda_n \mathbf{u}_n \end{pmatrix}, \quad (\text{D.26})$$

and extract the $2N$ eigenvectors $\mathbf{u}_n$ and eigenvalues λ_n of the quadratic eigenproblem.

C Identify degenerate eigenvalues of the form $\lambda = e^{ik}$ and rotate the corresponding eigenvectors by diagonalizing the velocity operator.

D Build up the matrices $U_<$ and $\Lambda_<$ from the right-going eigenstates.

E Calculate the surface Green's function as

$$g_R^r = (E - H_0 - H_1 U_< \Lambda_< U_<^{-1})^{-1} \quad (\text{D.27})$$

Implementation: The generalized eigenproblem can be solved numerically by the LAPACK routine ZGGEV [107]. Note that the numerical algorithms in LAPACK solve the generalized eigenvalue problem in the form

$$\beta A \mathbf{u} = \alpha B \mathbf{u} \quad (\text{D.28})$$

such that $\lambda = \alpha/\beta$. In this way, even $\lambda = \infty$ is well-defined as $\beta = 0$.

19.11.5 D.3.2 Schur decomposition based algorithms

The results from Section 3.3 allow for a more stable calculation of the lead Green's function in terms of the (generalized) Schur decomposition. We now summarize the computational steps, distinguishing again the cases of invertible and singular hopping matrices H_1 .

The algorithms are valid as long as $\lambda_{n,<}$ and $\lambda_{n,>}$ do not contain the same eigenvalues. Otherwise, i.e. in situations as depicted in Fig. 3.3(c), the energy E must be shifted slightly before starting the calculation.

Algorithm D.3. Invertible hopping matrix H_1

A From H_0 , H_1 , and H_{-1} build up the matrix

$$A = \begin{pmatrix} 0 & 1 \\ -(H_1)^{-1} H_{-1} & (H_1)^{-1} (E - H_0) \end{pmatrix}. \quad (\text{D.29})$$

B Calculate the Schur decomposition $Q^\dagger A Q = T$ and extract the eigenvalues λ_n .

C Calculate the eigenvectors ϕ_n of A for all propagating modes, i.e. $|\lambda_n| = 1$

D For those propagating modes, identify degenerate eigenvalues of the form $\lambda_n = \lambda_{n'}$ and rotate the corresponding eigenvectors by diagonalizing the velocity operator.

E Classify the eigenvalues into left- and right-going using the calculated eigenvectors ϕ_n for propagating modes.

F Reorder the Schur decomposition such that the eigenvalues $\lambda_{n,<}$ appear on the leading diagonal positions of the triangular matrix T , and update Q accordingly.

G Calculate the surface Green's function as

$$g_R^r = Q_{21} Q_{11}^{-1} (H_{-1})^{-1} \quad (\text{D.30})$$

Implementation: This algorithm can be fully implemented with routines from LAPACK [107]. The Schur decomposition is calculated using ZGHRD, ZUNGHR and ZHSEQR. The eigenvectors for propagating modes are computed using inverse iteration [112] with ZHSEIN. Finally, the routine ZTRSEN is applied to reorder the Schur decomposition.

Algorithm D.4. Singular hopping matrix H_1

A From H_0 , H_1 , and H_{-1} build up the matrices

$$A = \begin{pmatrix} 0 & 1 \\ -H_{-1} & E - H_0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & H_1 \end{pmatrix} \quad (\text{D.31})$$

B Calculate the generalized Schur decomposition $Q^\dagger A Z = T$ and $Q^\dagger B Z = S$, and extract the eigenvalues $\lambda_n = t_{nn}/s_{nn}$. Note that only the transformation Z needs to be calculated explicitly.

C Calculate the eigenvectors ϕ_n for all propagating modes, i.e. $|\lambda_n| = 1$

D For those propagating modes, identify degenerate eigenvalues of the form $\lambda_n = \lambda_{n'}$ and rotate the corresponding eigenvectors by diagonalizing the velocity operator.

E Classify the eigenvalues into left- and right-going using the velocities calculated from the eigenvectors ϕ_n for propagating modes.

F Reorder the generalized Schur decomposition such that the eigenvalues $\lambda_{n,<}$ appear on the leading diagonal positions of the triangular matrices T and S , and update Z accordingly.

G Calculate the surface Green's function as

$$g_R^r = (E - H_0 - H_1 Z_{21} Z_{11}^{-1})^{-1} \quad (\text{D.32})$$

Implementation: Again, LAPACK provides all routines necessary for the computation. The generalized Schur decomposition is computed using ZGGES and from that the eigenvectors for propagating modes using ZTGEVC. The routine ZTGSEN reorders the generalized Schur decomposition.

19.12 Appendix E The Fiduccia-Mattheyses algorithm

19.12.1 E. 1 Graphs and hypergraphs

The Fiduccia-Mattheyses algorithm was originally developed for hypergraph partitioning [132]. A hypergraph $\mathcal{H}$ is an ordered pair $\mathcal{H} = (\mathcal{V}, \mathcal{N})$, where $\mathcal{V}$ is a set of vertices, and $\mathcal{N}$ a set of nets (also called hyperedges) between them. A net $n_i \in \mathcal{N}$ is a set of vertices, i.e. $n_i \subset \mathcal{V}$. An undirected graph is a special realization of a hypergraph, where every net contains exactly two vertices. Thus, any algorithm for hypergraph partitioning can also be applied to an undirected graph.

During the FM bisection, we have to consider the graph structure arising from the Hamiltonian matrix in order to minimize the number of cut edges (min-cut), whereas for minimizing the number of surface vertices, i.e. the number of cut nets (min-netcut), the hypergraph structure arising from all nets $\text{net}(v)$ as defined in Eq. (4.11), $\mathcal{N} = \{\text{net}(v) \mid v \in \mathcal{V}\}$, is essential. For min-net-cut-min-cut optimization, we have to consider both structures simultaneously. A schematic representation of a graph and the corresponding hypergraph structure is shown in Fig. E.1.

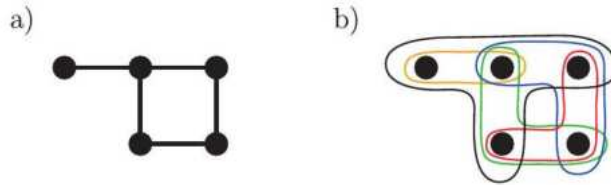


Figure E. 1 - Schematic representation of (a) a simple graph and (b) the corresponding hypergraph structure imposed through all nets, $\mathcal{N} = \{\text{net}(v) \mid v \in \mathcal{V}\}$.

19.12.2 E. 2 Fiduccia-Mattheyses bisection

The FM algorithm is based on the concept of gain. The gain of a vertex in an existing bisection is defined as the change in weight, i.e. the number of cut edges or nets, that occurs when this vertex is moved to the other part. This gain can also be negative, if such a move increases the number of cut edges or nets. The basic idea of the FM algorithm is to swap vertices with the highest gain between parts, while obeying some balance criterion. The fact that the highest gain can be negative, helps the FM algorithm to escape local minima. After moving, the respective vertex is locked in order to avoid an infinite loop, where a single vertex might be swapped back and forth repeatedly. The FM pass ends, when all (free) vertices have been moved, and the best bisection encountered during the pass is returned as result. Further passes can then successively improve on this bisection.

19.13 Appendix F The method of finite differences

19.13.1 F. 1 Basic ideas

The fundamental concept behind the method of finite differences go back to Newton and his contemporaries: Replace the derivative of a function by a difference quotient. With this replacement, a differential equation is reduced to a set of algebraic equations, that can easily be solved on a computer. In fact, the method of finite differences was already used for a numerical solution of the Schrödinger equation before the advent of electronic computers [77].

The approximations needed in the method of finite differences are most conveniently remembered from the Taylor expansion of a function f :

$$f(x + h) = f(x) + f'(x)h + \mathcal{O}(h^2) \quad (\text{F.1})$$

and

$$f(x - h) = f(x) - f'(x)h + \mathcal{O}(h^2). \quad (\text{F.2})$$

Subtracting Eq. (F.2) from Eq. (F.1), we obtain the symmetric approximation for the derivative of a function,

$$f'(x) = \frac{f(x + h) - f(x - h)}{2h} + \mathcal{O}(h^2), \quad (\text{F.3})$$

whereas adding Eqs. (F.1) and (F.2), we obtain an approximation for the function itself:

$$f(x) = \frac{1}{2}(f(x + h) + f(x - h)) + \mathcal{O}(h^2). \quad (\text{F.4})$$

At first glance, Eq. (F.4) may not seem very useful. However, it can be used conveniently to bring the finite difference expressions into a symmetric form, yielding a

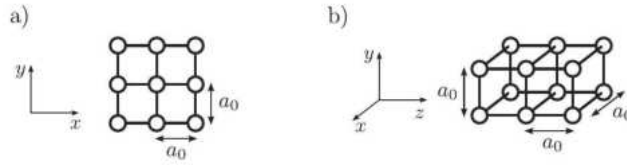


Figure F. 1 - Grids used for the method of finite differences: (a) square grid with lattice constant a_0 in two dimensions. (b) cubic grid with lattice constant a_0 in three dimensions.

Hermitian Hamiltonian. Note that Eqs. (F.3) and (F.4) are valid up to second order in h . It is possible to use higher-order approximation formulas (see, for example, Ref. [232]), but we will restrict ourselves to the lowest order in this work.

In the method of finite differences, functions are not evaluated in the continuum, but only on a discrete set of grid points. The simplest grids in two and three dimensions are the uniform square and cubic grids with lattice constant a_0 , shown in Fig. F.1. Using Eqs. (F.3) and (F.4) with $h = a_0$ and $h = a_0/2$ when appropriate, it is possible to reduce a differential equation to a closed set of linear equations involving only function values on the grid points.

19.13.2 F. 2 Example

In order to put the general ideas of the previous section into practice and to illustrate the typical steps in deriving a finite differences approximation to a Hamiltonian, we consider one example in great detail. Another detailed example can be found in the appendix of Ref. [155].

Consider the Hamiltonian H acting on the wave function $\varphi(x)$

$$\begin{aligned} H\varphi(x) &= \frac{1}{2} ((-i\hbar\partial_x + eA(x))\alpha(x)\varphi(x) + \alpha(x)(-i\hbar\partial_x + eA(x))\varphi(x)) \\ &= \frac{1}{2} (-i\hbar(\partial_x(\alpha(x)\varphi(x)) + \alpha(x)(\partial_x\varphi(x))) + 2e(A(x)\alpha(x)\varphi(x))) \end{aligned} \quad (\text{F.5})$$

Such terms arises from the Bychkov-Rashba Hamiltonian with a position-dependent Bychkov-Rashba parameter $\alpha(x)$ in a magnetic field with vector potential $A(x)$ (see Eq. (F.16)). Using (F.3), the first bracket in the last line of Eq. (F.5) can be rewritten as

$$\begin{aligned} \partial_x(\alpha(x)\varphi(x)) + \alpha(x)(\partial_x\varphi(x)) &\approx \\ \frac{1}{2a_0} ((\alpha(x) + \alpha(x + a_0))\varphi(x + a_0) - (\alpha(x) + \alpha(x - a_0))\varphi(x - a_0)) \end{aligned} \quad (\text{F.6})$$

The second bracket in the last line of Eq. (F.5) is conveniently recast into a more symmetric form as

$$\begin{aligned}
 A(x)\alpha(x)\varphi(x) &= \frac{1}{4}((A(x)\alpha(x))\varphi(x) + A(x)(\alpha(x)\varphi(x)) \\
 &\quad + \alpha(x)(A(x)\varphi(x)) + (A(x)\alpha(x)\varphi(x))) \\
 &\approx \frac{1}{8}((\alpha(x) + \alpha(x+a_0))(A(x) + A(x+a_0))\varphi(x+a_0) \\
 &\quad + (\alpha(x) + \alpha(x-a_0))(A(x) + A(x-a_0))\varphi(x-a_0)) \quad (F.7)
 \end{aligned}$$

where we used Eq. (F.4) on every one of the expressions in the inner brackets of the first line. Combining all terms, we finally arrive at

$$\begin{aligned}
 H\varphi(x) &\approx -\frac{i\hbar}{2a_0}\frac{1}{2}(\alpha(x) + \alpha(x+a_0))\varphi(x+a_0)\left(1 + \frac{iea_0}{\hbar}\frac{1}{2}(A(x) + A(x+a_0))\right) \\
 &\quad + \frac{i\hbar}{2a_0}\frac{1}{2}(\alpha(x) + \alpha(x-a_0))\varphi(x-a_0)\left(1 - \frac{iea_0}{\hbar}\frac{1}{2}(A(x) + A(x-a_0))\right). \quad (F.8)
 \end{aligned}$$

Introducing the notation $|x\rangle$ for a state at grid point x , we can write $H\varphi(x) = \langle x|H|\varphi\rangle$ and then read off the operator H from Eq. (F.8):

$$H \approx \sum_x -\frac{i\hbar}{2a_0}\frac{1}{2}(\alpha(x) + \alpha(x+a_0))e^{\frac{iea_0}{\hbar}\frac{1}{2}(A(x)+A(x+a_0))}|x\rangle\langle x+a_0| + \text{h.c.} \quad (F.9)$$

where h.c. denotes the Hermitian conjugate of the previous expression. In order to arrive at this expression, we have made use of the fact that all approximations made so far are valid up to second order in a_0 , and hence

$$1 + \frac{iea_0}{\hbar}\frac{1}{2}(A(x) + A(x+a_0)) \approx e^{\frac{iea_0}{\hbar}\frac{1}{2}(A(x)+A(x+a_0))} \quad (F.10)$$

The operator form of H in Eq. (F.9) has the form of a tight-binding operator. In this particular example, the Hamiltonian consists of a "hopping term" from point x to $x+a_0$, i.e. hopping between neighboring grid points. The magnetic field enters in the Hamiltonian in the form of a Peierls phase [233]

$$e^{\frac{iea_0}{\hbar}\frac{1}{2}(A(x)+A(x+a_0))} \approx e^{\frac{ie}{\hbar}\int_x^{x+a_0} d\mathbf{A}\cdot\mathbf{A}(\mathbf{x})} \quad (F.11)$$

Because of this formal equivalence, the discretization procedure in the method of finite differences is said to yield a tight-binding representation of the Hamiltonian H .

19.13.3 F. 3 Finite differences form of the Hamiltonian

After explicitly deriving the tight-binding representation of a particular example, we will now state the finite differences forms of several terms to the Hamiltonian encountered in this work. The expressions stated here are mostly valid in two ($\mathbf{x} = (x, y)$) and in three dimensions ($\mathbf{x} = x, y, z$). For brevity, we will not explicitly write out coordinates that remain unchanged, i.e. do not contribute a hopping term. In addition, we will not explicitly include a magnetic field. In order to include the effect of a magnetic field, every hopping term must be multiplied by the Peierls phase (as in the previous example, see Eq. (F.11)). As a general guideline, if a Hamiltonian contains parameters that exhibit a discontinuity, the grid should be chosen such that the discontinuity lies halfway between two grid points [234].

Kinetic energy: The kinetic energy Hamiltonian with a position-dependent mass $m(\mathbf{x})$ [235] reads

$$H_{\text{kin}} = \frac{1}{2} \mathbf{p} \frac{1}{m(\mathbf{x})} \mathbf{p} \quad (\text{F.12})$$

where $\mathbf{p} = -i\hbar \nabla$. The tight-binding representation is given as

$$\begin{aligned} H_{\text{kin}} = & \sum_{\mathbf{x}} \left(\frac{1}{m(x - \frac{a_0}{2})} + \frac{1}{m(x + \frac{a_0}{2})} \right) |x\rangle \langle x| \\ & - \left(\frac{1}{m(x + \frac{a_0}{2})} |x\rangle \langle x + a| + \text{h.c.} \right) \\ & + \text{analogous terms for the } y \text{ and } z\text{-direction.} \end{aligned} \quad (\text{F.13})$$

In the numerical calculation it is sometimes advantageous to substitute $m(x + \frac{a_0}{2}) \rightarrow \frac{1}{2}(m(x) + m(x + a_0))$ [234].

Potential energy: A potential enters the Hamiltonian as

$$H_{\text{pot}} = V(\mathbf{x}) \quad (\text{F.14})$$

and in tight-binding representation simply reads

$$H_{\text{pot}} = \sum_{\mathbf{x}} V(\mathbf{x}) |\mathbf{x}\rangle \langle \mathbf{x}|. \quad (\text{F.15})$$

Bychkov-Rashba spin-orbit coupling: The Bychkov-Rashba spin-orbit coupling Hamiltonian [21, 22] with a position-dependent Bychkov-Rashba parameter $\alpha(\mathbf{x})$ reads

$$H_{\text{BR}} = \frac{1}{2\hbar} \alpha(\mathbf{x}) (\sigma^x p_y - \sigma^y p_x) + \frac{1}{2\hbar} (\sigma^x p_y - \sigma^y p_x) \alpha(\mathbf{x}), \quad (\text{F.16})$$

where $\boldsymbol{\sigma}$ is the vector of Pauli matrices. Note that the Hamiltonian is symmetrized such that it is Hermitian (see Ref. [153, Chap. 3]). The tight-binding representation is then given as

$$\begin{aligned} H_{\text{BR}} = & \sum_{\mathbf{x}} \frac{-i}{2a_0} \sigma^x \frac{1}{2} (\alpha(y) + \alpha(y + a_0)) |y\rangle \langle y + a| + \text{h.c.} \\ & + \frac{i}{2a_0} \sigma^y \frac{1}{2} (\alpha(x) + \alpha(x + a_0)) |x\rangle \langle x + a| + \text{h.c.} \end{aligned} \quad (\text{F.17})$$

Dresselhaus spin-orbit coupling: The Dresselhaus spin-orbit coupling Hamiltonian [236] with a position-dependent Dresselhaus parameter $\gamma(z)$ reads

$$H_{\text{D}} = \frac{1}{\hbar} (\sigma^x p_x - \sigma^y p_y) \frac{\partial}{\partial z} \left(\gamma(z) \frac{\partial}{\partial z} \right) \quad (\text{F.18})$$

and in tight-binding representation

$$\begin{aligned}
H_D = & \frac{-i}{2a_0^3} \sigma^x \sum_{\mathbf{x}} \gamma \left(z + \frac{a_0}{2} \right) |x, y, z\rangle \langle x + a_0, y, z + a_0| \\
& + \gamma \left(z - \frac{a_0}{2} \right) |x, y, z\rangle \langle x + a_0, y, z - a_0| \\
& - \left(\gamma \left(z - \frac{a_0}{2} \right) + \gamma \left(z + \frac{a_0}{2} \right) \right) |x, y, z\rangle \langle x + a_0, y, z| \\
& \frac{i}{2a_0^3} \sigma^y \sum_{\mathbf{x}} \gamma \left(z + \frac{a_0}{2} \right) |x, y, z\rangle \langle x, y + a_0, z + a_0| \\
& + \gamma \left(z - \frac{a_0}{2} \right) |x, y, z\rangle \langle x, y + a_0, z - a_0| \\
& - \left(\gamma \left(z - \frac{a_0}{2} \right) + \gamma \left(z + \frac{a_0}{2} \right) \right) |x, y, z\rangle \langle x, y + a_0, z| \quad (F.19)
\end{aligned}$$

- Hermitian conjugate of the whole expression.

Note that in contrast to the previous expressions the tight-binding representation of the Dresselhaus parameter includes next-nearest neighbor hopping terms. Again, it may be useful to substitute $\gamma \left(z + \frac{a_0}{2} \right) \rightarrow \frac{1}{2} (\gamma(z) + \gamma(z + a_0))$.

19.13.4 F. 4 Bloch's theorem and periodic boundary conditions

If a tight-binding Hamiltonian has a translational symmetry in some direction, Bloch's theorem can be used to reduce the size of the problem.

As example, consider a square grid, which is translationally invariant under a shift of $N_y a_0$, i.e. is composed of periodically repeated supercells with N_y lattice points in the y -direction (see Fig. F.2(a)). From Bloch's theorem, we can write

$$\varphi(x, y + N_y a_0) = e^{ik_y N_y a_0} \varphi(x, y) \quad (F.20)$$

where $k_y \in \left[-\frac{\pi}{N_y a_0}, \frac{\pi}{N_y a_0} \right]$. In particular, this also holds for grid points at the boundary of the supercell. Let y_1 and y_2 denote the lower and upper row of grid points in the supercell. Then

$$\varphi(x, y_2 + a_0) = e^{ik_y N_y a_0} \varphi(x, y_1) \quad (F.21)$$

Within the tight-binding Hamiltonian we can then replace hopping terms at the boundary as

$$|y_2\rangle \langle y_2 + a_0| \rightarrow e^{ik_y N_y a_0} |y_2\rangle \langle y_1| \quad (F.22)$$

a)

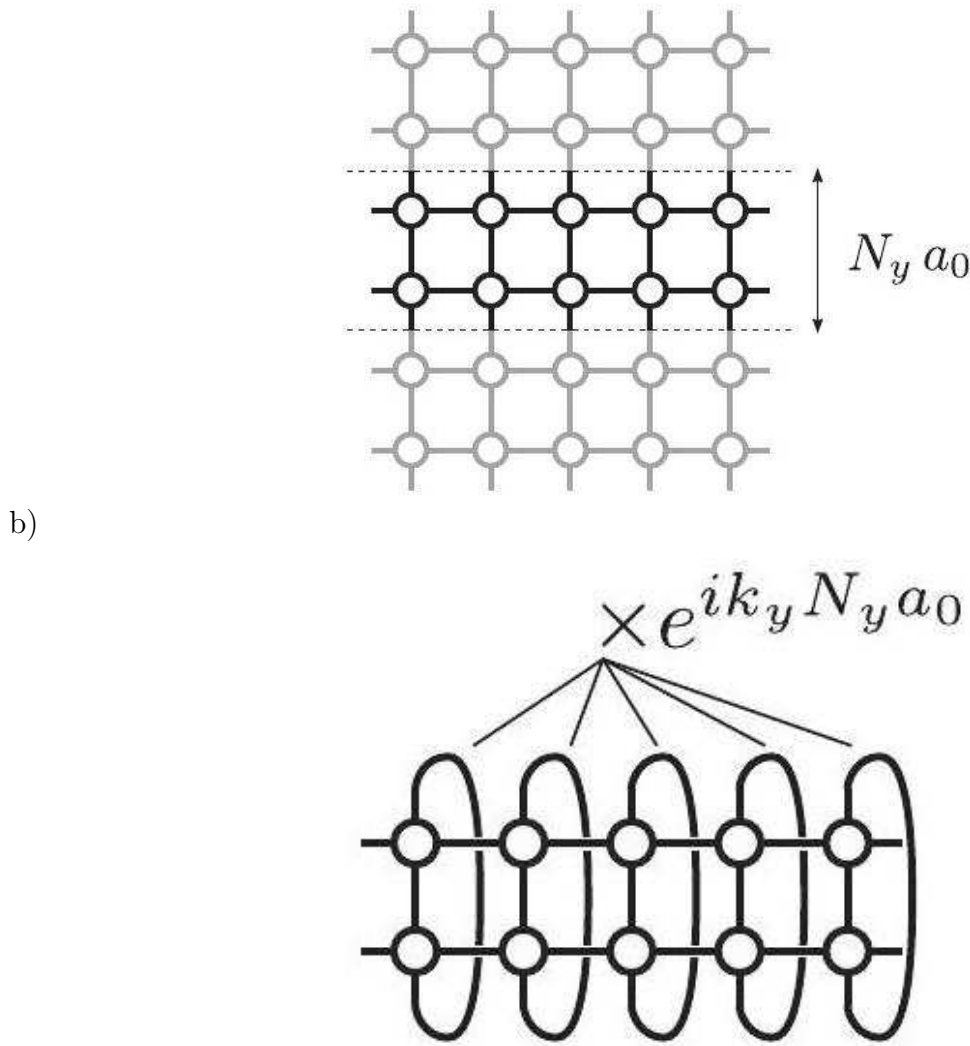


Figure F. 2 - (a) A square grid consisting of a periodic repetition of supercells in the y direction. (b) Effective structure after the application of Bloch's theorem.

and

$$|y_1\rangle \langle y_1 - a_0| \rightarrow e^{-ik_y N_y a_0} |y_1\rangle \langle y_2|. \quad (\text{F.23})$$

Hence, the system size has an effective width of N_y grid points in the y -direction, however with additional hopping terms from the top to the bottom of the supercell and vice versa, with a phase given by the Bloch wave vector k_y .

Note that if we set $k_y = 0$, we obtain the conventional "periodic boundary conditions". Neglecting all other Bloch vectors k_y is a good approximation, if $N_y a_0$ is large, i.e. k_y small. If this is not the case, the desired calculation must be performed for a set of k_y -values.

19.14 Appendix G Tight-binding model for graphene

19.14.1 G. 1 Lattice structure

Graphene is a two-dimensional realization of carbon, where the carbon atoms are arranged in a honeycomb network. The honeycomb network itself is not a Bravais lattice, but can be described as a triangular (or hexagonal) lattice with a basis of two atoms [93].

The triangular lattice is generated by the primitive basis vectors

$$\mathbf{a}_1 = \begin{pmatrix} a \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{a}_2 = \begin{pmatrix} \frac{1}{2}a \\ \frac{\sqrt{3}}{2}a \end{pmatrix} \quad (\text{G.1})$$

as shown in Fig. G.1(a). The primitive unit cell spanned by these vectors is a rhombus with side length a . In the case of the honeycomb network, there are two basis atoms per unit cell at the positions

$$\mathbf{d}_A = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{d}_B = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{3}}a \end{pmatrix} \quad (\text{G.2})$$

We will denote these basis atoms as atom A and B respectively. It should be emphasized that, although we label these basis atoms differently, within the infinitely extended honeycomb network, they are equivalent. The basis atoms form two triangular lattices (sublattice A and B, respectively), that are shifted with respect to each other to obtain the honeycomb network. In this network, each atom has three nearest neighbors of the opposite kind. The distance between neighboring carbon atoms is given as $a_{CC} = a/\sqrt{3}$.

In the case of a triangular lattice, the reciprocal is again triangular, rotated by 90° with respect to the original lattice, as shown in Fig. G.1(b). The reciprocal lattice is spanned by the primitive basis vectors [93]

$$\mathbf{b}_1 = \begin{pmatrix} \frac{2\pi}{a} \\ \frac{2\pi}{\sqrt{3}a} \end{pmatrix} \quad \text{and} \quad \mathbf{b}_2 = \begin{pmatrix} 0 \\ \frac{4\pi}{\sqrt{3}a} \end{pmatrix}. \quad (\text{G.3})$$

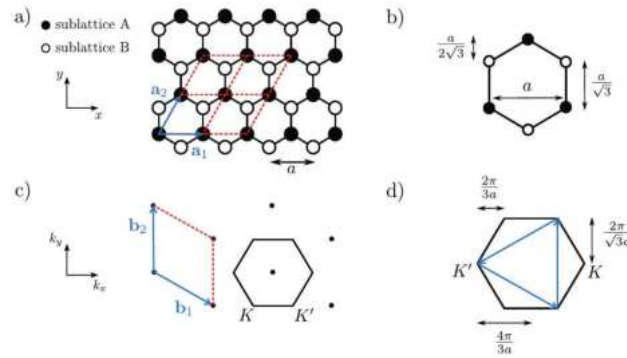


Figure G. 1 - (a) Honeycomb network of carbon atoms. The underlying triangular Bravais lattice is spanned by the primitive basis vectors $\mathbf{a}_1$ and $\mathbf{a}_2$, shown in blue. The primitive unit cell is a rhombus, as depicted by the red dashed lines. The two basis atoms are shown as black (sublattice A) and white (sublattice B) dots, respectively. (b) Geometrical details of a hexagon from the honeycomb network, formed by carbon atoms (c) The reciprocal lattice is spanned by the reciprocal basis vectors $\mathbf{b}_1$ and $\mathbf{b}_2$, shown in blue. The first Brillouin zone has a hexagonal shape. (d) Geometrical details of the first Brillouin zone, showing the reciprocal lattice vectors connecting the three equivalent K' -corners.

The first Brillouin zone is constructed as the Wigner-Seitz cell of the reciprocal lattice [93] and has the shape of a hexagon. However, a wave vector in a Bravais lattice is only defined up to some reciprocal lattice vector. Thus, only two of the six corners in the Brillouin zone are non-equivalent, as always three corners can be connected with a reciprocal lattice vector. These two inequivalent corners are called K -points and are labeled K and K' respectively. For simplicity, we choose the wave vectors

$$\mathbf{K} = \begin{pmatrix} \frac{4\pi}{3a} \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{K}' = \begin{pmatrix} -\frac{4\pi}{3a} \\ 0 \end{pmatrix} \quad (\text{G.4})$$

to represent the two K -points.

It should be emphasized that the hexagonal shape of the Brillouin zone, and thus the existence of two K -points, is solely due to the underlying triangular Bravais lattice structure of the

honeycomb network. The reciprocal lattice only depends on the Bravais lattice structure, not on the basis atoms per unit cell. In particular, the hexagonal shape of the Brillouin zone is not related to the hexagons formed by carbon atoms in the honeycomb network.

19.14.2 G. 2 Electronic structure

19.14.3 G.2.1 Tight-binding model

We consider the dynamics of electrons in the graphene lattice assuming that the effect of interactions can be incorporated by a mean field (non-interacting electrons; for a justification see e.g. [237, Chap. 5]). In this case, the single-particle Hamiltonian of the system is given as

$$H = -\frac{\hbar^2}{2m}\nabla^2 + V_{\text{eff}}(\mathbf{x}) + V_{\text{ext}}(\mathbf{x}) \quad (\text{G.5})$$

where $V_{\text{eff}}(\mathbf{x})$ incorporates both the potential arising from the ions within the graphene lattice and the effect of interaction in an approximate way, and V_{ext} is some external potential. Note that in the case of an infinite graphene sheet, $V_{\text{eff}}(\mathbf{x})$ has the full symmetry of the graphene lattice. We now intend to develop a tight-binding formulation for this problem, as was done first by Wallace in 1947 [238].

Carbon is the sixth element in the periodic table, and has 4 valence electrons in $2s^2 2p^2$ configuration. In graphene (and graphite, from which graphene originates), three of the valence electrons hybridize in a $2sp^2$ configuration. These electrons form bonds with the neighboring carbon atoms, leading to the honeycomb arrangement of carbon atoms. The fourth electron resides in a $2p_z$ -orbital perpendicular to the graphene plane. The electrons that form the bonds to the neighboring carbon atoms will not contribute to transport due to their bound nature, and we can therefore treat graphene as having one electron per carbon atom in the $2p_z$ state.

Following the arguments above, we describe graphene in tight-binding approximation with a single orbital per basis atom. The orbital at an atom situated at $\mathbf{R}_i + \mathbf{d}_\alpha$ is given as

$$\chi_{i,\alpha,s}(\mathbf{x}) = \chi_{2p_z,s}(\mathbf{x} - \mathbf{R}_i - \mathbf{d}_\alpha) \quad (\text{G.6})$$

where i denotes the (Bravais) lattice vector $\mathbf{R}_i = n\mathbf{a}_1 + m\mathbf{a}_2 (n, m \in \mathbb{Z})$, α the sublattice ($\alpha = A, B$) and s the spin. $\chi_{2p_z,s}(\mathbf{x} - \mathbf{R}_i - \mathbf{d}_\alpha)$ is the wave function of a $2p_z$ -orbital centered at atom α in the unit cell at the lattice point $\mathbf{R}_i$, with spin s . The quantum mechanical state corresponding to the wave function $\chi_{i,\alpha,s}(\mathbf{x})$ will be denoted as $|i, \alpha, s\rangle$. As discussed in Section 2.5.1, we assume that the states $|i, \alpha, s\rangle$ form an orthonormal set, thus neglecting overlap between wave functions located at different atoms. Within this basis, the tight-binding Hamiltonian reads

$$H = \sum_{i,\alpha,j,\beta,s,s'} \langle i, \alpha, s | H | j, \beta, s' \rangle |i, \alpha, s\rangle \langle j, \beta, s'| \quad (\text{G.7})$$

In the remainder of this chapter we will only consider systems where spin is a good quantum number, and suppress the spin index s in the notation. In this case, spin only enters in a twofold occupation of energy levels. When we consider spin-dependent transport properties of graphene nanoribbons in Chapter 7, we will explicitly discuss how spin-dependent terms enter the Hamiltonian.

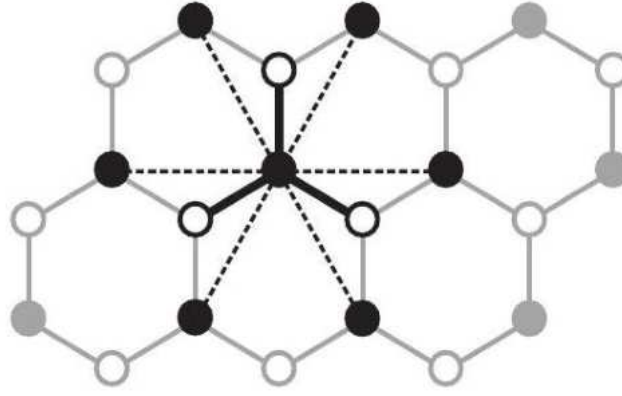


Figure G. 2 - The nearest neighbors of an A atom are three B atoms (connections shown as solid black lines), the next-nearest neighbors are six A atoms (connections shown as dashed black lines).

In the spin-less case, the Hamiltonian reads

$$H = \sum_{i,\alpha,j,\beta} t_{i\alpha j\beta} |i, \alpha\rangle \langle j, \beta|, \quad (\text{G.8})$$

where $t_{i\alpha j\beta} = \langle i\alpha | H | j\beta \rangle$. The solution has the form

$$|\varphi\rangle = \sum_{i,\alpha} \varphi_\alpha(i) |i, \alpha\rangle, \quad (\text{G.9})$$

where $\varphi_\alpha(i)$ is the probability amplitude of occupying the orbital state $|i, \alpha\rangle$. In realspace, the wave function reads

$$\varphi(\mathbf{x}) = \sum_{i,\alpha} \varphi_\alpha(i) \chi_{2p_z}(\mathbf{x} - \mathbf{R}_i - \mathbf{d}_\alpha). \quad (\text{G.10})$$

Thus, due to the localized nature of the atomic orbitals, $|\varphi_\alpha(i)|^2$ can be roughly interpreted as the probability of being close to the atomic site $\mathbf{R}_i + \mathbf{d}_\alpha$.

The fact that the atomic orbitals are localized is also the reason why only a few of the matrix elements $t_{i\alpha j\beta}$ are nonzero. Here, we will only consider matrix elements involving orbitals at the same site, and orbitals that are nearest or next-nearest neighbors.

The matrix elements involving only orbitals at the same atomic site can be written as $t_{i\alpha i\alpha} = \epsilon_0 + \epsilon_{i\alpha}$, where ϵ_0 arises from the lattice-periodic part of the Hamiltonian, $-\frac{\hbar^2}{2m}\nabla^2 + V_{\text{eff}}$, and $\epsilon_{i\alpha}$ from V_{ext} . Due to the lattice periodicity, ϵ_0 is the same for every atomic site and can thus be set to zero. For the external potential, it is useful to recast $\epsilon_{i\alpha}$ into a different form:

$$\epsilon_{i\alpha} = \begin{cases} V_i + M_i & \text{for } \alpha = \text{A}, \\ V_i - M_i & \text{for } \alpha = \text{B} \end{cases} \quad (\text{G.11})$$

Here, V_i has the meaning of a potential that has the same value for both atoms in the unit cell, and M_i is a staggered potential, i.e. has opposite sign for sublattice A and B. Such a staggered potential can thus arise from a potential varying rapidly on the inter-atomic scale. In addition, it is expected to arise naturally by placing graphene on a suitable substrate (breaking the symmetry between A and B atoms) [31], and due to antiferromagnetic ordering in zigzag graphene nanoribbons, as discussed in Chapter 7.

We now turn to the matrix elements involving orbitals at different sites, i.e. hopping matrix elements. As seen from Fig. G.2, every atom of a given type is surrounded by three nearest atoms of the opposite kind. The distance vectors from an A atom to the nearest B atoms are given by

$$\boldsymbol{\delta}_1 = \begin{pmatrix} 0 \\ \frac{a}{\sqrt{3}} \end{pmatrix}, \quad \boldsymbol{\delta}_2 = \begin{pmatrix} \frac{a}{2} \\ -\frac{a}{2\sqrt{3}} \end{pmatrix}, \quad \text{and} \quad \boldsymbol{\delta}_3 = \begin{pmatrix} -\frac{a}{2} \\ -\frac{a}{2\sqrt{3}} \end{pmatrix} \quad (\text{G.12})$$

the distance vectors from an B atom to the closest A atoms as $-\boldsymbol{\delta}_1$, $-\boldsymbol{\delta}_2$, and $-\boldsymbol{\delta}_3$. The next-nearest neighbors of an atom of a certain kind are given by six atoms of the same kind, with distance vectors

$$\boldsymbol{\delta}'_1 = \mathbf{a}_1, \boldsymbol{\delta}'_2 = \mathbf{a}_2, \boldsymbol{\delta}'_3 = \mathbf{a}_2 - \mathbf{a}_1, \boldsymbol{\delta}'_4 = -\mathbf{a}_1, \boldsymbol{\delta}'_5 = -\mathbf{a}_2, \text{ and } \boldsymbol{\delta}'_6 = \mathbf{a}_1 - \mathbf{a}_2 \quad (\text{G.13})$$

Since the $2p_z$ -orbital is symmetric with respect to rotations in the plane of the graphene sheet, we can parameterize all matrix elements involving nearest and next-nearest neighbors with a single parameter t and t' , respectively.

In summary, we consider a tight-binding model for graphene in next-nearest neighbor approximation with matrix elements

$$t_{i\alpha j\beta} = \begin{cases} V_i + M_i & \text{if } i = j, \alpha = \beta = \text{A} \\ V_i - M_i & \text{if } i = j, \alpha = \beta = \text{B} \\ -t & \text{if } i, \alpha \text{ and } j, \beta \text{ are nearest neighbors,} \\ -t' & \text{if } i, \alpha \text{ and } j, \beta \text{ are next-nearest neighbors.} \end{cases} \quad (\text{G.14})$$

In the spirit of the Slater-Koster approach to tight-binding [239], we do not attempt to calculate the hopping matrix elements from first principles, but as parameters that should be adjusted to fit *ab initio* calculations. The numerical values of the hopping matrix elements employed in this work are $t = 2.7\text{eV}$ and $t' = 0.1t$ [183, 184]. The effects of a magnetic field can be included through the Peierls phase [233], by replacing

$$t_{i\alpha j\beta} \rightarrow t_{i\alpha j\beta} \times \exp\left(\frac{ie}{\hbar} \int_{\mathbf{R}_j + \mathbf{d}_\beta}^{\mathbf{R}_i + \mathbf{d}_\alpha} d\mathbf{s} \cdot \mathbf{A}(\mathbf{x})\right) \quad (\text{G.15})$$

where $\mathbf{A}(\mathbf{x})$ is the vector potential.

Before we proceed to calculate the band structure of this tight-binding model, let us revisit a few of the approximations made in its derivation: First, we always assumed that $V_{\text{eff}}(\mathbf{x})$ has the full symmetry of the graphene lattice. This is not the case for a finite system, where especially atoms close to the boundary would feel a very different

effective lattice potential. In principle, this could result in quite different tight-binding parameters close to the boundary. However, we will neglect these effects when considering finite systems. This common approximation can be justified a posteriori by the observation that the results agree favorably with other approaches, such as DFT, that include the above mentioned effects of the boundaries (For an example, see the discussion of zigzag graphene nanoribbons in Chapter 7).

Second, the model as discussed above considers only up to next-nearest neighbor hopping. Reich et al. additionally include third-nearest neighbor hopping [190] on the grounds that the third-nearest neighbor distance is only slightly larger than the next-nearest neighbor distance. However, third-nearest neighbor hopping introduces only some corrections to the nearest-neighbor hopping, whereas next-nearest neighbor hopping introduces terms in the Hamiltonian that have different symmetry, as discussed below. These terms of different symmetry play a crucial role in the discussion of transport in graphene nanoribbons in Chapter 6. Therefore, the inclusion of nextnearest neighbors is essential, whereas third-nearest neighbor hopping only gives small corrections.

Finally, the tight-binding model is based on a single $2p_z$ -orbital per atomic site. In principle, it is possible to include all valence electrons on equal footing [240, 241]. However, in the energy

range relevant for transport, the band structures obtained from these extended tight-binding models are practically indistinguishable from the results of the much simpler single-orbital tight-binding approach employed here.

19.14.4 G.2.2 Band structure in tight-binding approximation

After introducing the tight-binding model for graphene, we now calculate the band structure for a homogeneous infinite graphene sheet, with $M_i = M$, i.e. a constant staggered potential, and $V_i = V_0$. Then, the system is invariant under lattice translations,¹ and Bloch's theorem holds. For this case, we can apply standard techniques to calculate the band structure [93, Chap. 10].

To this end, we define the tight-binding sums²

$$\Phi_\alpha(\mathbf{k}, \mathbf{x}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}_i} e^{i\mathbf{k}(\mathbf{R}_i + \mathbf{d}_\alpha)} \chi_{i,\alpha}(\mathbf{x}) \quad (\text{G.16})$$

where $\alpha = A, B$ and N is the number of unit cells. The tight-binding sums form an orthonormal set, i.e.

$$\langle \Phi_\alpha(\mathbf{k}) | \Phi_\beta(\mathbf{k}') \rangle = \delta_{\alpha,\beta} \delta_{\mathbf{k},\mathbf{k}'} \quad (\text{G.17})$$

and obey Bloch's theorem. From the tight-binding sums, we make the ansatz

$$\varphi(\mathbf{k}, \mathbf{x}) = \varphi_A(\mathbf{k}) \Phi_A(\mathbf{k}, \mathbf{x}) + \varphi_B(\mathbf{k}) \Phi_B(\mathbf{k}, \mathbf{x}) \quad (\text{G.18})$$

for the eigenfunctions with Bloch vector $\mathbf{k}$. Using the orthonormality of the tightbinding sums, the Schrödinger equation $H\varphi(\mathbf{k}, \mathbf{x}) = E\varphi(\mathbf{k}, \mathbf{x})$ turns into an eigenproblem involving only the coefficients $\varphi_{A,B}(\mathbf{k})$:

$$H(\mathbf{k}) \begin{pmatrix} \varphi_A(\mathbf{k}) \\ \varphi_B(\mathbf{k}) \end{pmatrix} = \begin{pmatrix} H_{AA}(\mathbf{k}) & H_{AB}(\mathbf{k}) \\ H_{BA}(\mathbf{k}) & H_{BB}(\mathbf{k}) \end{pmatrix} \begin{pmatrix} \varphi_A(\mathbf{k}) \\ \varphi_B(\mathbf{k}) \end{pmatrix} = E(\mathbf{k}) \begin{pmatrix} \varphi_A(\mathbf{k}) \\ \varphi_B(\mathbf{k}) \end{pmatrix}. \quad (\text{G.19})$$

Here, $H_{\alpha\beta} = \langle \Phi_\alpha(\mathbf{k}) | H | \Phi_\beta(\mathbf{k}) \rangle$. In order to calculate the band structure $E(\mathbf{k})$, we now evaluate these matrix elements using the tight-binding model in next-nearest neighbor approximation of the preceding section. We obtain

$$\begin{aligned} \langle \Phi_A(\mathbf{k}) | H | \Phi_A(\mathbf{k}) \rangle &= \frac{1}{N} \sum_{\mathbf{R}_i, \mathbf{R}_j} e^{i\mathbf{k}(\mathbf{R}_j - \mathbf{R}_i)} \langle i, A | H | j, A \rangle \\ &= \frac{1}{N} \sum_{\mathbf{R}_i, \mathbf{R}_j} e^{i\mathbf{k}(\mathbf{R}_j - \mathbf{R}_i)} \left((V_0 + M) \delta_{\mathbf{R}_i, \mathbf{R}_j} - t' \sum_{l=1}^6 \delta_{\mathbf{R}_j - \mathbf{R}_i, \delta'_l} \right) \\ &= V_0 + M - t' \sum_{l=1}^6 e^{i\mathbf{k}\delta'_l} = V_0 + M - 2t' \sum_{l=1}^3 \cos(\mathbf{k}\delta'_l), \end{aligned} \quad (\text{G.20})$$

¹²¹ At first glance, the staggered, and thus spatially dependent potential M seems to break translational invariance. However, the staggered potential changes sign only on A and B atoms within a unit cell, and exhibits the triangular lattice periodicity.

² Note that there are different conventions in the literature on how to define a tight-binding sum for a lattice with more than one basis atom. For example, the definition $\Phi_\alpha(\mathbf{k}, \mathbf{x}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}_i} e^{i\mathbf{k}\mathbf{R}_i} \chi_{i,\alpha}(\mathbf{x})$ is also used commonly [93, 237]. However, these different definitions are equivalent, as the corresponding wave functions only differ by a ($\mathbf{k}$ -dependent) phase. Here we follow the conventions set by the original calculation of Wallace [238] and the work of Haldane [242].

where we used $\delta_4 = -\delta_1$, etc. in the last line. In the same fashion we find the matrix element involving only the B sublattice as

$$\langle \Phi_B(\mathbf{k}) | H | \Phi_B(\mathbf{k}) \rangle = V_0 - M - 2t' \sum_{l=1}^3 \cos(\mathbf{k}\delta'_l). \quad (\text{G.21})$$

The off-diagonal matrix elements evaluate to

$$\begin{aligned} \langle \Phi_A(\mathbf{k}) | H | \Phi_B(\mathbf{k}) \rangle &= \frac{1}{N} \sum_{\mathbf{R}_i, \mathbf{R}_j} e^{i\mathbf{k}(\mathbf{R}_j + \mathbf{d}_B - \mathbf{R}_i)} \langle i, A | H | j, B \rangle \\ &= -t \sum_{l=1}^3 e^{i\mathbf{k}\delta_l} = -t \sum_{l=1}^3 \cos(\mathbf{k}\delta_l) + i \sin(\mathbf{k}\delta_l) \end{aligned} \quad (\text{G.22})$$

and $H_{BA}(\mathbf{k}) = (H_{AB}(\mathbf{k}))^*$.

The tight-binding Hamiltonian in $\mathbf{k}$ -space can thus be written in the form [242]

$$H(\mathbf{k}) = \left(V_0 - 2t' \sum_{l=1}^3 \cos(\mathbf{k}\delta'_l) \right) \mathbb{I} - \left(t \sum_{l=1}^3 \cos(\mathbf{k}\delta_l) \right) \sigma^x + \left(t \sum_{l=1}^3 \sin(\mathbf{k}\delta_l) \right) \sigma^y + M \sigma^z \quad (\text{G.23})$$

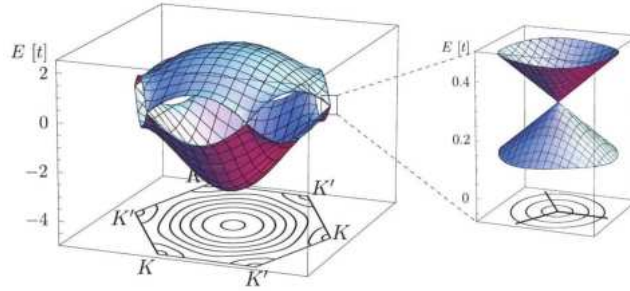


Figure G. 3 - Left: Band structure of graphene in next-nearest neighbor approximation for $V_0 = M = 0$. Right: zoom-in of the band structure close to the K -point.

where $\sigma^{x,y,z}$ are the Pauli matrices. Here we can see that third-nearest neighbor hopping would only contribute terms of the form σ^x and σ^y , just like the nearest neighbor hopping. This is due to the fact that third-nearest neighbor hopping again connects atoms in different sublattices. In contrast, next-nearest neighbor hopping, connecting atoms of the same sublattice, introduces new terms in this Hamiltonian. In this sense, the inclusion of next-nearest neighbor hopping is more important than third-nearest neighbors.

This decomposition of the Hamiltonian in terms of Pauli matrices is just a formal paraphrase, as any 2×2 Hermitian matrix can be written as a sum of Pauli matrices and unity. Still, since this form is reminiscent of a Hamiltonian involving spin, the twocomponent character of the wave function $(\varphi_A, \varphi_B)^t$ due to the sublattice structure of graphene is commonly referred to as pseudospin. This analogy to spin will become even more apparent in the next section, when we consider a low-energy approximation to the Hamiltonian.

Having evaluated $H(\mathbf{k})$ we can now turn to solving the eigenproblem in Eq. (G.19). The eigenenergies are

$$\begin{aligned} E_{\pm}(\mathbf{k}) &= V_0 - t' \left(2 \cos(k_x a) + 4 \cos\left(\frac{k_x a}{2}\right) \cos\left(\frac{\sqrt{3} k_y a}{2}\right) \right) \\ &\pm \sqrt{M^2 + t^2 \left(3 + 2 \cos(k_x a) + 4 \cos\left(\frac{k_x a}{2}\right) \cos\left(\frac{\sqrt{3} k_y a}{2}\right) \right)} \end{aligned} \quad (\text{G.24})$$

We first consider the case where $V_0 = 0$ (a finite V_0 simply shifts the zero of energy) and $M = 0$. In Fig. G. 3 we show $E_{\pm}(\mathbf{k})$ for $\mathbf{k}$ in the first Brillouin zone of graphene.

The two bands $E_+(\mathbf{k})$ and $E_-(\mathbf{k})$ are always restricted to positive and negative energies, respectively, and are called the conduction and valence band. The conduction and the valence band touch at the six corners of the Brillouin zone for $E = 0$. Since the two atoms in the unit cell each contribute one electron, the valence band is completely filled. Thus, the Fermi energy is given as $E_F = 0$, and states at the Fermi energy have wave vector $\mathbf{K}$ and $\mathbf{K}'$.

As seen from Fig. G.3, in our approximation the valence band has a larger band width than the conduction band. This is opposite to the usual band structure calculations for graphene, where the conduction band features a larger band width than the valence band [35, 190, 243]. This difference is due to the neglect of overlap from our side. However, for transport we are mainly concerned with states close to the Fermi energy, where overlap has only little effect³, and we will focus on this regime in the following.

Expanding the energy bands $E_{\pm}(\mathbf{k})$ to linear order around the two K -points, we obtain

$$E_{\pm}(\mathbf{K} + \mathbf{q}) = E_{\pm}(\mathbf{K}' + \mathbf{q}) \approx \hbar|\mathbf{q}|v_F \quad (\text{G.25})$$

where $v_F = \sqrt{3}ta/2\hbar$ is the Fermi velocity. Thus, close to the K -points, the energy spectrum is linear, as shown in the zoom-in of Fig. G.3. In terms of band structure, the full energy cone is obtained by combining three 120° segments from the three equivalent corners of the Brillouin zone. Furthermore, t' does not enter in Eq. (G.25), as next-nearest neighbor hopping only introduces terms quadratic or higher in $\mathbf{q}$ [35].

A linear energy spectrum is known from massless relativistic particles where $E = \pm\hbar kc$, and c is the speed of light. Close to the Fermi energy, the charge carriers in graphene thus have a "relativistic" energy dispersion, however with a much smaller "effective speed of light" $v_F \approx 10^6$ m/s. Because of this analogy, the cone-like shape of the energy bands around the K -points is commonly called Dirac cone. States within the Dirac cone at the K and K' -point are said to be in the K and K' -valley, respectively.

Finally, we consider the effect of a finite staggered potential. As seen from Eq. (G.24), the staggered potential opens a gap of size $2|M|$ at the corners of the Brillouin zone. (Fig. G.4). In this case, the spectrum close to the K -points can be approximated as

$$E_{\pm}(\mathbf{K} + \mathbf{q}) = E_{\pm}(\mathbf{K}' + \mathbf{q}) \approx \pm \sqrt{(M/v_F^2)^2 v_F^4 + (\hbar|\mathbf{q}|)^2 v_F^2} \quad (\text{G.26})$$

which is analogous to the energy spectrum of a massive relativistic particle, $E = \pm\sqrt{m_0^2 c^4 + (\hbar k)^2 c^2}$, where m_0 is the rest mass of the particle. Hence, the staggered potential introduces a "mass" M/v_F^2 for the charge carriers in graphene.

Although the low-energy spectrum of graphene resembles relativistic energy dispersions, one should not erroneously conclude that this is due to relativistic effects in

¹²³ When taking into account the overlap of neighboring orbitals, the Schrödinger equation in tightbinding approximation takes the form $H\varphi = ES\varphi$, where S is the overlap matrix. Roughly speaking, the smaller the energy E , the smaller the influence of the overlap, as E multiplies S . Thus, close to $E = 0$, the influence of overlap can be neglected. For example, this is reflected in Fig. 2 in Ref. [190], where the results of band structure calculations including and neglecting overlap are compared.

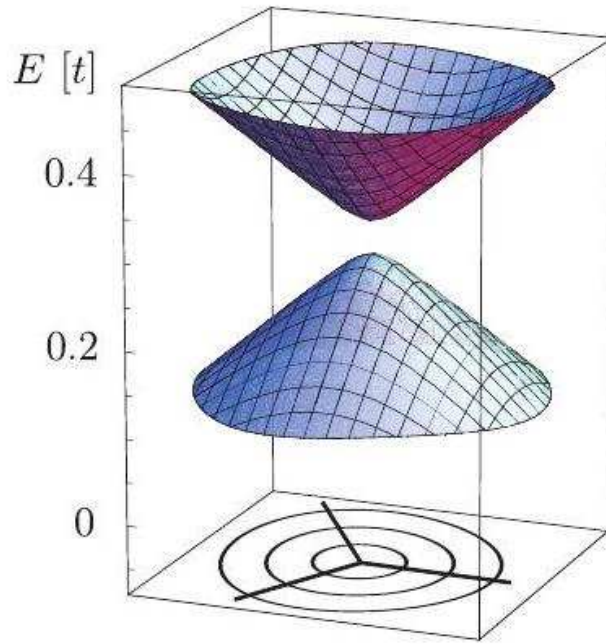


Figure G. 4 - The staggered potential opens a gap in the band structure close to the K points.

the motion of the valence electrons in the lattice of carbon atoms. Indeed, the starting point of the derivation was the non-relativistic Hamiltonian (G.5), and the linear low-energy spectrum is solely due to the lattice structure of the honeycomb network. Nevertheless, the Bloch electrons do behave in a way analogous to relativistic particles. In fact, this analogy goes even further than just a linear energy dispersion, as shown in the next section.

19.14.5 G.2.3 Effective Hamiltonian

The band structure of the previous section was obtained for a system with translational invariance. If this translational invariance is broken by, for example, a potential, it is in general not possible any more to solve the tight-binding problem analytically, and one has to resort to numerics or approximations. Here, we will derive such an approximation valid for slowly varying potentials and energies close to the Dirac point. This effective Hamiltonian for graphene was first derived by McClure including a magnetic field [244] and later by DiVincenzo and Mele including an external potential [245]. Both accounts followed the approach by Luttinger and Kohn for deriving an effective mass Hamiltonian [221], and so will we below.

In contrast to the lattice-periodic part of the Hamiltonian, a general potential may scatter Bloch electrons, connecting different Bloch vectors $\mathbf{k}$ and $\mathbf{k}'$. Thus, we have to generalize the ansatz of Eq. (G.18) to include a range of wave vectors:

$$\varphi(\mathbf{x}) = \int d\mathbf{k} \varphi_A(\mathbf{k}) \Phi_A(\mathbf{k}, \mathbf{x}) + \varphi_B(\mathbf{k}) \Phi_B(\mathbf{k}, \mathbf{x}) \quad (\text{G.27})$$

Note that this expression is still exact, as the $\Phi_\alpha(\mathbf{k}, \mathbf{x})$ form a complete set within the tight-binding subspace. Assuming energies close to the Dirac point and a potential that varies slowly on the scale of the lattice, we can neglect scattering events between the K and K' point, as they are separated by a wave vector on the order of the reciprocal lattice. Then, the ansatz (G.27) can be reduced to contain only wave vectors close to $\mathbf{K}$ and $\mathbf{K}'$ and we may write ⁴

$$\begin{aligned}\varphi(\mathbf{x}) = & \int d\mathbf{q} \varphi_A(\mathbf{q}) \Phi_A(\mathbf{K} + \mathbf{q}, \mathbf{x}) + \varphi_B(\mathbf{q}) \Phi_B(\mathbf{K} + \mathbf{q}, \mathbf{x}) \\ & + \varphi'_A(\mathbf{q}) \Phi_A(\mathbf{K}' + \mathbf{q}, \mathbf{x}) + \varphi'_B(\mathbf{q}) \Phi_B(\mathbf{K}' + \mathbf{q}, \mathbf{x})\end{aligned}\quad (\text{G.28})$$

When we use the ansatz (G.28) in Eq. (G.5), we can again make use of the fact that the tight-binding sums form an orthonormal set (see Eq. (G.17)). Thus, we need to calculate matrix elements of the form $\langle \Phi_\alpha(\mathbf{K}' + \mathbf{q}) | H | \Phi_\beta(\mathbf{K}' + \mathbf{q}') \rangle$. Since the lattice periodic part of the Hamiltonian does not mix different wave vectors, all corresponding matrix elements vanish for $\mathbf{q} \neq \mathbf{q}'$, and we obtain the result of the previous section, Eq. (G.23). However, for the external potential V_{ext} , we need to consider also matrix elements between different Bloch vectors. Making use of the smoothness of V_{ext} we can neglect matrix elements between different K -points, as was already done in the ansatz (G.28). Then, we find for the K -point

$$(E - H(\mathbf{K} + \mathbf{q})) \begin{pmatrix} \varphi_A(\mathbf{q}) \\ \varphi_B(\mathbf{q}) \end{pmatrix} + \int d\mathbf{q}' (V(\mathbf{q}, \mathbf{q}') + M(\mathbf{q}, \mathbf{q}') \sigma^z) \begin{pmatrix} \varphi_A(\mathbf{q}') \\ \varphi_B(\mathbf{q}') \end{pmatrix} = 0 \quad (\text{G.29})$$

where we again decomposed the external potential V_{ext} into a potential $V(\mathbf{x})$ that has the same value on sublattice A and B, and a staggered potential $M(\mathbf{x})$. $V(\mathbf{q}, \mathbf{q}')$ and $M(\mathbf{q}, \mathbf{q}')$ are defined as $\langle \Phi_A(\mathbf{K} + \mathbf{q}) | V | \Phi_A(\mathbf{K} + \mathbf{q}') \rangle$ and $\langle \Phi_A(\mathbf{K} + \mathbf{q}) | M | \Phi_A(\mathbf{K} + \mathbf{q}') \rangle$, respectively (These definitions only refer to sublattice A, but sublattice B is contained in the unit matrix and σ^z in Eq. (G.29)). The same equation holds for the K' -point, if we substitute $\mathbf{K}'$ for $\mathbf{K}$.

These equations can be simplified further by means of two approximations: First, we restrict ourselves to the low-energy regime and expand $H(\mathbf{K} + \mathbf{q})$ and $H(\mathbf{K}' + \mathbf{q})$, keeping only terms up to linear order in $\mathbf{q}$. For this, we introduce the notation

$$\mathbf{K}_\tau = \begin{pmatrix} \tau \frac{4\pi}{3a} \\ 0 \end{pmatrix} \quad (\text{G.30})$$

such that we can treat both the $K(\tau = +1)$ and the $K'(\tau = -1)$ in a unified way. Then we find from Eq. (G.23)

$$H(\mathbf{K}_\tau + \mathbf{q}) \approx \hbar v_F \begin{pmatrix} 0 & \tau q_x - i q_y \\ \tau q_x + i q_y & 0 \end{pmatrix} = \hbar v_F (\tau q_x \sigma^x + q_y \sigma^y). \quad (\text{G.31})$$

Second, we again make use of the assumption that the potential $V(\mathbf{x})$ varies smoothly on the lattice scale, so that we can replace $V(\mathbf{q}, \mathbf{q}')$ by $V(\mathbf{q} - \mathbf{q}')$, where $V(\mathbf{q}) = \frac{1}{(2\pi)^2} \int d\mathbf{x} e^{-i\mathbf{q}\mathbf{x}} V(\mathbf{x})$ is the Fourier transformation of $V(\mathbf{x})$, as discussed in Ref. [221].⁵ This also holds for the staggered potential $M(\mathbf{x})$.

With these approximations, Eq. (G.29) for the K -point reads

$$(E - \hbar v_F \mathbf{q} \cdot \boldsymbol{\sigma}) \begin{pmatrix} \varphi_A(\mathbf{q}) \\ \varphi_B(\mathbf{q}) \end{pmatrix} + \int d\mathbf{q}' (V(\mathbf{q} - \mathbf{q}') + M(\mathbf{q} - \mathbf{q}') \sigma^z) \begin{pmatrix} \varphi_A(\mathbf{q}') \\ \varphi_B(\mathbf{q}') \end{pmatrix} = 0 \quad (\text{G.32})$$

Introducing the envelope wave functions $\varphi_\alpha^{(j)}(\mathbf{x}) = \int d\mathbf{q} e^{i\mathbf{q}\mathbf{x}} \varphi_\alpha(\mathbf{q})$ and taking the Fourier transformation of Eq. (G.32), we finally arrive at

¹²⁴ Effective mass theory (also called $\mathbf{k} \cdot \mathbf{p}$ -theory) also uses often an ansatz different from Eq. (G.28): $\varphi(\mathbf{x}) \sim \int d\mathbf{q} \varphi(\mathbf{q}) e^{i\mathbf{q}\mathbf{x}} \Phi(\mathbf{K}, \mathbf{x})$. For example, this approach is used in the context of graphene in Refs. [244, 245]. As discussed in Ref. [221, Section II, III and Appendix A], these approaches are equivalent. However, the ansatz (G.28) seems more convenient for the present purposes.

$$(E - H_K) \begin{pmatrix} \varphi_A(\mathbf{x}) \\ \varphi_B(\mathbf{x}) \end{pmatrix} = 0 \quad (\text{G.33})$$

where H_K is the effective Hamiltonian for graphene at the K -point,

$$H_K = v_F(-i\hbar\nabla) \cdot \boldsymbol{\sigma} + V(\mathbf{x}) + M(\mathbf{x})\sigma^z \quad (\text{G.34})$$

This effective Hamiltonian is identical to the Dirac Hamiltonian in two spatial dimensions with an effective speed of light v_F [246]. The potential $V(\mathbf{x})$ plays the role of a scalar potential, whereas the staggered potential $M(\mathbf{x})$ acts as a mass term.

In the previous section, we have pointed out that the low-energy dispersion of graphene resembles relativistic spectra. This relativistic dispersion is a consequence of the fact that the effective low-energy Hamiltonian of graphene is the Dirac Hamiltonian. However, the Dirac Hamiltonian leads to effects beyond just a relativistic spectrum: In particular, the two components of the wave function arising from the two sublattices A and B behave like a spinor under transformation. Hence, the notion of pseudospin for the sublattice degree of freedom is justified. Several of the unique properties of graphene are a consequence of pseudospin, such as the odd-integer quantum Hall effect [32] and Klein tunneling [34].

For the K' -point we obtain the effective Hamiltonian

$$H_{K'} = -v_F(-i\hbar\nabla) \cdot \boldsymbol{\sigma}^* + V(\mathbf{x}) + M(\mathbf{x})\sigma^z. \quad (\text{G.35})$$

We can write the effective Hamiltonians for the K and K' -point in a unified way by defining

$$H_\tau = -i\hbar v_F (\tau \partial_x \sigma^x + \partial_y \sigma^y) + V(\mathbf{x}) + M(\mathbf{x})\sigma^z \quad (\text{G.36})$$

so that $H_K = H_+$ and $H_{K'} = H_-$. Note that the effective Hamiltonian does not contain any reference to the next-nearest neighbor hopping t' , as we only included terms linear in the momentum. Since t' only contributes starting from terms quadratic

in $\mathbf{q}$, the influence of next-nearest neighbor hopping is expected to be negligible in the low-energy regime (with the notable exception of the graphene edge state, as discussed in Chapter 6).

Finally, we establish a connection between the envelope wave functions $\varphi_\alpha^{(1)}(\mathbf{x})$ and the tight-binding wave function $\varphi_\alpha(j)$. From Eq. (G.28) and the definition of the tight-binding sums (G.16) we find

$$\begin{aligned} \varphi(\mathbf{x}) &= \sum_\alpha \int d\mathbf{q} \sum_{\mathbf{R}_i} \left(e^{i(\mathbf{K}+\mathbf{q})(\mathbf{R}_i+\mathbf{d}_\alpha)} \varphi_\alpha(\mathbf{q}) + e^{i(\mathbf{K}'+\mathbf{q})(\mathbf{R}_i+\mathbf{d}_\alpha)} \varphi'_\alpha(\mathbf{q}) \right) \\ &\quad \times \chi_{2p_z}(\mathbf{x} - \mathbf{R}_i - \mathbf{d}_\alpha) \\ &= \sum_\alpha \sum_{\mathbf{R}_i} \left(e^{i\mathbf{K}(\mathbf{R}_i+\mathbf{d}_\alpha)} \varphi_\alpha(\mathbf{R}_i + \mathbf{d}_\alpha) + e^{i\mathbf{K}'(\mathbf{R}_i+\mathbf{d}_\alpha)} \varphi'_\alpha(\mathbf{R}_i + \mathbf{d}_\alpha) \right) \\ &\quad \times \chi_{2p_z}(\mathbf{x} - \mathbf{R}_i - \mathbf{d}_\alpha) \end{aligned} \quad (\text{G.37})$$

Comparing this expression to Eq. (G.10), we can identify

$$\varphi_\alpha(j) = e^{i\mathbf{K}(\mathbf{R}_j+\mathbf{d}_\alpha)} \varphi_\alpha(\mathbf{R}_j + \mathbf{d}_\alpha) + e^{i\mathbf{K}'(\mathbf{R}_j+\mathbf{d}_\alpha)} \varphi'_\alpha(\mathbf{R}_j + \mathbf{d}_\alpha). \quad (\text{G.38})$$

¹²⁵ There are some subtleties for the case of graphene: Graphene is a two-dimensional crystal, but the p_z -orbitals and the potentials also extend into the third dimension. A two-dimensional potential is then obtained by averaging the potential over the electron density in z -direction [245]. Here, we simply assume that such an averaging procedure has already been performed and assume $V(\mathbf{x})$ and $M(\mathbf{x})$ as two-dimensional potentials.

19.14.6 G. 3 Graphene nanoribbons

19.14.7 G.3.1 Band structure

In this section we discuss how to understand the general features of the band structure of graphene nanoribbons from geometrical considerations and the band structure of an extended graphene sheet (for a similar discussion, see [169]).

A graphene nanoribbon (GNR) is a stripe cut out of the graphene honeycomb network. The boundary of the GNR depends on the direction of the cut. There are two basic types of GNRs: The GNR with zigzag edges (obtained by cutting e.g. in x direction) and with armchair edges (obtained by cutting e.g. in y -direction) as shown in Figs. G.5(a) and (b), respectively. Since a GNR does not have the full translational symmetry of the graphene lattice, the electronic states of the GNR will in general be a linear combination of Bloch waves from the extended graphene sheet, with different wave vectors such that they satisfy the boundary condition of the nanoribbon.

We first consider the zigzag nanoribbon. The zigzag GNR has translational symmetry in x -direction, with a unit cell of length a , as indicated in Fig. G.5(a), and thus $k_x \in [-\pi/a, \pi/a]$ is a good quantum number. The (bulk) states of the zigzag GNR are then a linear combination of Bloch waves with a fixed value of k_x , as indicated in the sketch of the Brillouin zone in Fig. G.5(a). In particular, the different K -points project uniquely to the k_x -axis, with $k_x = -\frac{2\pi}{3a}$ for the K -point and $k_x = +\frac{2\pi}{3a}$ for the K' -point.⁶ Thus, the K and K' -points are decoupled in zigzag nanoribbons. In the

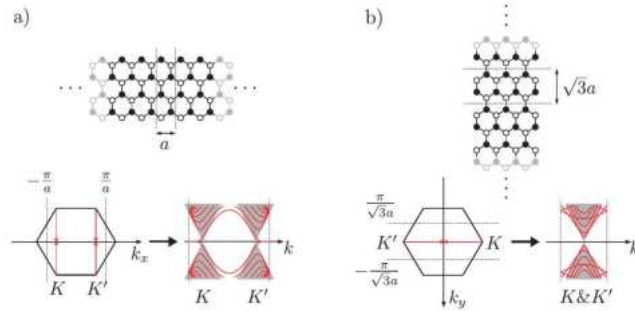


Figure G. 5 - The general features of the band structure of (a) zigzag and (b) armchair nanoribbons can be deduced from geometrical considerations. The upper panels show the graphene nanoribbons and the corresponding unit cells. The lower panels show the orientation of the one-dimensional nanoribbon Brillouin zone within the two-dimensional graphene Brillouin zone and the nanoribbon band structure.

low energy regime of the zigzag GNR band structure, we observe the Dirac cones at these points (sketched in grey in the band structure of Fig. G.5(a)).

In addition to the states obtained as linear combinations of bulk Bloch waves, the zigzag GNR band structure also features states close to $E = 0$ for $k_x < -\frac{2\pi}{3a}$ and $k_x > \frac{2\pi}{3a}$. These are states located at the nanoribbon edge and cannot be expressed as a linear combination of bulk Bloch waves, and are discussed in detail in Chapter 6.

In contrast to a zigzag ribbon, the armchair ribbon has translational symmetry in y -direction, with a unit cell of length $\sqrt{3}a$. Then, $k_y \in [-\frac{\pi}{\sqrt{3}a}, \frac{\pi}{\sqrt{3}a}]$ is a good quantum number. The electronic states of the armchair nanoribbon are thus linear combinations of bulk Bloch waves with a fixed k_y . In particular, both the K and K' -point map to $k_y = 0$. Hence, the armchair edge mixes the K and K' -valley.

¹²⁶ Note that this is not contradictory to our previous definition of $\mathbf{K} = (+\frac{4\pi}{3a}, 0)^t$ and $\mathbf{K}' = (-\frac{4\pi}{3a}, 0)^t$.

19.14.8 G.3.2 Boundary conditions

In order to handle finite-size graphene systems within the effective Dirac Hamiltonian, we need to derive appropriate boundary conditions. This will be done for the representative examples of zigzag and armchair nanoribbons.

The one-dimensional Brillouin zone of the nanoribbon has a smaller k_x -extent than the twodimensional Brillouin zone of graphene, and $k_x = +\frac{4\pi}{3a}$ and $k_x = -\frac{2\pi}{3a}$ are connected by a onedimensional reciprocal lattice vector of length $2\pi/a$.

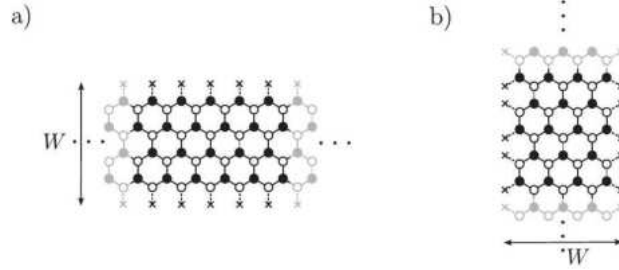


Figure G. 6 - Boundary conditions for (a) zigzag and (b) armchair nanoribbons: The wave function must vanish on the first row of missing atoms (indicated by crosses).

Since the effective Hamiltonian does not carry any reference to the next-nearest neighbor hopping, it is enough to take into account nearest neighbors when deriving the boundary condition. At an graphene edge, we therefore demand that the wave function must vanish on the missing atoms that are adjacent to the edge.

In the case of the zigzag nanoribbon, Fig. G.6(a), the edge atoms are all from the same sublattice, and so is the first row of missing atoms. Thus, the boundary conditions are

$$e^{iK_x x} \varphi_A(x, 0) + e^{iK'_x x} \varphi'_A(x, 0) = 0 \quad (\text{G.39a})$$

$$e^{iK_x x} \varphi_B(x, W) + e^{iK'_x x} \varphi'_B(x, W) = 0 \quad (\text{G.39b})$$

where we used the connection between the envelope functions from the effective Hamiltonian and the tight-binding wave function (G.38). Here, we chose a coordinate system such that the first row of missing atoms is located at $y = 0$, and W is the distance to the first row of missing atoms on the other edge [53] (see Fig. G.6(a)). Since $K_x = -K'_x$ and Eqs. (G.39) must hold for any point x on the boundary, the envelope functions $\varphi_\alpha^{(1)}(x)$ must satisfy the boundary conditions separately for the zigzag ribbon:

$$\varphi_A(x, 0) = 0, \quad (\text{G.40a})$$

$$\varphi'_A(x, 0) = 0, \quad (\text{G.40b})$$

$$\varphi_B(x, W) = 0, \quad (\text{G.40c})$$

$$\varphi'_B(x, W) = 0. \quad (\text{G.40d})$$

In the case of the armchair ribbon, both types of sublattices are present at the boundary, and hence the boundary condition reads

$$\varphi_A(x, 0) + \varphi'_A(x, 0) = 0, \quad (\text{G.41a})$$

$$\varphi_B(x, 0) + \varphi'_B(x, 0) = 0, \quad (\text{G.41b})$$

$$e^{iK_x W} \varphi_A(x, W) + e^{-iK_x W} \varphi'_A(x, W) = 0, \quad (\text{G.41c})$$

$$e^{iK_x W} \varphi_B(x, W) + e^{-iK_x W} \varphi'_B(x, W) = 0. \quad (\text{G.41d})$$

20 Diagram technique for nonequilibrium processes by Keldysh

Abstract

A graph technique analogous to the usual Feynman technique in field theory is developed for calculating Green's functions for particles in a statistical system which under the action of an external field deviates to any arbitrary extent from the state of thermodynamic equilibrium. It is found that in order to describe such a system it is necessary to introduce two Green's functions for each type of particles. The equation for one of these functions is a generalization of the usual Boltzmann kinetic equation.

THE use of quantum field theory methods has turned out to be very effective for the formulation and for the solution of problems in statistical physics. A number of prescriptions has been given for the description of thermodynamic characteristics of equilibrium systems [1-6] and of the ground state of a many-body system [7-9]. In the papers of Konstantinov and Perel, [10], Dzyaloshinski [11] and Mills [12] a technique has been developed for the calculation of time dependent quantities (kinetic coefficients) in systems in thermodynamic equilibrium. In carrying out this program the deviation of the system from the Gibbs equilibrium distribution was assumed to be small, i.e., the external field was taken to be sufficiently small. The present paper is devoted to the application of analogous methods to the description of essentially nonequilibrium systems, for example, such systems as the system of electrons in a strong electric field. In the simplest cases in order to describe such systems Boltzmann's kinetic equation or the system of coupled equations proposed by Bogolyubov [13] are used.

We shall obtain below a system of equations for the Green's functions of the particles in the system, and we shall show that one of these equations is essentially analogous to the kinetic equation. The appearance of such an equation-an additional one in comparison with systems in thermodynamic equilibrium-in the problem under consideration is quite natural, since we must have an equation which describes the distribution of particles in the system. In the equilibrium case this is given by the Gibbs canonical distribution. Our diagram technique will be close to Mills' technique for equilibrium systems. This technique has the advantage that it is completely analogous to the usual Feynman technique in field theory, with the only difference that the number of Green's functions appearing in it is increased. In the course of the discussion there appears automatically a function which describes the distribution of the particles in the system, and the equation for this function plays the role of the kinetic equation.

Following the method of Konstantinov and Perel' we shall define the density matrix ρ for the system by the usual equation

$$i\partial\rho/\partial t = [H_i, \rho]_- \equiv H_i(t)\rho(t) - \rho(t)H_i(t) \quad (1)$$

with the boundary condition

$$\rho(t = -\infty) = \rho_0 = \exp \{ \Psi_0 - H_0(-\infty) \} / kT \quad (2)$$

where $H = H_0 + H_i$ is the Hamiltonian of the system

$$H_0(t) = \int \psi_0 + (\mathbf{r}t) \left\{ \varepsilon \left(-i\hbar\nabla - \frac{e}{c}\mathbf{A} \right) + e(\mathbf{D}) \right\} \times \psi_0(\mathbf{r}t) d\mathbf{r} + H_T \quad (3)$$

The first term in H_0 is the energy of noninteracting electrons in an external electromagnetic field

$\{A(\mathbf{r}t), \Phi(\mathbf{r}t)\}$, the second is the Hamiltonian of the heat bath in contact with the system under consideration, $\epsilon(p)$ is the dispersion law for the electrons, Ψ_0 is the initial free energy.

The field operators ψ_0 and ψ_0^+ are defined in the following manner: let $\varphi_{\mathbf{p}}(\mathbf{r}t)$ be the complete system of functions determined by the equation

$$\left\{ \varepsilon \left(-i\nabla - \frac{e}{c} \mathbf{A}(\mathbf{r}t) \right) + e\Phi(\mathbf{r}t) - i\frac{\partial}{\partial t} \right\} \varphi_{\mathbf{p}}(\mathbf{r}t) = 0 \quad (4)$$

and by the boundary condition at $t \rightarrow -\infty$

$$\varphi_{\mathbf{p}}(\mathbf{r}t) \rightarrow \exp \{ i(\mathbf{p}\mathbf{r} - \varepsilon_{\mathbf{p}}t) \}, \quad (5)$$

where we consider the external field to be switched off at $t \rightarrow -\infty$, i.e., $(A, \Phi) \rightarrow 0$. Then the field operator $\psi_0(\mathbf{r}t)$ is defined by the relation

$$\psi_0(\mathbf{r}t) = \sum_{\mathbf{p}} a_{\mathbf{p}} \varphi_{\mathbf{p}}(\mathbf{r}t) \quad (6)$$

where $a_{\mathbf{p}}$ are the usual Fermi operators which satisfy the commutation relations

$$[a_{\mathbf{p}}^+ a_{\mathbf{p}'}]_+ \equiv a_{\mathbf{p}}^+ a_{\mathbf{p}'} + a_{\mathbf{p}'} a_{\mathbf{p}}^+ = \delta_{\mathbf{p}\mathbf{p}'} \quad (7)$$

Utilizing the fact that the functions $\varphi_{\mathbf{p}}(\mathbf{r}t)$ form at any arbitrary instant of time a complete orthogonal system, and formulas (6) and (7), it can be easily shown that for coincident times

$$[\psi_0^+(\mathbf{r}t), \psi_0(\mathbf{r}'t)]_+ = \delta(\mathbf{r} - \mathbf{r}') \quad (8)$$

and in virtue of this the operators $\psi_0(\mathbf{r}t)$ satisfy the free equations of motion

$$i\partial\psi_0/\partial t = [\psi_0, H_0]_- \quad (9)$$

i.e., we have from the outset defined the field operators in the interaction representation.

We note that the operator $H_0(t)$ differs in the Schrödinger representation and in the interaction representation in contrast to the usual case, when all the external fields are independent of the time. Having in mind such a definition of the field operators we have from the outset written equation (1) in the interaction representation leaving in it only the operator for the interaction energy H_i which, for the sake of definiteness, we shall in future write in the form

$$H_i(t) = g \int \psi_0^+(\mathbf{r}t) \psi_0(\mathbf{r}t) \varphi_0(\mathbf{r}t) d\mathbf{r} \quad (10)$$

where g is a dimensionless coupling constant. Such an expression describes the interaction between electrons and phonons in solids, and also can be utilized to describe the Coulomb interaction of charged particles if we write a separate equation for the Coulomb field.

Equation (1) for the density matrix $\rho(t)$ can be formally solved with the aid of the S-matrix

$$S(t, -\infty) = T \exp \left\{ -\frac{i}{\hbar} \int_{-\infty}^t H_i(\tau) d\tau \right\} \quad (11)$$

which satisfies the equation

$$i\partial S/\partial t = H_i(t)S(t) \quad (12)$$

The symbol T in (11) denotes a time-ordered product defined in the usual manner. Then we have

$$\begin{aligned}\rho(t) &= S(t, -\infty)\rho_0 S^+(t, -\infty) \\ &= S(t, -\infty)\rho_0 S(-\infty, t)\end{aligned}\quad (13)$$

The density matrix defined in this manner depends explicitly on the time. The average value of an arbitrary operator $L_0(t)$ at time t has the form

$$\langle L_0(t) \rangle = \text{Tr} \{ \rho(t) L_0(t) \} \quad (14)$$

where the subscript zero on the operator L_0 shows that this operator is taken in the interaction representation, i.e., its time dependence is determined by the free equation of motion in the external field

$$i\partial L_0/\partial t = [L_0(t), H_0(t)]_- \quad (15)$$

since the density matrix $\rho(t)$ itself was defined by us in the interaction representation. In future, however, we shall have occasion to deal with correlation functions of several field operators taken at different instants of time. In these cases it will be more convenient for us to transfer the entire time dependence to these operators, and to regard the density matrix as being independent of the time, i.e., to go over to the Heisenberg representation. In this case the field operators satisfy the complete equations of motion

$$i\partial\psi/\partial t = [\psi, H]_- \quad (16)$$

For the time-independent density matrix we can take the value of the matrix determined by expression (13) at a certain fixed instant of time, for example, $t = 0$, having thus included in it all the changes which the distribution $\rho_0(2)$ had undergone when the external field and the interaction in the system were switched on. In order to convince ourselves that such a procedure is correct we shall, for example, show that in the case of a system in thermodynamic equilibrium it leads to the Gibbs canonical distribution, as it ought. In order to do this we utilize the well-known relation ^[14,15]

$$\begin{aligned}S(0, -\infty)H_0 S^+(0, -\infty) \\ = H - i\delta g \frac{dS(0, -\infty)}{dg} S^+(0, -\infty)\end{aligned}\quad (17)$$

where δ is a parameter describing the adiabatic switching-on of the interaction, i.e., the interaction constant g in (10) is considered to be time dependent, in accordance with $g(t) = g e^{\delta t}$. After going to the limit $\delta \rightarrow 0$ the second term in (17) turns into a c-number equal to the sum of all the vacuum loops and describing the correction to the free energy of the system ^[1-3]. Therefore, for the given case of an equilibrium system

$$\rho = \rho(0) = S(0, -\infty)\rho_0 S^+(0, -\infty) = e^{(\Psi-H)/kT},$$

where H is the total Hamiltonian, while Ψ is the total free energy.

Having defined the density matrix as $\rho(0)$ we must utilize such Heisenberg field operators $\psi(\mathbf{r}, t)$, which at $t = 0$ go over into the free $\psi_0(\mathbf{r}t)$. In other words,

$$\psi(\mathbf{r}t) = S(0, t)\psi_0(\mathbf{r}t)S(t, 0) \quad (19)$$

where, by definition,

$$\begin{aligned}
S(t', t) &= T \exp \left\{ -i \int_t^{t'} H_i(\tau) d\tau \right\} \\
&= S(t', -\infty) S^+(t, -\infty)
\end{aligned} \tag{20}$$

Utilizing these relations and the definition of the Heisenberg density matrix ρ , the average value of the T-product of arbitrary Heisenberg operators $L(t), M(t') \dots$ can be brought to the form

$$\begin{aligned}
\langle TL(t)M(t') \dots \rangle &= \text{Tr} \{ \rho TL(t)M(t') \dots \} \\
&= \text{Tr} \{ S(0, -\infty) \rho_0 S(-\infty, 0) TL(t)M(t') \dots \}
\end{aligned}$$

or, commuting in the argument of Tr the operator $S(0, -\infty)$ and going over to operators in the interaction representation,

$$\begin{aligned}
\langle TL(t)M(t') \dots \rangle &= \text{Tr} \{ \rho_0 S(-\infty, 0) [TS(0, t)L_0(t)S(t, t')M_0(t') \dots] \\
&\times S(0, -\infty) \} = \text{Tr} \{ \rho_0 T_c [S_c L_0(t)M(t') \dots] \}
\end{aligned} \tag{21}$$

where T_c indicates ordering along the contour c , going from $-\infty$, passing the points $t, t' \dots$ and then returning back to $-\infty$; S_c is the complete S-matrix defined along the whole contour c . Ordering along c means that the points on the return branch of this contour correspond to later times than the points on the original direct branch, and of any two points on the return path the later one is that which is closer to $-\infty$.

However, evaluation of integrals along the contour c is inconvenient. Therefore, we perform in (21) the following identity transformation: we insert in the argument of T_c the factor $S(t_0, +\infty) \times S(+\infty, t_0)$ which is identically equal to unity, taking the point t_0 to lie on the contour c later than that one of the points $t, t' \dots$ which lies on the extreme right. This is equivalent to cutting the contour C to the right of the point lying on the extreme right and inserting into it a piece which goes to $+\infty$ and returns. The resultant contour C traverses the whole time axis from $-\infty$ to $+\infty$ and then back again from $+\infty$ to $-\infty$. Formula (21) is thus brought to the final form

$$\langle TL(t)M'(t') \dots \rangle = \text{Tr} \{ \rho_0 T_C (S_C L_0(t)M_0(t') \dots) \} \tag{22}$$

with all the arguments $t, t' \dots$ considered as lying on the positive $(-\infty, +\infty)$ branch of the contour C .

But use of the contour C enables us to write in the form of a T-product also the usual correlation function of two operators $L(t)$ and $M(t')$. Indeed, we shall assign to points lying on the positive (proceeding from $-\infty$ to $+\infty$) branch of the contour the subscript $+$, and to points lying on the negative branch the subscript $-$. Then evidently we have

$$\langle L(t)M(t') \rangle = \text{Tr} \{ \rho_0 T_C [S_C L_0(t_-)M_0(t'_+)] \} \tag{23}$$

since in accordance with the definition of the T_C -product the point t'_+ always precedes the point t_- . This possibility leads to the fact that the diagram technique obtained below enables us to evaluate directly not only the Green's functions for the particles, but also the usual average values of operators.

The averages of T-products of the type (22) and (23) can be decomposed in the usual manner ^[2] into a sum of Feynman graphs evaluated in accordance with the following rules.

1. An n -th order graph contains $2n$ points along the real time axis. Moreover, each point can lie either on the positive or on the negative branch of the contour C , and this is denoted respectively by the subscripts $+$ or $-$. An integration with respect to the coordinates of all the internal points is carried out over the whole volume and over time from $-\infty$ to $+\infty$. Each internal point is put in correspondence with a factor g , if its subscript is $+$, and with the factor $-g$, if its subscript is $-$.
2. One phonon and two electron lines emerge from (or enter) each point. Each of these lines is set in correspondence with one of four functions depending on the signs of the points which this line joins:
 - a) if the line goes from the point $-$ to the point $+$, it is set in correspondence with the function

$$\begin{aligned} G_0^+ (\mathbf{r}t_+, \mathbf{r}'t_-) &= -i \text{Tr} \{ \rho_0 T_C [\psi_0 (\mathbf{r}t_+), \psi_0 + (\mathbf{r}'t_-)] \} \\ &= i \langle \psi_0^+ (\mathbf{r}'t_+) \psi_0 (\mathbf{r}t_-) \rangle_0 \end{aligned} \quad (24)$$

the symbol $\langle \dots \rangle_0$ denotes averaging using the matrix ρ_0 ;

- b) if the line goes from the point $+$ to the point $-$, it is set in correspondence with the function

$$\begin{aligned} G_0^- (\mathbf{r}t_-, \mathbf{r}'t_+) &= -i \text{Tr} \{ \rho_0 T_c [\psi_0 (\mathbf{r}t_-) \psi_0^+ (\mathbf{r}'t_+)] \} \\ &= -i \langle \psi_0 (\mathbf{r}t_-) \psi_0^+ (\mathbf{r}'t_+) \rangle_0 \end{aligned} \quad (25)$$

- c) if the line goes from the point $+$ to the point $+$, then it is set in correspondence with the usual causal Green's function

$$\begin{aligned} G_0^c (\mathbf{r}t_+, \mathbf{r}'t_+) &= -i \text{Tr} \{ \rho_0 T_c [\psi_0 (\mathbf{r}t_+) \psi_0^+ (\mathbf{r}'t_+)] \} \\ &= -i \langle T \psi_0 (\mathbf{r}t_+) \psi_0^+ (\mathbf{r}'t_+) \rangle_0 \end{aligned} \quad (26)$$

- d) and, finally, the line connecting the point with the point $-$ is set in correspondence with the function

$$\begin{aligned} \tilde{G}_0^c (\mathbf{r}t_-, \mathbf{r}'t_-) &= -i \text{Tr} \{ \rho_0 T_C [\psi_0 (\mathbf{r}t_-) \psi_0^+ (\mathbf{r}'t_-)] \} \\ &= -i \langle \tilde{T} \psi_0 (\mathbf{r}t_-) \psi_0^+ (\mathbf{r}'t_-) \rangle_0 \end{aligned} \quad (27)$$

In the last formula $\tilde{T}$ denotes the product ordered with respect to the reversed time, i.e.,

$$\tilde{T} \psi_0(t) \psi_0^+(t') = \begin{cases} \psi_0(t) \psi_0^+(t'), & \text{if } t < t' \\ -\psi_0^+(t') \psi_0(t), & \text{if } t > t' \end{cases}$$

We shall henceforth omit everywhere the indices $+$ and $-$ in the time arguments of the Green's functions, since the index ascribed to the Green's function itself uniquely determines the positions of its final and initial points on the branches of the contour C . The functions G_0^c and $\tilde{G}_0^c$ are related to $G_0^\pm$ by the obvious relations

$$\begin{aligned} G_{0^c} (\mathbf{r}t, \mathbf{r}'t') &= \theta(t - t') G_0^- (\mathbf{r}t, \mathbf{r}'t') \\ &\quad + \theta(t' - t) G_0^+ (\mathbf{r}t, \mathbf{r}'t') \end{aligned} \quad (28)$$

$$\begin{aligned} \tilde{G}_0^c (\mathbf{r}t, \mathbf{r}'t') &= \theta(t - t') G_{0'}^- (\mathbf{r}t, \mathbf{r}'t') \\ &\quad + \theta(t' - t) G_0^- (\mathbf{r}t, \mathbf{r}'t') \\ \theta(x) &= 1/2(1 + \text{sign } x) \end{aligned} \quad (29)$$

The Green's functions for Bose particles are defined in an analogous manner.

The further rules for calculation are completely analogous to the usual Feynman rules, and, therefore, we shall not state them in detail. We only note that a diagram with $2n$ points contains the factor i^{n+2F} , where F is the number of closed loops in it, and that the unconnected diagrams must be omitted since they cancel with the trace of the density matrix which we have for this reason omitted in advance in the denominators of all the formulas (21)-(29).

It should also be noted that the applicability of Wick's theorem to the expansion of the T-products (22), (23) is not entirely obvious. Strictly speaking, when four or more field operators having the same momentum are averaged over the distribution ρ_0 , the T-product does not decompose into a sum of products of T-products taken a pair at a time. But for systems of Fermi-particles and of phonons (phonons correspond to a real field) considered by us Wick's theorem is completely rigorous. In other cases the contribution of two or more pairs of operators whose momenta are the same is generally of order $1/N$, where N is the number of degrees of freedom, since integration is carried out over all the internal momenta. An exception can occur in the case of systems in which phenomena of the type of Bose-condensation are observed. The technique under consideration is not directly applicable to these systems.

We now define the complete Green's functions:

$$G^c(\mathbf{r}t, \mathbf{r}'t') = -i \text{Tr} \{ \rho_0 T_C [\psi_0(\mathbf{r}t_+) \psi_0^+(\mathbf{r}'t'_+) S_C] \}, \quad (30)$$

$$\tilde{G}^c(\mathbf{r}t, \mathbf{r}'t') = -i \text{Tr} \{ \rho_0 T_C [\psi_0(\mathbf{r}t_-) \psi_0^+(\mathbf{r}'t'_-) S_C] \}, \quad (31)$$

$$G^\pm(\mathbf{r}t, \mathbf{r}'t') = -i \text{Tr} \{ \rho_0 T_C [\psi_0(\mathbf{r}t_\pm) \psi_0^+(\mathbf{r}'t'_\pm) S_C] \}, \quad (32)$$

which are sums of all the connected diagrams with one entering and one emerging line which differ from one another by the location of the initial and the final points on different branches of the contour C . The following relations involving them are preserved

$$G^c(\mathbf{r}t, \mathbf{r}'t') = \theta(t - t') G^-(\mathbf{r}t, \mathbf{r}'t') + \theta(t' - t) G^+(\mathbf{r}t, \mathbf{r}'t') \quad (33)$$

$$(34)$$

$$\tilde{G}^c(\mathbf{r}t, \mathbf{r}'t') = \theta(t - t') G^+(\mathbf{r}t, \mathbf{r}'t') + \theta(t' - t) G^-(\mathbf{r}t, \mathbf{r}'t')$$

i.e., only the two functions G^+ and G^- are independent, and even they are related by the obvious expression

$$\lim_{t' \rightarrow t} \{ G^+(\mathbf{r}t, \mathbf{r}'t') - G^-(\mathbf{r}t, \mathbf{r}'t') \} = -i\delta(\mathbf{r} - \mathbf{r}') \quad (35)$$

The subsequent technique of calculating the various quantities and the form of the equations are both considerably simplified if, following Mills ^[12], we go over to the matrix form of writing the equations. We define a two-rowed matrix G in the following manner:

$$G = \begin{pmatrix} G^c & G^- \\ G^+ & \tilde{G}^c \end{pmatrix} \quad (36)$$

Then, as can be easily seen, summation over the subscript $\pm$ at each point reduces to matrix multiplication of matrices corresponding to lines emerging from that point. But since there are three such lines—two electron lines and one phonon line—it is necessary to connect them by means of a vertex matrix of the third rank

$$\gamma_{ij}^k = \delta_{ij}(\sigma_z)_{jk} \quad (37)$$

where the subscripts refer to the electron lines, and the superscript to the phonon line; σ_z is the third Pauli matrix; the fact that for $k = 2$ it is negative takes into account the factor -1 for points on the return branch of the contour C (in matrix notation the points + correspond to subscripts 1, and the points - correspond to subscripts 2).

An arbitrary matrix element can now be written in the form of a product of blocks of the type

$$G_{li}\gamma_{ij}^k G_{jm} D_{kn}$$

in which summation is assumed to be carried out over repeated subscripts. D is the matrix for the Green's functions for Bose-particles defined in analogy with (36). With this notation we no longer need to distinguish the points denoted by + and -. The technique for the evaluation of any arbitrary graph now becomes a purely Feynman one with the only difference that to each line we now set in correspondence the matrix (36), and to each point the vertex matrix (37).

Summation of the diagrams for the complete Green's function leads to an equation of the type of Dyson's equation:

$$\begin{aligned} G(\mathbf{r}t, \mathbf{r}'t') &= G_0(\mathbf{r}t, \mathbf{r}'t') \\ &+ ig^2 \int G_0(\mathbf{r}t, \mathbf{r}_1 t_1) \Sigma(\mathbf{r}_1 t_1, \mathbf{r}_2 t_2) G(\mathbf{r}_2 t_2, \mathbf{r}'t') d\mathbf{r}_1 d\mathbf{r}_2 dt_1 dt_2 \end{aligned} \quad (38)$$

The self energy matrix Σ is defined by relation (39) in which the complete vertex matrix $\Gamma_{ij}^k(x, x'; y)$ is the sum of all the irreducible diagrams with one incoming (at the point x) and one outgoing (at the point x') electron line and one external phonon line at the point y :

$$\begin{aligned} \Sigma_{ij}(x, x') &= \int \gamma_{ii'k} G_{i,j}(x, x_1) \Gamma_{j'j'}(x_1, x'y) D_{k'k}(y, x) d^4 x_1 d^4 y \\ x &= (\mathbf{r}, t) \end{aligned} \quad (39)$$

Dyson's equation in integral form (38) appears to be unsatisfactory since it contains the function G_0 which explicitly depends on the initial distribution ρ_0 . At the same time from physical considerations it follows that if the system is in contact with a heat bath, then after sufficient time has elapsed it must reach a state which is, generally speaking, not stationary, since the external field depends on the time, but which is independent of the initial conditions. This inconsistency can be easily avoided by going over to the differential form of Dyson's equation, i.e., by multiplying (38) by the operator matrix

$$\hat{G}_0^{-1} = G_0^{-1}(x) \sigma_z \quad (40)$$

where the differential operator G_0^{-1} is defined by

$$G_0^{-1}(x) = i \frac{\partial}{\partial t} - \varepsilon \left(-i \nabla - \frac{e}{c} \mathbf{A}(x) \right) - e \Phi(x), \quad (41)$$

$x = (\mathbf{r}t)$. Then Eq. (38) is reduced to the form

$$\begin{aligned} \hat{G}_0^{-1}(x) G(x, x') &= ig^2 \int \Sigma(x, x_1) G(x_1, x') dx_1 + \delta(x - x') \end{aligned} \quad (42)$$

In deriving this formula we have utilized the obvious relations

$$\begin{aligned} G_0^{-1}(x)G_0^{\pm}(x, x') &= 0, \quad G_0^{-1}(x)G_0^c(x, x') = \delta(x - x'), \\ G_0^{-1}(x)G_0^c(x, x') &= -\delta(x - x') \end{aligned}$$

Equation (42) contains only complete Green's functions which do not explicitly depend on ρ_0 . At the same time its solution is determined only up to an arbitrary solution of the homogeneous equation

$$\hat{G}_0^{-1}(x)h(x, x') = 0 \quad (43)$$

We shall discuss the problem of the unique choice of the solution and of the role of initial conditions later.

Equation (42) represents a system of four in-tegro-differential equations. However, between the four unknown functions G_{ij} there exist two linear relations (33) and (34) which, naturally, are consistent with the equations. Therefore, by means of a certain linear canonical transformation of the matrices G, γ and D we can reduce the number of independent equations to two. Such a transformation is given by

$$G \rightarrow \frac{1 - i\sigma_y}{\sqrt{2}} \begin{pmatrix} G^c & G^- \\ G^+ & \tilde{G}^c \end{pmatrix} \frac{1 + i\sigma_y}{\sqrt{2}} = \begin{pmatrix} 0 & G^a \\ G^r & F \end{pmatrix}, \quad (44)$$

where three new Green's functions have been introduced: the advanced one G^a , the retarded one G^r , and the correlation function F which, as we shall see later, is essentially a single-particle density matrix. The matrix σ_y is the second of the three Pauli matrices:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The functions G^a, G^r , and F are defined as follows in terms of average values of Heisenberg operators using the Heisenberg density matrix:

$$G^a(x, x') = -i\theta(t - t') \langle [\psi(x), \psi^+(x')]_+ \rangle \quad (45)$$

$$G^r(x, x') = i\theta(t' - t) \langle [\psi(x), \psi^+(x')]_+ \rangle \quad (46)$$

$$F(x, x') = -i \langle [\psi(x), \psi^+(x')]_- \rangle \quad (47)$$

They are related to the G -functions used above by means of the relations

$$\begin{aligned} G^a &= G^c - G^+ = -\tilde{G}^c + G^- \\ &= -\theta(t - t') (G^+ - G^-) \end{aligned} \quad (48)$$

$$\begin{aligned} G^r &= G^c - G^- = -\tilde{G}^c + G^+ \\ &= \theta(t' - t) (G^+ - G^-) \end{aligned} \quad (49)$$

$$F = G^c + \tilde{G}^c = G^+ + G^- \quad (50)$$

The D -matrix transforms in a similar fashion. The vertex matrix γ_{ij}^k and the operator $\hat{G}_0^{-1}$ are also altered:

$$\gamma_{ij}^1 = 2^{-1/2} \delta_{ij}, \quad \gamma_{ij}^2 = 2^{-1/2} (\sigma_x)_{ij} \quad (51)$$

$$\hat{G}_0^{-1} = \sigma_x G_0^{-1}(x) \quad (52)$$

The self energy matrix Σ is brought to the form

$$\Sigma = \begin{pmatrix} \Omega & \Sigma^r \\ \Sigma^a & 0 \end{pmatrix} \quad (53)$$

$$\begin{aligned} \Sigma^r &= \Sigma^c + \Sigma^- = -(\tilde{\Sigma}^c + \Sigma^+) \\ \Sigma^a &= \Sigma^c + \Sigma^+ = -(\tilde{\Sigma}^c + \Sigma^-) \\ \Omega &= \Sigma^c + \tilde{\Sigma}^c = -(\Sigma^+ + \Sigma^-) \end{aligned} \quad (54)$$

The difference between (53), (54) and (44), (48)-(50) is related to the fact that the quantities $\Sigma^\pm$ in contrast to Σ^c and $\tilde{\Sigma}^c$ contain the factor -1, which appears in the vertex (37) for $k = 2$, an odd number of times.

The general equations (39) and (42) are obviously invariant under the transformation (44); we simply go over to a new representation for the matrices G and D . But the number of independent equations is now obviously equal to two since the functions G^a and G^r are Hermitian conjugates and the equations for them follow from each other. The physical meaning of the functions G^r and F is quite different. Roughly speaking, G^r and G^a characterize the dynamic properties of the particles in the system under consideration, while F characterizes their static distribution. This can already be seen from the definition of these functions in the absence both of an external field and of an interaction, i.e., for $t \rightarrow -\infty$. Then, going over to Fourier transforms with respect to the differences $\mathbf{r} - \mathbf{r}'$ and $t - t'$ we obtain

$$G_0(\mathbf{p}\varepsilon) = \int G_0(\mathbf{r} - \mathbf{r}', t - t') \exp \{ -i [\mathbf{p}(\mathbf{r} - \mathbf{r}') - \varepsilon(t - t')] \} d(\mathbf{r} - \mathbf{r}') d(t - t') \quad (55)$$

and in the limiting case mentioned above

$$G_0^r(\mathbf{p}\varepsilon) = G_0^{a*}(\mathbf{p}\varepsilon) = 1/(\varepsilon - \varepsilon_{\mathbf{p}} - i\delta) \quad \delta \rightarrow +0 \quad (56)$$

$$\begin{aligned} G_0^+(\mathbf{p}\varepsilon) &= 2\pi i n_{\mathbf{p}} \delta(\varepsilon - \varepsilon_{\mathbf{p}}) \\ G_0^-(\mathbf{p}\varepsilon) &= -2\pi i (1 - n_{\mathbf{p}}) \delta(\varepsilon - \varepsilon_{\mathbf{p}}) \end{aligned} \quad (57)$$

$$F_0(\mathbf{p}\varepsilon) = 2\pi i (2n_{\mathbf{p}} - 1) \delta(\varepsilon - \varepsilon_{\mathbf{p}}) \quad (58)$$

where

$$n_{\mathbf{p}} = \left[\exp \frac{\varepsilon_{\mathbf{p}} - \mu}{kT} + 1 \right]^{-1}$$

is the Fermi distribution function, μ is the chemical potential in the equilibrium state. Thus, the functions F and $G^\pm$ essentially represent the distribution function for the particles written in different forms.

When the interaction is taken into account the functions G^r and G^a also indirectly depend on the distribution of the particles, but the fact that $G^\pm(\mathbf{r}\mathbf{t}, \mathbf{r}'\mathbf{t})$ and $F(\mathbf{r}\mathbf{t}, \mathbf{r}'\mathbf{t})$ taken at coincident times $t' = t$ directly determine the distributions of electrons and of holes follows directly from their definition (32). We shall now show that the equation for G^+ (or the equation for F equivalent to it) contains within itself the usual Boltzmann kinetic equation under the assumptions under which the latter exists, i.e., when the external field is quasiclassical and the interaction is weak in the sense

$$\hbar/\tau \nless \bar{\epsilon} \quad (59)$$

where τ is the time taken to traverse a mean free path, and $\bar{\epsilon}$ is the average energy of the particles.

The equation for G^+ has the form

$$\begin{aligned} G_0^{-1}(x)G^+(x, x') &= ig^2 \int \left\{ \Sigma^+(x, x'') G^c(x'', x') \right. \\ &\quad \left. + \tilde{\Sigma}^c(x, x'') G^+(x'', x') \right\} dx'' \end{aligned} \quad (60)$$

The corresponding Hermitian conjugate equation is

$$\begin{aligned} G_0^{-1*}(x')G^+(x, x') &= ig^2 \int \left\{ \tilde{G}^c(x, x'') \Sigma^+(x'', x') \right. \\ &\quad \left. + G^+(x, x'') \Sigma^c(x'', x') \right\} dx'' \end{aligned} \quad (61)$$

We add these two equations, then go over from the variables x, x' to the new variables $(r_{stS}) = x_S = (x + x')/2$ and $(r_{at_a}) = x_a = (x - x')/2$, and then take a Fourier transform of the type (55) with respect to the difference variable x_a . As a result terms of the type

$$\begin{aligned} ig^2 \int \Sigma(x_s + x'_a, p') G(x_s - x_a + x'_a, p'') \exp \{ i[(p' - p)x_a \\ + (p'' - p')x_a] \} dx_a dx'_a dp' dp'', \quad p = (\mathbf{p}\epsilon) \end{aligned}$$

will appear on the right hand side of the equation.

The quasiclassical nature of the external field means that quantities of the type $p\nabla_S$ are small. At the same time the right hand side of the equation is proportional to the small interaction constant g^2 . Therefore, neglecting small terms of the form $g^2(p\nabla_S)$ and introducing a further change of variables $p \rightarrow p - eA(x_S)/c, \epsilon \rightarrow \epsilon - e\Phi(x_S)$, we obtain

$$\begin{aligned} \left\{ \frac{\partial}{\partial t_s} + \mathbf{v}_p \nabla_s + \left(e\mathbf{E} + \left[\frac{e\mathbf{v}_p}{c} \mathbf{H} \right] \right) \nabla_p \right\} G^+(x_s p) \\ = g^2 \left\{ \Sigma^+(x_s p) [G^c(x_s p) + \tilde{G}^c(x_s p)] \right. \\ \left. + [\Sigma^c(x_s p) + \tilde{\Sigma}^c(x_s p)] G^+(x_s p) \right\} \end{aligned} \quad (62)$$

where $\mathbf{v}_p = \nabla_p \epsilon_p$ is the velocity of the particles, $\mathbf{E}$ and $\mathbf{H}$ are the electric and the magnetic fields:

$$* [v_p H] = v_p \times H.$$

$$\mathbf{E} = -\nabla\Phi - c^{-1}\partial\mathbf{A}/\partial t; \quad \mathbf{H} = \text{rot } \mathbf{A} \quad (63)$$

In deriving (62) we have expanded in the left hand side of the equation the operator $G_0^{-1}(x) - G_0^{-1*}(x')$ in series in terms of the small difference $x_a = (x - x')/2$ up to the first order on the basis of the same quasiclassical condition. The right hand side of (62) can be further simplified by utilizing (50) and (54):

$$\begin{aligned} \left\{ \frac{\partial}{\partial t} + \mathbf{v} \nabla + e \left(\mathbf{E} + \left[\frac{\mathbf{v}}{c} \mathbf{H} \right] \right) \nabla_p \right\} G^+(xp) \\ = g^2 \{ \Sigma^+(xp) G^-(xp) - \Sigma^-(xp) G^+(xp) \} \end{aligned} \quad (64)$$

The right hand side of (64) describes the effect of collisions on the distribution of the particles, with the first term being related to particles entering the state of momentum $\mathbf{p}$ from all the other states with momenta $\mathbf{p}'$, while the second term is related to the particles leaving the state. In order to bring this equation to the usual form of the Boltzmann equation we introduce the distribution function $f(\mathbf{p}t)$ which, as has been noted previously, coincides with G^+ for coincident times $t' = t$:

$$f(\mathbf{p}t) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} G^+(\mathbf{r}t; \mathbf{p}\varepsilon) d\varepsilon \quad (65)$$

Integration of equation (64) over ε then yields

$$\begin{aligned} & \left\{ \frac{\partial}{\partial t} + \mathbf{v} \nabla + e \left(\mathcal{E} + \left[\frac{\mathbf{v}}{c} \mathbf{H} \right] \right) \nabla \mathbf{p} \right\} f(\mathbf{p}t) \\ &= \frac{g^2}{2\pi i} \int_{-\infty}^{\infty} \left\{ \Sigma^+(\mathbf{r}t; \mathbf{p}\varepsilon) G^-(\mathbf{r}t; \mathbf{p}\varepsilon) \right. \\ & \quad \left. - \Sigma^-(\mathbf{r}t; \mathbf{p}\varepsilon) G^+(\mathbf{r}t; \mathbf{p}\varepsilon) \right\} d\varepsilon \end{aligned} \quad (66)$$

In accordance with the assumption that the interaction is small we must evaluate the right hand side of (66) up to the first order in g^2 , i.e., we must substitute into (66) and (39) the functions $G_0^\pm$, satisfying the free equations

$$G_0^{-1} G_0 \pm (\mathbf{r}t; \mathbf{p}\varepsilon) = 0 \quad (67)$$

but normalized with respect to an as yet unknown distribution function f by equations (65) and (35). Symbolically they can be written in the form

$$G_0^+(\mathbf{r}t; \mathbf{p}\varepsilon) = 2\pi i \delta(\varepsilon - \varepsilon_{\mathbf{p}} + i\hat{B}) f(\mathbf{p}t) \quad (68)$$

$$G_0 - (\mathbf{r}t; \mathbf{p}\varepsilon) = -2\pi i \delta(\varepsilon - \varepsilon_{\mathbf{p}} + i\hat{B}) (1 - f(\mathbf{p}t)) \quad (69)$$

where $\hat{B}$ is the operator in the left hand side of (66). The quantity $\delta(G_0^{-1})$ in formulas (68), (69), is an integral operator with a kernel satisfying equation (67) and equal to $\delta(r - r')$ for $t' = t$. It can be easily expressed in terms of the functions φ_p (4).

In order to evaluate $\Sigma^\pm$ we must also define $D_0^\pm$. In the majority of problems of solid state physics the phonons can be regarded as being in thermodynamic equilibrium. Then

$$\begin{aligned} D_0 \pm (k\omega) &= -i\pi |m_{\mathbf{k}}|^2 \{ (1 + N_{\mathbf{k}}) \delta(\omega \pm \omega_{\mathbf{k}}) \\ & \quad + N_{\mathbf{k}} \delta(\omega \mp \omega_{\mathbf{k}}) \} \end{aligned} \quad (70)$$

where $\mathbf{k}$ and $\omega_{\mathbf{k}}$ are the propagation vector and the frequency of the phonon, $m_{\mathbf{k}}$ is the matrix element for the interaction with the electron, and

$$N_{\mathbf{k}} = \left[\exp \frac{\omega_{\mathbf{k}}}{kT} - 1 \right]^{-1} \quad (71)$$

is the Planck distribution function.

Substituting (68)-(70) into (66) we obtain

$$\begin{aligned}
\hat{B}f(\mathbf{prt}) = & 2\pi^2 g^2 \int |m_{\mathbf{k}}|^2 \{ [(1 + N_{\mathbf{k}}) \delta(\varepsilon_{\mathbf{p}} - \varepsilon_{\mathbf{p}+\mathbf{k}} + \omega_{\mathbf{k}}) \\
& + N_{\mathbf{k}} \delta(\varepsilon_{\mathbf{p}} - \varepsilon_{\mathbf{p}+\mathbf{k}} - \omega_{\mathbf{k}})] f(\mathbf{p} + \mathbf{k}, \mathbf{rt})(1 - f(\mathbf{prt})) \\
& - [(1 + N_{\mathbf{k}}) \delta(\varepsilon_{\mathbf{p}} - \varepsilon_{\mathbf{p}+\mathbf{k}} - \omega_{\mathbf{k}}) + N_{\mathbf{k}} \delta(\varepsilon_{\mathbf{p}} - \varepsilon_{\mathbf{p}+\mathbf{k}} + \omega_{\mathbf{k}})] \\
& \times (1 - f(\mathbf{p} + \mathbf{k}, \mathbf{rt})) f(\mathbf{prt}) \} d\mathbf{k} / (2\pi)^3
\end{aligned} \tag{72}$$

i.e., the usual form of Boltzmann's equation. Thus, in deriving it we have utilized two conditions: the quasiclassical nature of the external field

$$\frac{d}{dx}(e\mathbf{A}, e\Phi) \ll \varepsilon \bar{p} \tag{73}$$

and the smallness of the interaction

$$\bar{\varepsilon} \gg g^2 \Sigma^{\pm} \sim \hbar/\tau \tag{74}$$

The quasiclassical condition is not obligatory. Proceeding in exactly the same manner, but without expanding (60) and (61) in terms of the external field, we would have obtained the so-called quantum kinetic equation^[16,17]. But the condition of smallness of the interaction (74) is necessary in this case also. If it is violated, then the functions $G^{\pm}$ do not have a δ -like character (68), (69) as functions of ϵ , and the integration over ϵ in equation (66) cannot be effectively carried out. In such a case we shall not obtain, generally speaking, a closed equation for the distribution function $f(\mathbf{prt})$, i.e., for the G^+ -function for coincident times $t' = t$. The equation exists only for the complete G -function considered as a function of all its arguments. In the most general case this corresponds to Dyson's equations (39) and (42). If the interaction is not weak, they can be solved if the system has some other parameters, for example density, which enable us to select a set of principal diagrams. However, in this case there will arise the previously mentioned problem of the nonuniqueness of the solution of (42) up to an arbitrary solution of the free equation. We shall now show that such arbitrariness does not in fact exist, and that there is a unique prescription for the choice of the correct solution.

In order to do this we compare Eqs. (42) and (38) in the representation (44), by writing them separately for each of the functions G^a, G^r and F . In differential form

$$G_0^{-1} G^a = 1 + \Sigma^a G^a, \tag{75}$$

$$G_0^{-1} G^r = 1 + \Sigma^r G^r \tag{76}$$

$$G_0^{-1} F = \Omega G^a + \Sigma^r F \tag{77}$$

in integral form, corresponding to equation (38),

$$G^a = G_0^a (1 + \Sigma^a G^a) \tag{75a}$$

$$G^r = G_0^r (1 + \Sigma^r G^r) \tag{76a}$$

$$F = F_0 (1 + \Sigma^a G^a) + G_0^r (\Omega G^a + \Sigma^r F) \tag{77a}$$

Comparison of (75) - (77) with (75a) - (77a) shows that equations (75), (76) do not contain any arbitrariness: they are solved by the advanced operator G_0^a or the retarded operator G_0^r respectively. For $t' \rightarrow t$ their solutions must go over into $\delta(r - r')$. A possible lack of uniqueness, i.e., dependence on initial conditions, is associated only with the term proportional to F_0

in equation (77 a), since F_0 depends on the initial distribution, while in the differential form of the equation this term is completely absent. However, it can be easily seen that this term is identically equal to zero. Indeed, we substitute (75) into (77a). Then we have

$$F = F_0 G_0^{-1} G^a + G_0^r (\Omega G^a + \Sigma^r F) = G_0^r (\Omega G^a + \Sigma^r F), \quad (78)$$

since $F_0 G_0^{-1} = 0$. Thus, (77) is solved by means of the function G_0^r operating from the left. Equation (77) and its solution only appear to be nonsymmetric with respect to retarded and advanced functions. In place of (77) we can consider the Hermitian conjugate equation equivalent to it

$$G_0^{-1} F = G^r \Omega + F \Sigma^a \quad (79)$$

which is solved by the function G_0^a operating from the right. Thus, solutions of Eqs. (75) - (77) are determined uniquely and independently of the initial distribution at $t = -\infty$.

We should now make clear to what systems the results obtained above are applicable. Evidently, we are dealing with well established solutions, i.e., such solutions which are determined by the effect of the external field at the present and at earlier instants of time, by the interaction with the heat bath, but not by the initial conditions. Generally speaking, these are not stationary solutions if the external field depends on the time. The lack of dependence on the initial conditions means physically

that the initial conditions, even if they had been imposed, were specified in the sufficiently remote past, formally at $t \rightarrow -\infty$, while in fact we are dealing with times large in comparison to the relaxation time of the system. However, in discussions of kinetic phenomena, another way of posing the problem is possible: at time t_0 an arbitrary density matrix $\rho(t_0)$ is given and we are dealing with the relaxation of this state towards some well established regime. Such a problem, apparently, can not in principle be described by any diagram technique, since the latter is always based on the concept of the possibility of expressing two-, three-, etc. particle Green's functions in terms of the oneparticle Green's function at least in the form of an infinite series. But an arbitrary specification of the density matrix is equivalent to an arbitrary and independent specification of all these quantities.

The existence of well established solutions independent of initial conditions is evident in the case of systems in contact with a heat bath. In the case of an adiabatic system the problem is less clear and the technique under discussion is applicable with certainty only in the case when the system had been actually in a state of thermodynamic equilibrium and was then taken out of that state by a given external field. The problem of approach to equilibrium from an arbitrary initial state has been discussed in papers by Van Hove ^[18], Prigogine and Resibois ^[19], Fujita ^[20] and others.

In conclusion I wish to express my gratitude to V. L. Gurevich, I. E. Dzyaloshinskiĭ, D. A. Kirzhnits, and E. S. Fradkin for discussions of a number of points dealt with in the present article.

21 Hamiltonian approach to the transport properties of superconducting quantum point contacts by Cuevas, Martin-Rodero, Levy Yeyati

Abstract

A microscopic theory of the transport properties of quantum point contacts giving a unified description of the normal conductor-superconductor ($N-S$) and superconductor-superconductor ($S-S$) cases is presented. It is based on a model Hamiltonian describing

charge transfer processes in the contact region and makes use of nonequilibrium Green function techniques for the calculation of the relevant quantities. It is explicitly shown that when calculations are performed up to infinite order in the coupling between the electrodes, the theory contains all known results predicted by the more usual scattering approach for $N - S$ and $S - S$ contacts. For the latter we introduce a specific formulation for dealing with the nonstationary transport properties. An efficient algorithm is developed for obtaining the dc and ac current components, which allows a detailed analysis of the different current-voltage characteristics for all range of parameters. We finally address the less understood small bias limit, for which some analytical results can be obtained within the present formalism. It is shown that four different physical regimes can be reached in this limit depending on the values of the inelastic scattering rate and the contact transmission. The behavior of the system in these regimes is discussed together with the conditions for their experimental observability. [S0163-1829(96)02034-6]

21.0.1 I. INTRODUCTION

Electronic transport through $N - S$ and $S - S$ junctions and weak links has been the object of interest for many years.^{1,2} Much of the theoretical understanding of these systems has been obtained by analyzing simple models where a onedimensional character is assumed.^{2,3} These kinds of models have been very useful for clarifying the microscopic origin of basic phenomena like the excess current in $N - S$ and $S - S$ contacts and the subgap structure in $S - S$ and S -insulator- S (S-I-S) junctions. In particular, the one-dimensional scattering model introduced by Blonder et al.³ (hereafter referred to as BTK) has provided a simple way of analyzing the transport properties in terms of normal electron transmission and Andreev reflection probabilities.

With the recent advances in the fabrication of nanoscale devices,^{4,5} a closer comparison between the predictions of simple quasi-one-dimensional theories and experimental results is now at hand. References 4 and 5 provide two different examples of the progress made in the experimental achievement of a superconducting quantum point contact (SQPC). This possibility has provoked a renewed interest in more detailed and quantitative analysis of models involving a few conducting channels. It turns out that, in spite of its apparent simplicity, the case of a single-channel SQPC still contains nontrivial physical behavior at certain regimes. Single-mode contacts have been analyzed recently by different authors⁶⁻⁸ who have obtained new quantitative results for the case of small applied voltage where multiple Andreev reflections (MAR's) play a crucial role.

Some of these recent works^{7,8} illustrate the existence of a variety of regimes controlled by the contact transmission and quasiparticle damping in the small bias limit. A full understanding of the different regimes, together with the transitions between them remains to be explored. This will be thoroughly analyzed in this paper.

Traditionally, quantum transport in microelectronic devices has been mainly addressed by two different approaches: one is based on the scattering picture first introduced by Landauer⁹ in which the transport properties are expressed in terms of transmission and reflection scattering amplitudes. This approach is generally used in a phenomenological way by replacing the device with a simple scattering structure. The natural extension of this picture to the superconducting case was provided by the already mentioned BTK model.

A different point of view arises when the problem is analyzed starting from a microscopic Hamiltonian. We shall very generally call 'Hamiltonian approach' the theories which take this starting point. The origin of this approach can be traced back to the work by Bardeen who introduced the tunnel Hamiltonian approximation for describing a tunnel junction.¹⁰ Most of the calculations based on this tunnel Hamiltonian were restricted to the lowest order transport processes like in the calculation of the Josephson current in a $S - I - S$ junction.¹¹

Multiparticle tunneling processes (MPT's) were first discussed by Schrieffer and Wilkins¹² as a possible explanation for the observed subgap structure in superconducting tunnel junctions. The contributions of these higher order processes were found to be divergent, which has led to the quite extended belief that the Hamiltonian approach is pathological except for describing the lowest order tunneling processes.

One of the aims of the present work is to show that starting from a very simple model Hamiltonian describing a single-channel contact it is possible to obtain results for $N-S$ and $S-S$ contacts in agreement with those provided by scattering theory. As will be shown below inclusion of higher order processes up to infinite order eliminates the pathologies associated with finite order perturbation theory.

On the other hand, the Hamiltonian approach in combination with nonequilibrium Green function techniques presents

its own advantages. In recent publications we have shown how this approach can be generalized for dealing with situations where self-consistency of the superconducting order parameter is needed.¹³ Moreover, the formulation in terms of Green functions is especially well suited for dealing with correlation effects when strong $e-e$ interactions are present.¹⁴

The paper is organized as follows: In Sec. II the general model for a superconducting weak link in a local representation is introduced. In particular, we discuss how this model can be simplified in order to describe a single-mode contact. We then introduce the nonequilibrium Green function formalism by means of which the total current through the contact can be expressed. In Sec. III we analyze the more simple $N-N$ and $N-S$ contacts. The study of the $N-N$ case allows us to define the contact transmission coefficient in terms of the microscopic parameters of our model. The complete correspondence with the results of the BTK scattering theory is then established by analyzing the $N-S$ contact. For this last case we derive an analytical expression for the total current as a function of the contact transmission. Section IV is devoted to the $S-S$ case. We first show how the problem of the calculation of the ac current components can be reduced to the evaluation of an algebraic set of linear equations. This allows us to study in detail the general features of the dc and ac $I-V$ characteristics for the whole range of voltages and transmissions. We finally concentrate on the limit $V \ll \Delta$ where the more interesting and less understood physics takes place. Section V is devoted to some concluding remarks.

21.0.2 II. MODEL AND METHOD

In this paper we shall consider "point-contact-like" geometries consisting very generally of two wide (3D or 2D) electrodes connected by a narrow constriction. The constriction length, L_C , is assumed to be much smaller than the superconducting coherence length and its width, W_C , comparable to the Fermi wavelength, λ_F . For a sufficiently short constriction the detailed form of the self-consistent order parameter and electrostatic potential in the region between the electrodes becomes irrelevant¹³ allowing us to represent them by simple step functions. On the other hand, the condition $W_C \sim \lambda_F$ implies that there is only a reduced number of quantum transverse channels through the constriction. One can think of different experimental realizations of such a physical situation, like for instance the recently developed atomic size break junctions⁴ and the split-gate S-2DEG-S SQPC of Takayanagi et al.⁵ These two situations are schematically represented in Fig. 1.

The analysis of the transport and electronic properties of such a system would involve, for the most general case where both electrodes are superconductors, the solution of the time-dependent Bogoliubov-de Gennes (BdeG) equations.¹⁵ In previous works we have developed an approach to these kind of problems in which the BdeG equations are formulated in a site representation.¹³ This representation can be viewed either as a tight-binding description of

the electronic states or as a discretization of the BdeG equations. The first case would be more suitable for describing systems like an atomic size contact [Fig. 1(a)], while the second could be used to represent constrictions involving a 2DEG in semiconducting heterostructures [Fig. 1(b)]. In the

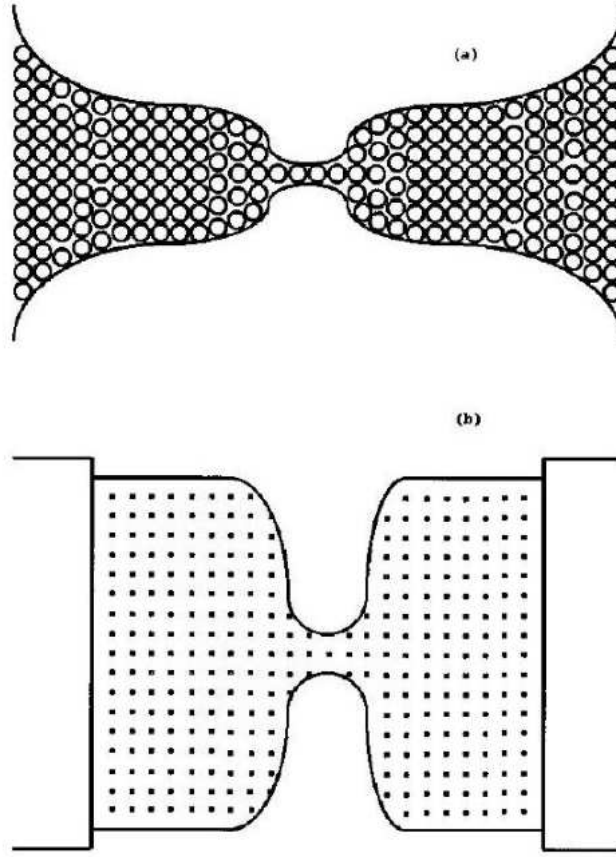


FIG. 1. Schematic representations of the two different situations discussed in Sec. II. In the absence of applied fields the mean-field Hamiltonian giving rise to the BdeG equations would take the form

$$\hat{H} = \sum_{i,\sigma} (\epsilon_i - \mu) c_{i\sigma}^\dagger c_{i\sigma} + \sum_{i \neq j, \sigma} t_{ij} c_i^\dagger c_j + \sum_i \left(\Delta_i^* c_{i\downarrow}^\dagger c_{i\uparrow}^\dagger + \Delta_i c_{i\uparrow} c_{i\downarrow} \right), \quad (1)$$

where i, j run over the sites used to represent the system. By appropriately choosing the different parameters ϵ_i, t_{ij} , and Δ_i one can model different junction geometries.¹³ In order to analyze the case of biased contacts it is convenient to simplify this model Hamiltonian taking the steplike behavior of the order parameter and electrostatic potentials in a point contact geometry into account. To do this we take the complex order parameter and the chemical potentials as constants on the left and right electrodes [denoted by (Δ_L, μ_L) and (Δ_R, μ_R) , respectively]. The case of a single quantum channel connecting both electrodes can then be described by the following Hamiltonian:

$$\hat{H} = \hat{H}_L + \hat{H}_R + \sum_{\sigma} \left(t c_{L\sigma}^\dagger c_{R\sigma} + t^* c_{R\sigma}^\dagger c_{L\sigma} \right) - \mu_L \hat{N}_L - \mu_R \hat{N}_R, \quad (2)$$

where $\hat{H}_L$ and $\hat{H}_R$ correspond to the uncoupled electrodes while the hopping term describes the electron transfer processes between the outermost sites on both electrodes. We would like to stress that, although this is a very simple model Hamiltonian (formally equivalent to a

tunnel Hamiltonian), it contains the relevant physics of a quantum point contact which depends essentially on the contact normal transmission coefficient. In our model this transmission can vary between ~ 0 (tunnel limit) and ~ 1 (ballistic regime) as a function of the coupling parameter t . This will be discussed in Sec. III.

For the case of a symmetric superconducting contact in the presence of an applied bias voltage $eV = \mu_L - \mu_R$ it is convenient to perform a gauge transformation¹⁶ by means of which the Hamiltonian (2) adopts the following timedependent form:

$$\hat{H}(\tau) = \hat{H}_L + \hat{H}_R + \sum_{\sigma} \left(t e^{i\phi(\tau)/2} c_{L\sigma}^{\dagger} c_{R\sigma} + t^* e^{-i\phi(\tau)/2} c_{R\sigma}^{\dagger} c_{L\sigma} \right),$$

where the superconducting phase difference $\phi(\tau) = \phi_0 + 2eV\tau/\hbar$ appears as a phase factor in the hopping elements.

Within the previously introduced model, the average total current through the contact is given by

$$I(\tau) = \frac{ie}{\hbar} \sum_{\sigma} \left[t \langle c_{L\sigma}^{\dagger}(\tau) c_{R\sigma}(\tau) \rangle - t^* \langle c_{R\sigma}^{\dagger}(\tau) c_{L\sigma}(\tau) \rangle \right], \quad (4)$$

where, depending on the gauge choice, t can include a timedependent phase like in Eq. (3). The averaged quantities in Eq. (4) can be expressed in terms of the nonequilibrium Keldysh Green functions $\hat{G}^{+-}$.¹⁷ For the superconducting state it is convenient to introduce the 2×2 Nambu representation in which $\hat{G}_{i,j}^{+-}$ adopts the form¹³

$$\hat{G}_{i,j}^{+-}(\tau, \tau') = i \begin{pmatrix} \langle c_{j\uparrow}^{\dagger}(\tau') c_{i\uparrow}(\tau) \rangle & \langle c_{j\downarrow}(\tau') c_{i\uparrow}(\tau) \rangle \\ \langle c_{j\uparrow}^{\dagger}(\tau') c_{i\downarrow}^{\dagger}(\tau) \rangle & \langle c_{j\downarrow}(\tau') c_{i\downarrow}^{\dagger}(\tau) \rangle \end{pmatrix}. \quad (5)$$

In terms of the $\hat{G}^{+-}$, the current is given by

$$I(\tau) = \frac{2e}{\hbar} \left[\hat{t} \hat{G}_{RL}^{+-}(\tau, \tau) - \hat{t}^* \hat{G}_{LR}^{+-}(\tau, \tau) \right]_{11} \quad (6)$$

where $\hat{t}$ is the Nambu representation of the hopping elements

$$\hat{t} = \begin{pmatrix} t & 0 \\ 0 & -t^* \end{pmatrix} \quad (7)$$

The problem then consists in the determination of these nonequilibrium Green functions $\hat{G}^{+-}$. This can be formally achieved by treating the coupling $\hat{t}$ as a perturbation. The unperturbed Green functions correspond to the uncoupled electrodes in equilibrium. For a symmetric contact, and neglecting finite bandwidth effects, the uncoupled retarded and advanced Green functions can be expressed as

$$\begin{aligned} \hat{g}_{LL}^{r,a}(\omega) &= \hat{g}_{RR}^{r,a}(\omega) \\ &= \frac{1}{W \sqrt{\Delta^2 - (\omega \pm i\eta)^2}} \begin{pmatrix} -\omega \pm i\eta & \Delta \\ \Delta & -\omega \pm i\eta \end{pmatrix}, \end{aligned} \quad (8)$$

where W is an energy scale related to the normal density of states at the Fermi level by $\rho(\epsilon_F) \sim 1/(\pi W)$ and η is a small energy relaxation rate that takes into account the damping of the quasiparticle states due to inelastic processes inside the electrodes.¹⁸ On the other hand, the unperturbed $\hat{g}^{+-}(\omega)$ satisfy the relation

$$\hat{g}^{+-}(\omega) = 2\pi i \hat{\rho}(\omega) n_F(\omega) \quad (9)$$

where $\hat{\rho}(\omega) = (1/\pi) \text{Im} [\hat{g}^a(\omega)]$ and $n_F(\omega)$ is the Fermi function.

For the coupled system, the functions $\hat{G}_{i,j}^{+-}$ can be obtained from the retarded and advanced functions by means of the integral equation

$$\hat{G}^{+-}(\tau, \tau') = \int d\tau_1 d\tau_2 \left[\hat{I} \delta(\tau - \tau_1) + \hat{G}^r(\tau, \tau_1) \hat{\Sigma}^r(\tau_1) \right] \hat{g}^{+-}(\tau_1 - \tau_2) \left[\hat{I} \delta(\tau_2 - \tau') + \hat{\Sigma}^a(\tau_2) \hat{G}^a(\tau_2, \tau') \right] \quad (10)$$

where $\hat{G}^{r,a}$ can be derived from their corresponding Dyson equations

$$\hat{G}^{r,a}(\tau, \tau') = \hat{g}^{r,a}(\tau - \tau') + \int d\tau_1 \hat{g}^{r,a}(\tau - \tau_1) \hat{\Sigma}^{r,a}(\tau_1) \hat{G}^{r,a}(\tau_1, \tau') \quad (11)$$

In the above equations the self-energies take the simple form $\hat{\Sigma}_{LL}^{r,a} = \hat{\Sigma}_{RR}^{r,a} = 0$ and $\hat{\Sigma}_{LR}^{r,a} = (\hat{\Sigma}_{RL}^{r,a})^* = \hat{t}$. In Secs. III and IV details of the solution of these integral equations are given.

21.0.3 III. $N - N$ AND $N - S$ CONTACTS

In this section we briefly review the $N - N$ and $N - S$ cases. The analysis of the $N - N$ case permits one to define the normal transmission coefficient of the contact in terms of the microscopic parameters of the model, which is necessary for making contact with the scattering approach. The equivalence of both approaches is illustrated by comparing our results for the $N - S$ contact with those of the BTK model.³

As we shall deal with a stationary situation it is convenient to adopt the time-independent formulation based on the Hamiltonian (2). In this case the Green functions depend on the difference of its temporal arguments and integral equations (10) and (11) become simple algebraic equations when Fourier transformed. Then we have

$$\hat{G}^{r,a}(\omega) = \hat{g}^{r,a}(\omega) + \hat{g}^{r,a}(\omega) \hat{\Sigma}^{r,a}(\omega) \hat{G}^{r,a}(\omega), \quad (12)$$

$$\hat{G}^{(+--),(-++)}(\omega) = [\hat{I} + \hat{G}^r(\omega) \hat{\Sigma}^r(\omega)] \hat{g}^{(+--),(-++)}(\omega) [\hat{I} + \hat{\Sigma}^a(\omega) \hat{G}^a(\omega)], \quad (13)$$

where $\hat{g}^{-+}(\omega) = -2\pi i \hat{\rho}(\omega) [1 - n_F(\omega)]$. On the other hand, it is convenient to rewrite expression (6) for the current as^{13,19}

$$I = \frac{2e}{h} t^2 \int_{-\infty}^{\infty} [g_{LL,11}^{+-} G_{RR,11}^{-+}(\omega) - g_{LL,11}^{-+} G_{RR,11}^{+-}(\omega)] d\omega \quad (14)$$

where we have assumed that the left electrode is in the normal state, while the right one can be either normal or superconducting.

Let us first analyze the case where both electrodes are in the normal state. Details on the calculations of the Green functions for this case have been given elsewhere.²⁰ As in the scattering approach, the current can be written as

$$I = \frac{2e}{h} \int_{-\infty}^{\infty} T(\omega, V) [n_F(\omega - eV) - n_F(\omega)] d\omega, \quad (15)$$

where $T(\omega, V)$ is an energy-dependent transmission coefficient which is given by²¹

$$T(\omega, V) = \frac{4\pi^2 t^2 \rho_{LL}(\omega - eV) \rho_{RR}(\omega)}{|1 - t^2 g_{LL}(\omega - eV) g_{RR}(\omega)|^2} \quad (16)$$

Equation (16) adopts a simpler form in the small bias limit where the normal system exhibits an Ohmic response. This regime is valid for a bias voltage range much smaller than the typical

energy scale for the variation of the normal density of states around the Fermi energy. This condition certainly holds for the case we are interested in in this paper ($eV \sim \Delta$) and is consistent with the assumptions leading to Eq. (8). In this limit the coefficient $T(\omega, V)$ becomes a constant α , given by

$$T(\omega, V) \simeq \frac{4t^2/W^2}{(1 + t^2/W^2)^2} \equiv \alpha \quad (17)$$

The normal linear conductance of this single-mode contact is then given by the Landauer formula $G_{NN} = (2e^2/h) \alpha$. Notice that α can vary between zero and one as a function of t . The $\alpha \rightarrow 0$ limit is reached both for $t/W \ll 1$ and for $t/W \gg 1$, while the ballistic limit, i.e., $\alpha \sim 1$, is reached when $t/W \sim 1$.

In Ref. 21 a detailed discussion on the derivation of Eq. (17) by a different perturbative formalism and its range of validity can be found. As stated in this reference, finite order perturbation in the coupling t/W is only valid for $t/W < 1$, while for $t/W > 1$ finite order perturbation theory breaks down together with the appearance of bound states outside the metallic band. This fact could lead to the conclusion that Eq. (17) is not valid for $t/W > 1$. However, being the result of the summation of the complete infinite series, Eq. (17) is actually a nonperturbative result valid even for the somewhat academic case $t/W > 1$. In fact, by comparing this expression with the normal transmission coefficient of the δ -like scattering potential model used by BTK,³ the equivalence of both models arises by identifying $Z = [1 - (t/W)^2]/(2t/W)$, where Z is the parameter controlling the barrier strength. Thus, one can associate the case $t/W < 1$ with a repulsive barrier ($Z > 0$), and the case $t/W > 1$ with an attractive δ -like potential ($Z < 0$), which certainly would lead to a transmission lower than 1 and the presence of bound states.

In the $N-S$ case ($\Delta_L = 0, \Delta_R = \Delta$) starting from Eqs. (12) and (13) after some simple algebra we obtain the following expression for the current as the sum of four different contributions $I = I_1 + I_2 + I_3 + I_A$, where

$$\begin{aligned} I_1 &= \frac{8e}{h} \pi^2 t^2 \int_{-\infty}^{\infty} d\omega |1 + tG_{RL,11}^r(\omega)|^2 \rho_{LL,11}(\omega - eV) \rho_{RR,11}(\omega) [n_F(\omega - eV) - n_F(\omega)] \\ I_2 &= -\frac{16e}{h} \pi^2 t^2 \int_{-\infty}^{\infty} d\omega \text{Re} \{ tG_{LR,21}^a(\omega) [1 + tG_{RL,11}^r(\omega)] \} \rho_{LL,11}(\omega - eV) \rho_{RR,12}(\omega) [n_F(\omega - eV) - n_F(\omega)] \\ I_3 &= \frac{8e}{h} \pi^2 t^4 \int_{-\infty}^{\infty} d\omega |G_{RL,12}(\omega)|^2 \rho_{LL,11}(\omega - eV) \rho_{RR,22}(\omega) [n_F(\omega - eV) - n_F(\omega)] \\ I_A &= \frac{8e}{h} \pi^2 t^4 \int_{-\infty}^{\infty} d\omega |G_{RR,12}(\omega)|^2 \rho_{LL,11}(\omega - eV) \rho_{LL,22}(\omega + eV) [n_F(\omega - eV) - n_F(\omega + eV)] \end{aligned} \quad (18)$$

Written in this form, each contribution has a clear interpretation in terms of elementary processes that can be identified by inspection of the intervening spectral densities. Although there is a formal resemblance of the above expressions with those of tunnel theory²² Eq. (18) contains all possible processes up to infinite order in t . Thus, I_1 corresponds to normal electron transfer between the electrodes, while I_2 corresponds also to a net transfer of a single electron with creation or annihilation of pairs as an intermediate state. On the other hand, I_3 arises from processes where an electron in the normal electrode is converted into a hole in the superconducting side, i.e., processes with "branch crossing" in the BTK language. Finally, I_A arises from Andreev reflection processes in which an electron [with an associated spectral weight $\rho_{LL,11}(\omega - eV)$] is transmitted from the left to the right electrode with a hole reflecting backwards into the normal electrode [with an associated

spectral weight $\rho_{LL,22}(\omega + eV)$] while a Cooper pair is created in the superconducting side with a probability proportional to $|G_{RR,12}(\omega)|^2$.

As expected, the only nonzero contribution for $eV < \Delta$ is I_A , while all processes contribute for $eV > \Delta$. With the same simplifying assumptions leading to Eq. (17) the differential conductance at zero temperature adopts the simple form

$$\begin{aligned} G_{NS}(V) &= \frac{4e^2}{h} \frac{\alpha^2}{(2-\alpha)^2 - 4(1-\alpha)(eV/\Delta)^2}, \quad eV \leq \Delta, \\ G_{NS}(V) &= \frac{4e^2}{h} \frac{\alpha}{\alpha + (2-\alpha)\sqrt{1-(\Delta/eV)^2}}, \quad eV > \Delta. \end{aligned} \quad (19)$$

This expression can be shown to be equivalent to the one obtained from the BTK model³ with the correspondence $Z = [1 - (t/W)^2] / (2t/W)$ commented on above.

It is worth noticing that the differential conductance rises from $(4e^2/h) \alpha^2 / (2-\alpha)^2$ at $V = 0$ to the value $4e^2/h$ at $eV = \Delta$, this last value being independent of the contact transmission. This result can never be obtained within any finite order perturbative approximation in t , but requires an infinite order calculation.²³ These sorts of nonperturbative features are also very important in the $S-S$ case,²⁴ as will be discussed in the next section.

Finally, it is possible to obtain from Eq. (18) an analytical expression of the total current at zero temperature. For $eV > \Delta$ this can be written as $I = I_1 + I_2$ with

$$\begin{aligned} I_1 &= \frac{e\Delta}{h} \frac{\alpha^2}{(2-\alpha)\sqrt{1-\alpha}} \ln \left[\frac{1 + [2\sqrt{1-\alpha}/(2-\alpha)]}{1 - [2\sqrt{1-\alpha}/(2-\alpha)]} \right] \\ I_2(x) &= \frac{4e}{h} \Delta \left[\frac{\alpha^2}{4(1-\alpha)} + \frac{\alpha}{x[\alpha + (2-\alpha)\sqrt{1-x^2}]} - \frac{\alpha(2-\alpha)^2 x}{4(1-\alpha)[\alpha + (2-\alpha)\sqrt{1-x^2}]} \right. \\ &\quad \left. + \frac{\alpha^2(2-\alpha)}{8(1-\alpha)^{3/2}} \ln \left(\frac{(x/(1+\sqrt{1-x^2}) + 1/\sqrt{1-\alpha})(1-1/\sqrt{1-\alpha})}{(x/(1+\sqrt{1-x^2}) - 1/\sqrt{1-\alpha})(1+1/\sqrt{1-\alpha})} \right) \right] \end{aligned} \quad (20)$$

where I_1 is the gap contribution to the total current, I_2 is the contribution coming from energies outside the gap and $x = \Delta/eV$. This allows one to analyze with detail the 'excess' current, defined as $I_{\text{exc}} = \lim_{V \rightarrow \infty} (I_{NS} - I_{NN})$, as a function of the contact transparency. We find $I_{\text{exc}} = I_{\text{exc}_1} + I_{\text{exc}_2}$

$$\begin{aligned} I_{\text{exc}_1} &= \frac{e\Delta}{h} \frac{\alpha^2}{(2-\alpha)\sqrt{1-\alpha}} \ln \left[\frac{1 + [2\sqrt{1-\alpha}/(2-\alpha)]}{1 - [2\sqrt{1-\alpha}/(2-\alpha)]} \right] \\ I_{\text{exc}_2} &= \frac{e\Delta}{h} \alpha^2 \left[\frac{1}{1-\alpha} + \frac{2-\alpha}{2(1-\alpha)^{3/2}} \ln \left(\frac{1 - \sqrt{1-\alpha}}{1 + \sqrt{1-\alpha}} \right) \right] \end{aligned}$$

where I_{exc_1} and I_{exc_2} are, respectively, the contributions coming from energies inside and outside the gap. As can be easily checked, $I_{\text{exc}_1} > 0$ while $I_{\text{exc}_2} < 0$ the total excess current being always positive. For $\alpha = 1$ the well-known result^{25,29} for the ballistic contact $I_{\text{exc}} = (8/3)e\Delta/h$ is recovered.

21.0.4 IV. THE S-S CONTACT

21.0.5 A. An efficient algorithm for evaluating the ac current

As commented on in Sec. II, for the case of a voltagebiased $S-S$ contact it is convenient to start from the Hamiltonian (3) in which the applied bias is taken into account through a

time-dependent phase factor in the hopping element, which in a Nambu representation has the form

$$\hat{t} = \begin{pmatrix} te^{i\phi(\tau)/2} & 0 \\ 0 & -te^{-i\phi(\tau)/2} \end{pmatrix} \quad (21)$$

where $\phi = \phi_0 + 2eV\tau/\hbar$ is the time-dependent superconducting phase difference.

This explicit time dependence indicates that all dynamic quantities can be expanded as a Fourier series in all possible harmonics of the fundamental frequency $\omega_0 = 2eV/\hbar$.⁷ For instance, the total current can be written as

$$I(\tau) = \sum_m I_m e^{im\omega_0\tau} \quad (22)$$

We shall now show how these Fourier coefficients, I_m , can be efficiently evaluated within the nonequilibrium Green function formalism. Let us first notice that the nonequilibrium Green functions appearing in Eqs. (10) and (11) do not depend only on the difference of their temporal arguments, and have therefore a generalized Fourier expansion of the form^{26,27}

$$\hat{G}(\tau, \tau') = \frac{1}{2\pi} \sum_n \int d\omega e^{-i\omega\tau} e^{i(\omega+n\omega_0/2)\tau'} \hat{G}(\omega, \omega + n\omega_0/2).$$

Hereafter we shall use the notation $\hat{G}_{nm}(\omega) = \hat{G}(\omega + n\omega_0/2, \omega + m\omega_0/2)$. Different Fourier components $\hat{G}_{nm}$ are related by $\hat{G}_{nm}(\omega) = \hat{G}_{n-m,0}(\omega + m\omega_0/2)$. For the particular gauge choice adopted here, it is useful to express all quantities in terms of a renormalized hopping which satisfies its own Dyson equation

$$\begin{aligned} \hat{T}^{a,r}(\tau, \tau') &= \hat{t}(\tau) \delta(\tau - \tau') + \int d\tau_1 d\tau_2 \hat{t}(\tau) \hat{g}^{a,r}(\tau - \tau_1) \hat{t}^\dagger(\tau_1) \\ &\quad \times \hat{g}^{a,r}(\tau_1 - \tau_2) \hat{T}^{a,r}(\tau_2, \tau') \end{aligned} \quad (24)$$

This quantity can be viewed as the total hopping amplitude arising from summing up all processes in which one electron is transferred. Clearly, it is formally equivalent to use a renormalized hopping instead of renormalized propagators as they are linked by relations like $\hat{G}_{LL}(\tau, \tau') \hat{t}(\tau') = \int d\tau_1 \hat{g}_L(\tau - \tau_1) \hat{T}_{LR}(\tau_1, \tau')$. The current components can now be expressed in terms of the renormalized hopping elements $\hat{T}_{nm}^{a,r}(\omega)$ as

$$\begin{aligned} I_m &= \frac{2e}{h} \int d\omega \sum_n \left[\hat{T}_{0n}^r \hat{g}_{nn}^{+-} \hat{T}_{nm}^{r\dagger} \hat{g}_{mm}^a - \hat{g}_{00}^r \hat{T}_{0n}^r \hat{g}_{nn}^{+-} \hat{T}_{nm}^{r\dagger} \right. \\ &\quad \left. + \hat{g}_{00}^r \hat{T}_{0n}^{a\dagger} \hat{g}_{nn}^{+-} \hat{T}_{nm}^a - \hat{T}_{0n}^{a\dagger} \hat{g}_{nn}^{+-} \hat{T}_{nm}^a \hat{g}_{mm}^a \right]_{11}, \end{aligned} \quad (25)$$

where we have eliminated the site indexes L and R in the uncoupled Green functions due to the left-right symmetry of the contact.

The problem is then reduced to that of the evaluation of the components $\hat{T}_{nm}$. From Eq. (24) it can be shown that the components $\hat{T}_{nm}$ (both retarded and advanced parts) satisfy a set of linear equations of the form

$$\hat{T}_{nm} = \hat{t}_{nm} + \hat{\epsilon}_n \hat{T}_{nm} + \hat{V}_{n,n-2} \hat{T}_{n-2,m} + \hat{V}_{n,n+2} \hat{T}_{n+2,m} \quad (26)$$

These equations are mathematically equivalent to those describing the motion of electrons in a tight-binding linear chain with 'site energies,' $\hat{\epsilon}_n$, and 'nearest-neighbor couplings,' $\hat{V}_{n,n-2}$ and $\hat{V}_{n,n+2}$. The detailed expression of $\hat{\epsilon}_n$ and $\hat{V}_{n,m}$ in terms of the unperturbed Green functions are given in Appendix A. This analogy allows us to obtain the Fourier coefficients $\hat{T}_{nm}$ using

standard recursion techniques. One can show (see Appendix A) that the following recursive relation holds:

$$\begin{aligned}\hat{T}_{n+2,m}(\omega) &= \hat{z}^+ [\omega + (n-1)\omega_0] \hat{T}_{nm}(\omega), \quad n \geq 1, \\ \hat{T}_{n-2,m}(\omega) &= \hat{z}^- [\omega + (n+1)\omega_0] \hat{T}_{nm}(\omega), \quad n \leq -1,\end{aligned}\quad (27)$$

where the transfer matrix $\hat{z}^\pm$ satisfy the equation

$$\hat{z}^\pm(\omega) = [\hat{I} - \hat{\epsilon}_{\pm 3} - \hat{V}_{\pm 3, \pm 5} \hat{z}^\pm(\omega \pm \omega_0)]^{-1} \quad (28)$$

Clearly, as the transfer matrix $\hat{z}^\pm$ connects consecutive harmonics of $\hat{T}$, it can be viewed as a generating function which introduces the effect of a unitary Andreev reflection process. The problem has been reduced to the calculation of only two matrix coefficients like, for instance, $\hat{T}_{1,0}$ and $\hat{T}_{-1,0}$ as a starting point for the generating equations (27) (see Appendix A for details).

In summary, the basic mathematical difficulty lies in the evaluation of the transfer matrix functions $\hat{z}^\pm$ from Eq. (28). Although Eq. (28) looks simple, it is nevertheless hard to solve analytically for arbitrary values of V . The analytical results presented in this paper are limited to the $eV/\Delta \rightarrow 0$

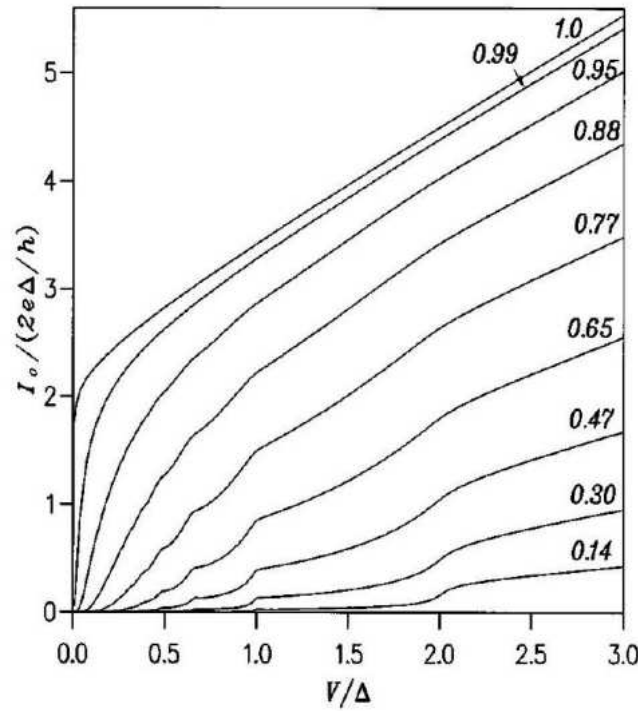


FIG. 2. The dc current-voltage characteristic of a $S - S$ contact for different values of the normal transmission at zero temperature.

and $eV/\Delta \rightarrow \infty$ cases where some simplifying relations hold. For intermediate voltages, an accurate numerical solution of Eq. (28) can be obtained.

21.0.6 B. Analysis of ac and dc $I - V$ characteristics

In this subsection we analyze the general features of the $I - V$ characteristics of an $S - S$ contact obtained using our formalism. Let us start by briefly discussing the dc current, $I_0(V)$, for different values of the transmission, as shown in Fig. 2. Although the overall qualitative features of these curves have been known since the works of Octavio et al.,²⁸ Zaitsev,²⁵ and Arnold,²⁶ more quantitative and detailed analysis are being reported in recent publications.

^{6,8} The results of Fig. 2 are in agreement with those reported recently by Averin and Bardas ⁸ which were obtained using the scattering approach.

Two relevant features of these curves are the subharmonic gap structure for $eV < 2\Delta$ and the excess current for $eV \gg \Delta$. As can be observed, the subgap structure becomes progressively more pronounced with decreasing transmission. Eventually, when $\alpha \ll 1$ the current steps at positions $eV \sim 2\Delta/n$ can be clearly resolved. In this limit one can isolate the elementary processes which give rise to these steps. It can be shown that in the tunnel limit the n th step can be calculated as

$$\begin{aligned} \delta I_0^{(n)} &= \lim_{eV \rightarrow 2\Delta^+/n} \frac{8e}{\hbar} \pi^2 n t^{2n} \\ &\times \int_{\Delta}^{neV-\Delta} d\omega \left[\prod_{i=1}^{n-1} |g_{12}(\omega - ieV)|^2 \right] \\ &\times \rho_{11}(\omega) \rho_{22}(\omega - neV). \end{aligned} \quad (29)$$

By comparison of Eq. (29) with the expression of I_A in Eq. (18) it becomes clear that the steps inside the gap are due to the opening of a new Andreev reflection channel whenever $eV = 2\Delta/n$. Calculation of the integrals in Eq. (29) leads to

$$\delta I_0^{(n)} = \frac{e\Delta\alpha^n}{\hbar} \left(\frac{2n}{4^{2n-1}} \right) \left(\frac{n^n}{n!} \right)^2 \quad (30)$$

in agreement with the recent prediction of Bratus et al. ⁶ On the other hand, when $\alpha \rightarrow 1$ the subgap structure is completely washed out and there appears an excess current even in the small bias limit.

The other relevant feature of the dc current, namely the large voltage excess current, can be analytically evaluated within our model for any transmission value. The main simplification in this limit comes from the fact that only the lowest order Andreev reflection process gives a significant contribution to the excess current. ²⁸ This implies that one can truncate the system of equations (26) for harmonic indexes $n > 1$, the resulting simplified system can then be solved explicitly for $\hat{T}_{1,0}(\omega)$ and $\hat{T}_{-1,0}(\omega)$ (for details see Appendix B). As shown in this Appendix the simple result $I_{\text{exc}}^{SS} = 2I_{\text{exc}}^{NS}$ is obtained for any value of the transmission. Although this is the expected physically sound result, to our knowledge, it has not been shown explicitly before except for the ballistic case. ^{25,29} Moreover, some authors have reported the existence of a negative excess current for low transmissions. ³⁰ which seems to be in contradiction with the above result. However, one should notice that the excess current as defined above is an asymptotic quantity (only valid in the $eV/\Delta \rightarrow \infty$ limit). When corrections of order Δ/eV are taken into account one actually can have a defect instead of an excess current for sufficiently low transmission.

The algorithm described in Sec. IV A allows an efficient evaluation of the higher order ac components of the current. For the following analysis we decompose the ac current, Eq. (22), into its dissipative and nondissipative contributions given, respectively, by

$$I_D = I_0 + \sum_m I_m^D \cos(m\omega_0\tau) \quad (31)$$

and

$$I_S = \sum_m I_m^S \sin(m\omega_0\tau) \quad (32)$$

where $I_m^D = 2 \operatorname{Re}(I_m)$ and $I_m^S = -2 \operatorname{Im}(I_m)$.

The results obtained for the first three I_m^D and I_m^S components are depicted in Fig. 3 and Fig. 4. As can be observed, these components become exponentially small for bias voltages larger than Δ/n . On the contrary, when $eV < \Delta/n$ the decay of the ac components with increasing n becomes slower. The analysis of the higher order components reveals a decay for $eV < \Delta/n$ close to an inverse power law. As a consequence of this slow decay, one is forced to take an increasing number of ac components into account in order to adequately describe the behavior at small bias. This will be the subject of the next subsection.

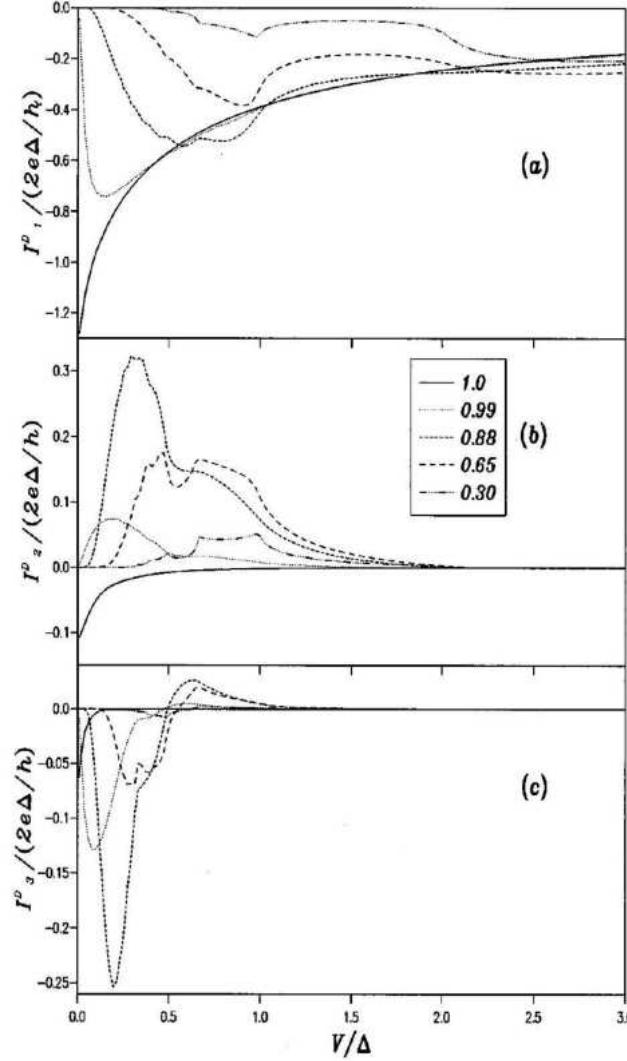


FIG. 3. The first three ac components of the dissipative current for different values of the normal transmission at zero temperature.

21.0.7 C. Small bias regime

In this subsection we concentrate on the $eV/\Delta \rightarrow 0$ case, which turns out to exhibit a remarkable variety of different regimes according to the values of the parameters $\alpha\Delta$ and the inelastic scattering rate η . As it is well known, the main difficulty for obtaining quantitative results in this limit lies in the fact that the number of MAR's contributing to the current grows with decreasing V as $\sim \Delta/eV$.²⁶ Furthermore, the amplitudes of these multiple processes do not decay when $(V, \eta) \rightarrow 0$ leading to the appearance of divergencies in the perturbative expansion in the coupling $\hat{t}$.⁷ Again, a complete summation of the perturbative series is needed in order to regularize these divergencies. An additional difficulty arises, as will be discussed below, from the fact that the limits $V \rightarrow 0$ and $\eta \rightarrow 0$ do not actually commute.

When decreasing the bias voltage two main situations can be reached depending on the strength of the inelastic scattering rate at the superconducting electrodes: the case of $eV/\Delta \rightarrow 0$ and finite η has been discussed by the present authors in recent publications;⁷ on the other hand, the case of small V/Δ and negligible η has been recently addressed to

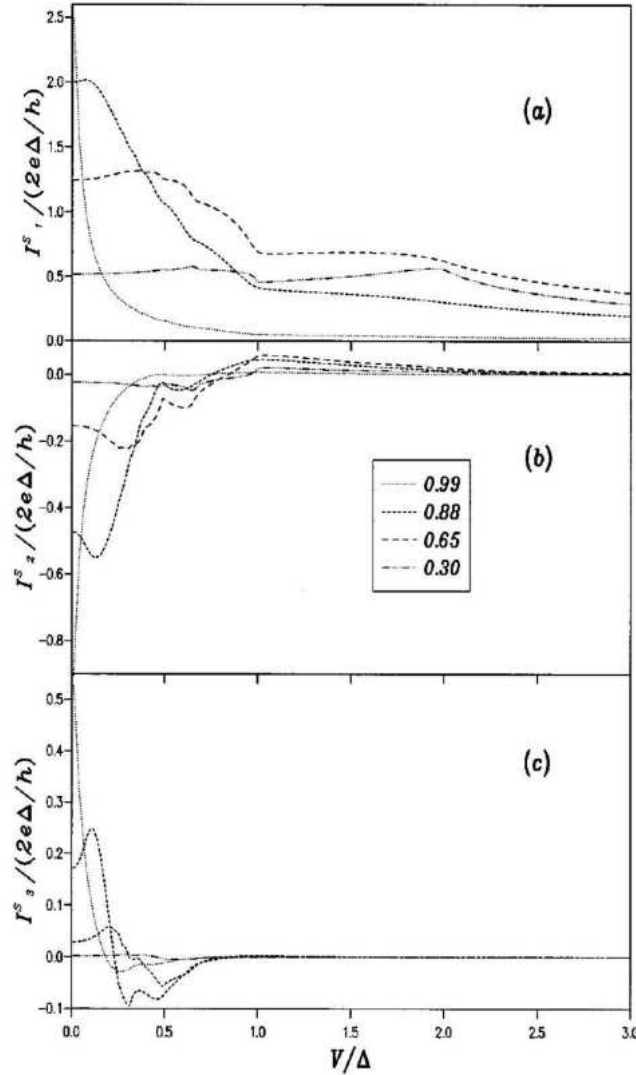


FIG. 4. Same as Fig. 4 for the nondissipative current. by Averin and Bardas.⁸ In what follows we summarize the main results for both regimes and analyze the conditions for their actual observability in a real SQPC.

Within our formalism, analytical results in the small bias limit become feasible since the transfer matrix $\hat{z}^\pm(\omega)$ tends to a scalar quantity having the form of a simple phase factor inside the gap region. As shown in Appendix C for $(\eta, V) \rightarrow 0$ we find

$$\hat{z}^\pm(\omega) = z(\omega) = e^{i\varphi(\omega)}, \quad \Delta\sqrt{1-\alpha} \leq |\omega| \leq \Delta \quad (33)$$

where

$$\varphi(\omega) = \arcsin \left(\frac{2}{\alpha\Delta^2} \sqrt{\Delta^2 - \omega^2} \sqrt{\omega^2 - (1-\alpha)\Delta^2} \right)$$

As the multiple Andreev processes are generated by successive applications of $\hat{z}(\omega)$, Eq. (33) indicates that these processes do not decay in this limit inside the gap region. This infinite series of MAR gives rise to the well-known bound state spectrum of a current-carrying SQPC at zero bias voltage.³¹ As shown in Appendix C, the positions of these bound states are determined by the condition $\varphi(\omega) = \phi$. The

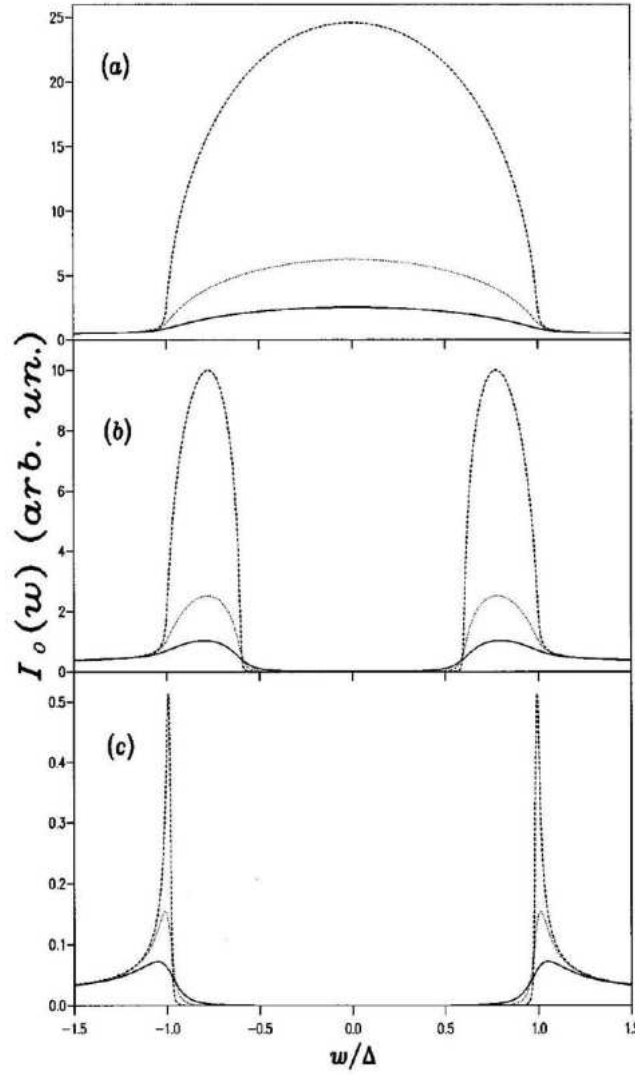


FIG. 5. Current density for the dc component I_0 within the linear regime discussed in Sec. IV C. (a), (b), and (c) correspond to transmission values $\alpha = 1, 0.65$, and 0.04 , respectively. In all cases the full line corresponds to $\eta/\Delta = 1/10$, the dotted line to $\eta/\Delta = 1/25$, and the broken line to $\eta/\Delta = 1/100$. The thermal factor $\text{sech}^2(\beta\omega/2)$ [see Eq. (C8)] has been extracted from the current density.

presence of a small but finite η or V (whichever is larger) introduces an effective damping into the otherwise infinite series of MAR.

When this effective damping is due to a finite η ($eV \ll \eta$) a linear regime can be defined where the total current is given by $I_S(\phi) + G(\phi)V$, $G(\phi)$ being a phasedependent linear conductance.^{25,7} Within this linear regime, one can identify two different subregimes according to the ratio $\eta/\alpha\Delta$. The case $\eta/\alpha\Delta \ll 1$ corresponds to a situation where MAR's are very weakly damped and give the dominant contribution to the current. The physical picture one can have in mind is that of electrons and holes Andreevreflecting between the electrodes for a very long time before being inelastically scattered.³² In order to illustrate the dominant contribution of the processes inside the gap for the weakly damped case, we represent in Fig. 5 the current den-

sity corresponding to the I_0 component for three values of the transmission. Three important features of this linear regime are displayed in this figure: first, the current density inside the gap increases as $\sim 1/\eta$ therefore giving the dominant contribution in the weakly damped case; second, there is a region inside the gap of width $2\Delta\sqrt{1-\alpha}$ in which the current density vanishes. This is the forbidden energy region for bound states at a given transmission. Finally, in Fig. 5(c) one can observe that the contribution of the continuum outside the gap becomes important

as $\alpha < \eta/\Delta$. In the limit $\eta/\alpha\Delta \gg 1$ a second subregime is reached where the contributions of MAR's are heavily damped and the current is dominated by single quasiparticle tunneling processes. The transition between these two subregimes has been analyzed with detail in Refs. 7.

In order to identify the actual subregime for a real SQPC an estimation of the order of magnitude of η is needed. In Ref. 18 η is estimated from the electron-phonon interaction to be a small fraction of the gap for traditional superconductors. Thus, our theory predicts that a real SQPC would generally fall into the weakly damped case except for extremely low transmissions.

For this subregime the supercurrent I_S and the linear conductance $G(\phi)$ can be obtained analytically as discussed in Appendix C. In particular, for $G(\phi)$ one obtains

$$G(\phi) = \frac{2e^2}{h} \frac{\pi}{16\eta} \left[\frac{\Delta \alpha \sin \phi}{\sqrt{1 - \alpha \sin^2(\phi/2)}} \operatorname{sech} \left(\frac{\beta \omega_S}{2} \right) \right]^2 \beta \quad (34)$$

where ω_S is the position of the bound states inside the gap and $\beta = 1/k_B T$. This expression for large transmission and small temperatures gives a phase-dependence which is in qualitative agreement with the few available experimental results, performed in nonmesoscopic contacts.³³ The unusual phase dependence of Eq. (34) which deviates strongly from the $\cos(\phi)$ form predicted by the standard tunnel theory may explain the old controversy between tunnel theory and experiments known as the $\cos(\phi)$ problem.^{34,35}

On the other hand, when the truncation of the infinite series of MAR's is caused by a finite V (with negligible η), analytical results have only been obtained in the quasiballistic limit, i.e., $\alpha \rightarrow 1$. A closer inspection of the $I - V$ curves in the small bias region and for $\alpha \sim 1$ reveals that the supercurrent components decay exponentially from its value at $V = 0$ with a collapsing width $\sim (1 - \alpha)\Delta$. This is illustrated in Fig. 6(a) where a blow up of the behavior of I_1^S for small bias voltages and quasiballistic transmissions is shown. In the limit $\alpha \rightarrow 1$ the supercurrent becomes a δ function at $V = 0$. On the contrary, the dissipative components in this same limit tend to a finite value outside the region of width $\sim (1 - \alpha)\Delta$. This behavior is shown in Fig. 6(b) where I_1^D is plotted in the same magnified scale as I_1^S . The summation of these dissipative components for $\alpha = 1$ and very small V yields (see Appendix C)

$$I_D(\phi) = \frac{e\Delta}{h} |\sin(\phi/2)| \operatorname{sgn} V \quad (35)$$

in agreement with the result recently derived by Averin and Bardas.⁸ The existence of a region of decreasing width

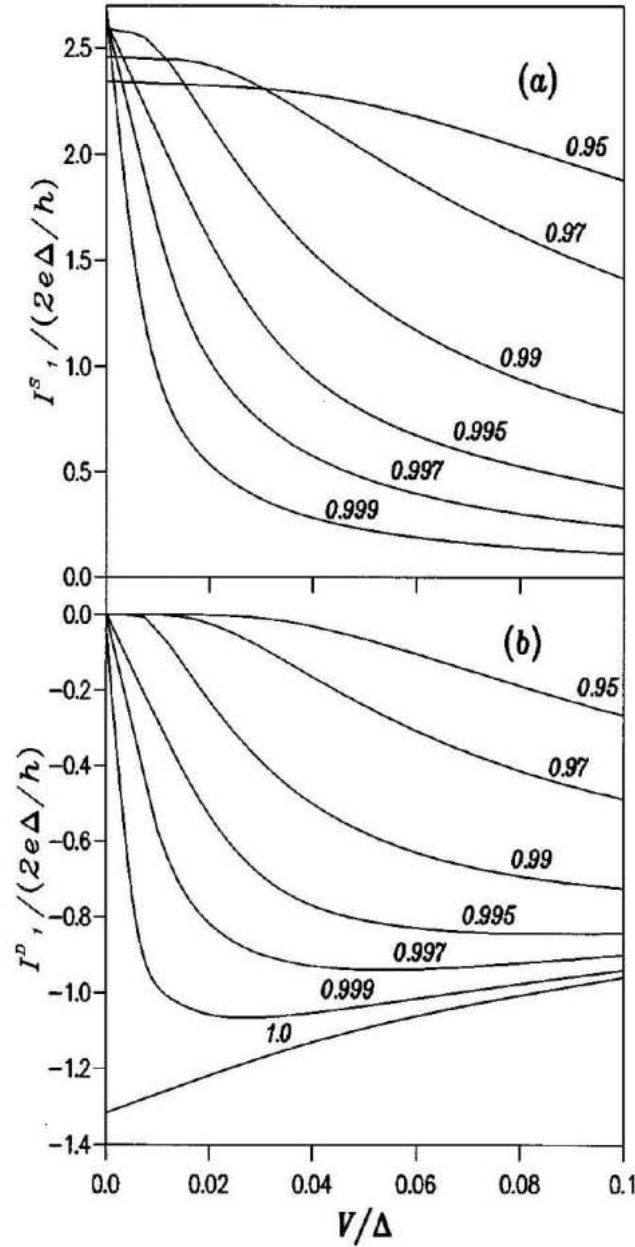


FIG. 6. Behavior of the first nondissipative (a) and dissipative (b) ac components in the very small voltage range close to ballistic conditions. These results have been obtained for negligible η .

$V \sim (1 - \alpha)\Delta$ in which this crossover from supercurrent to dissipative current takes place can be associated with the collapse of the forbidden region for MAR's inside the superconducting gap taking place when $\alpha \rightarrow 1$. In this way, when V is small compared to the width of the forbidden region the excitation of quasiparticles from states at $\omega < -\Delta\sqrt{1 - \alpha}$ into states at $\omega > \Delta\sqrt{1 - \alpha}$ is negligible and there is no appreciable dissipative current; whereas the opposite situation holds for $V > \sqrt{1 - \alpha}\Delta$. Averin and Bardas⁸ have described this crossover as a Landau-Zener transition in which the nondissipative and the dissipative components scale with α and V as $(1 - p)$ and p , respectively, where $p = \exp[-\pi(1 - \alpha)\Delta/eV]$. The numerical results for sufficiently small eV/Δ and $(1 - \alpha)$ are well fitted by these scaling laws. However, a careful analysis reveals that their range of validity around $V = 0$ and $\alpha = 1$ decreases strongly when increasing the component number.

In summary, in this small bias limit one can identify four different subregimes depending on the relative values of parameters η , $\alpha\Delta$, and eV . The prediction of the actual behavior of a real SQPC in this limit would therefore require a careful estimation of all these parameters. In

this respect, one

should keep in mind that while eV and α can be varied experimentally in a rather controlled way, the inelastic scattering rate η is an intrinsic property of the superconducting electrodes much more difficult to control. The unavoidable presence of some degree of inelastic scattering can prevent the actual observability of the crossover from nondissipative to dissipative behavior described after Eq. (35). The requirement of $eV \sim (1 - \alpha)\Delta$ together with that of $\alpha \sim 1$ can actually imply $eV < \eta$ which would rather correspond to the linear regime. Finally, when considering the experimental test of all these theoretical predictions the relevance of noise in a real SQPC should be taken into account. Recent theoretical predictions^{36,37} suggest that the magnitude of thermal noise in a single-mode superconducting device can be extremely large near ballistic conditions.

21.0.8 V. CONCLUDING REMARKS

In the present work we have presented a Hamiltonian approach for describing the transport properties of single-mode $N - S$ and $S - S$ contacts.

It has been explicitly demonstrated that this approach is, with some simplifying assumptions, equivalent to the phenomenological scattering approach. We believe that the present work can help clarify the somewhat recurrent discussion about the unsuitability of a Hamiltonian approach for obtaining the transport properties of a $N - S$ or $S - S$ contact: when performing the calculations up to infinite order in the coupling t all the unphysical divergences are eliminated and the results become equivalent to those of scattering theory.

On the other hand, the present approach has been applied to discuss in detail the dc and ac $I - V$ characteristics of a SQPC. In particular, we have concentrated on the less understood small bias voltage limit where one can identify four different subregimes depending on the values of the contact transmission and inelastic scattering rate. Finally, we have discussed the conditions for the experimental observability of the theoretical predictions in this limit.

Although in the present work we have restricted the discussion to the simplest single-mode case with an energyindependent transmission coefficient, the general model introduced in Sec. II can describe more complex situations which may be relevant for a closer comparison with recent experiments.⁵ Work along this line is under progress.

21.0.9 APPENDIX A

In this appendix we describe the algorithm for evaluating the ac current components in a $S - S$ biased contact. As has been pointed out in Sec. IV, we can express the current in terms of the retarded and advanced Fourier components of the renormalized hopping $\hat{T}_{nm}(\omega)$ satisfying Eq. (26)

$$\hat{T}_{nm} = \hat{t}_{nm} + \hat{\epsilon}_n \hat{T}_{nm} + \hat{V}_{n,n-2} \hat{T}_{n-2,m} + \hat{V}_{n,n+2} \hat{T}_{n+2,m} \quad (\text{A1})$$

where the matrix coefficients $\hat{\epsilon}_n$ and $\hat{V}_{nm}$ can be expressed in terms of the uncoupled Green functions as

$$\hat{\epsilon}_n = t^2 \begin{pmatrix} g_{n+1}g_n & g_{n+1}f_n \\ g_{n-1}f_n & g_{n-1}g_n \end{pmatrix}, \quad (\text{A2})$$

$$\hat{V}_{n,n+2} = -t^2 f_{n+1} \begin{pmatrix} f_{n+2} & g_{n+2} \\ 0 & 0 \end{pmatrix}, \quad (\text{A3})$$

$$\hat{V}_{n,n-2} = -t^2 f_{n-1} \begin{pmatrix} 0 & 0 \\ g_{n-2} & f_{n-2} \end{pmatrix}, \quad (\text{A4})$$

where $f(\omega) \equiv g_{12}(\omega) = g_{21}(\omega)$ and $g(\omega) \equiv g_{11}(\omega) = g_{22}(\omega)$ due to electron-hole symmetry. In the above equations the shorthand notation $g_n = g(\omega + n\omega_0/2)$ is used. Moreover, the site indexes in the Green functions have been omitted since we are considering a symmetric contact.

As commented on in Sec. IV, the linear equations (A1) are analogous to those describing a tight-binding chain with nearest-neighbor hopping parameters $\hat{V}_{n,n+2}$ and $\hat{V}_{n,n-2}$. A solution can then be obtained by standard recursive techniques. It is straightforward to show that the following recursive relations between the coefficients $\hat{T}_{nm}$ hold:

$$\begin{aligned}\hat{T}_{n+2,m}(\omega) &= \hat{z}^+ [\omega + (n-1)\omega_0] \hat{T}_{nm}(\omega), \quad n \geq 1, \\ \hat{T}_{n-2,m}(\omega) &= \hat{z}^- [\omega + (n+1)\omega_0] \hat{T}_{nm}(\omega), \quad n \leq -1,\end{aligned}\tag{A5}$$

where the transfer matrix $\hat{z}^\pm(\omega)$ satisfies the equation

$$\hat{z}^\pm(\omega) = \left[\hat{I} - \hat{\epsilon}_{\pm 3} - \hat{V}_{\pm 3, \pm 5} \hat{Z}^\pm(\omega \pm \omega_0) \right]^{-1}\tag{A6}$$

One can see from Eq. (A6) that $\hat{z}^+(\omega)$ and $\hat{z}^-(\omega)$ are related by $\hat{z}^-(\omega, V) = \hat{\sigma}_x \hat{z}^+(\omega, -V) \hat{\sigma}_x$, where $\hat{\sigma}_x$ is the corresponding Pauli matrix.

By virtue of the relation $\hat{T}_{nm}(\omega) = \hat{T}_{n-m,0}(\omega + m\omega_0/2)$, one can write the current components given by Eq. (25) in terms of $\hat{T}_{n0}(\omega) \equiv \hat{T}_n$. Using recursive relations (A5) the calculation can be reduced to a closed system for coefficients $\hat{T}_1$ and $\hat{T}_{-1}$:

$$\begin{aligned}\left[\hat{I} - \hat{\epsilon}_1 - \hat{V}_{13} \hat{z}^+(\omega) \right] \hat{T}_1 &= \hat{t}_{10} + \hat{V}_{1,-1} \hat{T}_{-1} \\ \left[\hat{I} - \hat{\epsilon}_{-1} - \hat{V}_{-1,-3} \hat{z}^-(\omega) \right] \hat{T}_{-1} &= \hat{t}_{-10} + \hat{V}_{-1,1} \hat{T}_1\end{aligned}\tag{A7}$$

The remaining task is the calculation of the transfer matrix $\hat{z}^+(\omega)$. It can be shown that the solution of Eq. (A6) is a diagonal matrix whose elements can be expressed in terms of a scalar function $\lambda^+(\omega)$

$$\hat{z}^+(\omega) = -t^2 \begin{pmatrix} f_2 f_3 \frac{\delta_0^+ \delta_1^+}{\lambda_0^+ \lambda_1^+} & 0 \\ 0 & f_1 f_2 \frac{\delta_{-1}^+ \delta_0^+}{\lambda_0^+ \lambda_1^+} \end{pmatrix}\tag{A8}$$

where $\lambda_n^+ = \lambda^+(\omega + n\omega_0/2)$ and $\delta_n^+ = (\lambda_n^+ - g_{n+2})/t^2 f_{n+2}^2$. The function $\lambda^+(\omega)$ satisfies the following equation

$$\lambda_0^+ = a + b\lambda_0^+ + c\lambda_2^+ + d\lambda_0^+ \lambda_2^+,\tag{A9}$$

where, taking into account Eq. (17), the coefficients can be written as $a = g_2 + (t^2/W^2)g_3$, $b = t^2 g_2 g_3$, $c = -t^2 [g_2 g_3 - (t^2/W^4)]$, and $d = t^2 [g_3 + (t^2/W^2)g_2]$. For an arbitrary bias voltage, Eq. (A9) can only be solved numerically. However, as we show in Appendices B and C it is possible to obtain analytical solutions in the two special cases $\omega_0 \rightarrow 0$ and $\omega_0 \rightarrow \infty$.

Once $\hat{z}^+(\omega)$ has been determined one can calculate the coefficients $\hat{T}_1$ and $\hat{T}_{-1}$ from Eq. (A7)

$$\hat{T}_1 = \frac{-t}{1 - t^2 \lambda_2^- \lambda_{-1}^-} \begin{pmatrix} t^4 f_0 f_1 \delta_2^- \delta_{-1}^+ & t^2 f_1 \delta_{-1}^+ \\ t^2 f_0 \delta_2^- & 1 \end{pmatrix},\tag{A10}$$

$$\hat{T}_{-1}(\omega, V) = -\hat{\sigma}_x \hat{T}_1(\omega, -V) \hat{\sigma}_x,\tag{A11}$$

where $\lambda_n^-(\omega, V) = \lambda_n^+(\omega, -V)$ and $\delta_n^-(\omega, V) = \delta_n^+(\omega, -V)$. The rest of the coefficients $\hat{T}_n$ can be calculated from Eqs. (A5):

$$\begin{aligned}\hat{T}_{2n+1} &= \left[\prod_{i=1}^n \hat{z}^+[\omega + (i-1)\omega_0] \right] \hat{T}_1, \quad n > 0 \\ \hat{T}_{-n} &= -\hat{\sigma}_x \hat{T}_n(\omega, -V) \hat{\sigma}_x, \quad n > 0\end{aligned}\tag{A12}$$

Finally, the current components, separated into dissipative and nondissipative parts [Eqs. (31) and (32)], can be calculated from the expressions

$$I_0 = -\frac{4e}{h} \int_{-\infty}^{\infty} d\omega \sum_{n=\text{odd}} \text{Re} \left\{ \text{Tr} \left[\hat{\sigma}_z \hat{T}_n^\dagger(\omega) \hat{g}_n^{+-} \hat{T}_n(\omega) \hat{g}_0^a \right] \right\}, \tag{A13}$$

$$\begin{aligned}I_m^D &= -\frac{4e}{h} \int_{-\infty}^{\infty} d\omega \sum_{n=\text{odd}} \text{Re} \left(\text{Tr} \left\{ \hat{\sigma}_z \left[\hat{T}_{n+m}^\dagger(\omega - m\omega_0/2) \right. \right. \right. \\ &\quad \left. \left. \left. + \hat{T}_{n-m}^\dagger(\omega + m\omega_0/2) \hat{g}_n^{+-} \hat{T}_n(\omega) \hat{g}_0^a \right] \right\} \right),\end{aligned}\tag{A14}$$

$$\begin{aligned}I_m^S &= \frac{4e}{h} \int_{-\infty}^{\infty} d\omega \sum_{n=\text{odd}} \text{Im} \left(\text{Tr} \left\{ \hat{\sigma}_z \left[\hat{T}_{n+m}^\dagger(\omega - m\omega_0/2) \right. \right. \right. \\ &\quad \left. \left. \left. - \hat{T}_{n-m}^\dagger(\omega + m\omega_0/2) \hat{g}_n^{+-} \hat{T}_n(\omega) \hat{g}_0^a \right] \right\} \right).\end{aligned}\tag{A15}$$

21.0.10 APPENDIX B

In this appendix we give details on the evaluation of the excess current for $S-S$ contacts. In the limit $eV/\Delta \rightarrow \infty$ only the dc current component I_0 survives. The infinite summation over n in Eq. (A13) can be truncated in this case neglecting the $n > 1$ terms. This is justified as the products $f_n f_{n+1}$ are negligible in this limit leading to a vanishing transfer matrix $\hat{z}^\pm(\omega)$. Physically, this is equivalent to neglecting multiple Andreev processes for $eV/\Delta \gg 1$. Then, Eq. (A13) reduces to

$$I_0 \sim \tag{B1}$$

$$-\frac{4e}{h} \int_{-\infty}^{\infty} d\omega \sum_{n=-1,1} \text{Re} \left\{ \text{Tr} \left[\hat{\sigma}_z \hat{T}_n^\dagger(\omega) \hat{g}_n^{+-}(\omega) \hat{T}_n(\omega) \hat{g}_0^a \right] \right\},$$

with

$$\begin{aligned}\hat{T}_1 &\sim \frac{-t}{1 - t^2 \lambda_2^- \lambda_{-1}^+} \begin{pmatrix} 0 & t^2 f_1 \delta_{-1}^+ \\ t^2 f_0 \delta_2^- & 1 \end{pmatrix}, \\ \hat{T}_{-1}(\omega, V) &= -\hat{\sigma}_x \hat{T}_1(\omega, -V) \hat{\sigma}_x.\end{aligned}$$

On the other hand, when neglecting contributions of order Δ/eV Eq. (A9) simply yields $\lambda_n^+ \sim (g_{n+2} + it^2/W^3)/(1 - it^2 g_{n+2}/W)$. We then obtain from Eq. (B1) the simple result $I_{\text{exc}}^{SS} = 2I_{\text{exc}}^{NN}$ for the excess current at zero temperature and any value of the transmission coefficient.

21.0.11 APPENDIX C

In this appendix we give the main steps in the analytical calculation of the current components in the limit $eV/\Delta \rightarrow 0$.

21.0.12 1. Linear regime ($\eta \gg eV$)

The small voltage response can be straightforwardly derived from Eqs. (A13)-(A15) by expanding the Fermi functions appearing in $\hat{g}^{+-}$ up to first order in eV : $n_F(\omega + n\omega_0/2) \sim n_F(\omega) - (\beta/8)n\omega_0 \text{sech}^2(\beta\omega/2)$ and evaluating the rest of these expressions at $eV = 0$. The current components can be then written as

$$I_0 = \frac{2e^2}{h} \beta V \int_{-\infty}^{\infty} d\omega \text{sech}^2\left(\frac{\beta\omega}{2}\right) \sum_{n=\text{odd} > 0} n \text{Re} \left\{ \text{Tr} \left[\hat{\sigma}_z \hat{T}_n^\dagger (\hat{g}_0^a - \hat{g}_0^r) \hat{T}_n \hat{g}_0^a \right] \right\}, \quad (\text{C1})$$

$$I_m^D = \frac{2e^2}{h} \beta V \int_{-\infty}^{\infty} d\omega \text{sech}^2\left(\frac{\beta\omega}{2}\right) \sum_{n=\text{odd} > 0} n \text{Re} \left(\text{Tr} \left\{ \hat{\sigma}_z \left[\hat{T}_{n+m}^\dagger + \hat{T}_{n-m}^\dagger \right] (\hat{g}_0^a - \hat{g}_0^r) \hat{T}_n \hat{g}_0^a \right\} \right), \quad (\text{C2})$$

$$I_m^S = \frac{8e}{h} \int_{-\infty}^{\infty} d\omega n_F(\omega) \sum_{n=0, \text{dd} > 0} \text{Im} \left(\text{Tr} \left\{ \hat{\sigma}_z \left[\hat{T}_{n+m}^\dagger - \hat{T}_{n-m}^\dagger \right] (\hat{g}_0^a - \hat{g}_0^r) \hat{T}_n \hat{g}_0^a \right\} \right), \quad (\text{C3})$$

Thus, the dissipative contribution, I_D , goes to zero as $I_D(\phi) \sim G(\phi)V$, $G(\phi)$ being the phase-dependent linear conductance, while the supercurrent part, $I_S(\phi)$, tends to a finite value at $V = 0$.

In the zero-voltage limit the coefficients $\hat{T}_n$ adopt a simple form. The transfer matrix $\hat{z}^\pm(\omega)$ becomes a scalar function: $\hat{z}^+(\omega) = \hat{z}^-(\omega) \equiv z(\omega)\hat{I}$; with $z(\omega) = -t^2 f \delta^2 / \lambda^2$, where $\lambda^+(\omega) = \lambda^-(\omega) \equiv \lambda(\omega)$ satisfies the simple quadratic equation

$$t^2 g(\omega) \lambda^2(\omega) - \left(1 - \frac{t^2}{W^2}\right) \lambda(\omega) + g(\omega) = 0 \quad (\text{C4})$$

Finally, the coefficients $\hat{T}_n$ adopt the form

$$\hat{T}_1(\omega) = \frac{-t}{1 - t^2 \lambda^2} \begin{pmatrix} t^4 f^2 \delta^2 & t^2 f \delta \\ t^2 f \delta & 1 \end{pmatrix} \quad (\text{C5})$$

$$\hat{T}_{2n+1}(\omega) = z^n(\omega) \hat{T}_1(\omega), \quad n \geq 0 \quad (\text{C6})$$

Due to these simple recursive relations, the series appearing in the current components become geometrical series, which can be summed up without difficulty. In the weakly damped case, $\eta/\alpha\Delta \ll 1$, these summations lead to analytical expressions for the dissipative and nondissipative parts of the current. By solving Eq. (C4) up to corrections of order $\eta/\alpha\Delta$ one obtains

$$z(\omega) = e^{i\varphi(\omega)} - \frac{4\omega\eta}{\alpha\Delta^2} \{i + \cotg[\varphi(\omega)]\}, \quad \Delta\sqrt{1-\alpha} \leq |\omega| \leq \Delta, \quad (\text{C7})$$

where

$$\varphi(\omega) = \arcsin \left(\frac{2}{\alpha\Delta^2} \sqrt{\Delta^2 - \omega^2} \sqrt{\omega^2 - (1-\alpha)\Delta^2} \right)$$

The summation of the geometrical series yields

$$I_0 = \frac{2e^2}{h} \beta V \int_{-\infty}^{\infty} d\omega \operatorname{sech}^2 \left(\frac{\beta\omega}{2} \right) \operatorname{Re} \{ A(\omega) \} \frac{1 + |z|^2}{(1 - |z|^2)^2}, \quad (\text{C8})$$

$$I_m^D = \frac{2e^2}{h} \beta V \int_{-\infty}^{\infty} d\omega \operatorname{sech}^2 \left(\frac{\beta\omega}{2} \right) \operatorname{Re} \left\{ A(\omega) \left[z^m + (z^*)^m \right] \frac{1 + |z|^2}{(1 - |z|^2)^2} + \frac{2mz^m}{1 - |z|^2} \right\}, \quad (\text{C9})$$

$$I_m^S = \frac{8e}{h} \int_{-\infty}^{\infty} d\omega n_F(\omega) \operatorname{Im} \left\{ A(\omega) \frac{(z^*)^m - z^m}{1 - |z|^2} \right\}, \quad (\text{C10})$$

where $A(\omega) \equiv \operatorname{Tr} [\hat{\sigma}_z \hat{T}_1^\dagger (\hat{g}^a - \hat{g}^r) \hat{T}_1 \hat{g}^a]$. It can be noticed that in this weakly damped limit the integrands goes like $1/\eta$ and the energy interval $\Delta\sqrt{1-\alpha} \leq |\omega| \leq \Delta$ give the main contribution to the current.

Finally, when summing up all ac components to obtain the total dissipative and nondissipative parts, the current densities become singular at the condition $\varphi(\omega) = \phi$. This condition is satisfied for $\omega = \omega_S = \pm \Delta\sqrt{1-\alpha} \sin^2(\phi/2)$ (i.e., at the bound-state energy levels) leading to

$$I_D(\phi) = \frac{e^2 \alpha^2 \Delta^4}{8\eta h} \beta V \int_{-\infty}^{\infty} d\omega \operatorname{sech}^2 \left(\frac{\beta\omega}{2} \right) \frac{\sin^2 \varphi(\omega)}{\omega} \operatorname{Im} \left\{ \frac{1}{(\omega - |\omega_S| - i\eta)(\omega + |\omega_S| - i\eta)} \right\} \quad (\text{C11})$$

$$I_S(\phi) = -\frac{2e}{h} \alpha \Delta^2 \sin \phi \int_{-\infty}^{\infty} d\omega n_F(\omega) \operatorname{Im} \left\{ \frac{1}{(\omega - |\omega_S| - i\eta)(\omega + |\omega_S| - i\eta)} \right\} \quad (\text{C12})$$

These integrals can be straightforwardly evaluated. Equation (C11) gives the expression for the phase-dependent linear conductance given in Sec. IV [Eq. (34)], while Eq. (C12) yields the well-known expression for the supercurrent in a singlemode SQPC.^{13,3}

$$I_S(\phi) = \frac{e \Delta^2 \alpha}{2\hbar} \frac{\sin(\phi)}{|\omega_S(\phi)|} \tanh \left(\frac{\beta |\omega_S(\phi)|}{2} \right). \quad (\text{C13})$$

21.0.13 2. Nonlinear regime ($n \ll eV$)

We first rewrite Eqs. (A13)-(A15) for the current components as

$$\begin{aligned} I_0 &= \frac{4e}{h} \int_{-\infty}^{\infty} d\omega \tanh \left(\frac{\beta\omega}{2} \right) \sum_{n=\text{odd}>0} \operatorname{Re} \left\{ \operatorname{Tr} \left[\hat{\sigma}_z \hat{T}_n^\dagger (\hat{g}_0^a - \hat{g}_0^r) \hat{T}_n \hat{g}_{-n}^a \right] \right\}, \\ I_m^D &= \frac{4e}{h} \int_{-\infty}^{\infty} d\omega \tanh \left(\frac{\beta\omega}{2} \right) \sum_{n=\text{odd}>0} \operatorname{Re} \left(\operatorname{Tr} \left\{ \hat{\sigma}_z \left[\hat{T}_{n+m}^\dagger \hat{T}_{n-m}^\dagger \right] (\hat{g}_0^a - \hat{g}_0^r) \hat{T}_n \hat{g}_{-n}^a \right\} \right), \\ I_m^S &= -\frac{4e}{h} \int_{-\infty}^{\infty} d\omega \tanh \left(\frac{\beta\omega}{2} \right) \sum_{n=\text{odd}>0} \operatorname{Im} \left(\operatorname{Tr} \left\{ \hat{\sigma}_z \left[\hat{T}_{n+m}^\dagger - \hat{T}_{n-m}^\dagger \right] (\hat{g}_0^a - \hat{g}_0^r) \hat{T}_n \hat{g}_{-n}^a \right\} \right), \end{aligned} \quad (\text{C14})$$

where a rigid shift of $n\omega_0/2$ in the energy arguments of the different $\hat{T}_n$ with respect to the ones appearing in Eqs. (A13)-(A15) has been introduced.

In the limit $eV \rightarrow 0$, the solution of Eq. (A9) is $\lambda_n^+ = \lambda [\omega + (n+2)\omega_0/2]$, where $\lambda(\omega)$ satisfies the quadratic equation (C4). The coefficients $\hat{T}_n (n > 0)$ can then be generated starting from $\hat{T}_1$ and using the transfer matrix $\hat{z}^+(\omega)$. These quantities are obtained from Eqs. (A8) and (A10) making use of the $eV \rightarrow 0$ solution for λ_n^+

The resulting expressions simplify considerably in the ballistic limit where $\lambda_n^+ = i/t$. For energies inside the gap one obtains

$$\hat{T}_n(\omega) = -\frac{t^{n-1}}{2} \prod_{j=1} e^{i\alpha_j} \begin{pmatrix} e^{i(\alpha_0+\alpha_n)} & e^{i\alpha_n} \\ e^{i\alpha_0} & 1 \end{pmatrix} \quad (\text{C15})$$

where $\alpha_n = \arccos[(\omega + neV)/\Delta]$. As discussed in Ref. 8, when written as in Eq. (C14), the main contribution to the current in this limit comes from a small energy range around the gap edges. Evaluation of these integrals leads directly to Eq. (35).

22 An introduction to the quasiclassical theory of superconductivity for diffusive proximity-coupled system by Chandrasekhar

Although the physics of normal metals (N) in close proximity to superconductors (S) has been studied extensively for more than thirty years, it is only in the past decade that experiments have been able to probe directly the region close to the NS interface at temperatures far below the transition temperature T_c of the superconductor. These experiments have been made possible by the availability of microlithography techniques that enable the fabrication of heterostructure devices with submicron scale features. This size scale is comparable to the relevant physical length scales of the problem, and consequently, a number of new effects have been observed. A variety of systems have been studied, the variation primarily being in the type of normal metal coupled to the superconductor. Canonical normal metals such as copper and gold, semiconductors, insulators, and ferromagnets have been employed. Although a variety of theoretical techniques have been used to describe proximity-coupled systems, the quasiclassical theory of superconductivity [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12] has proved to be a remarkably powerful tool in understanding the microscopic basis for the remarkable effects observed in these systems.

A number of excellent recent articles [13, 14, 15, 16, 17, 18, 19] have explored the application of the quasiclassical theory of superconductivity to proximity coupled systems. In this review, a self-contained development of the quasiclassical theory is presented, starting from non-equilibrium Keldysh Green's functions for normal metal systems. If the normal metal is clean, quasiparticles in the normal metal travel ballistically over long length scales; in the samples studied in the majority of recent experiments, however, the quasiparticles are scattered elastically within a short distance, so that the quasiparticle motion is diffusive. Here, we shall concentrate on this diffusive case, where the elastic scattering length ℓ is the shortest relevant length scale in the problem. We shall also restrict ourselves to the case where the superconductor is of the canonical s -wave type, avoiding complications with non-spatially symmetric order parameters within the superconductor.

22.1 Transport equations in the diffusive approximation

As an illustration of some of the issues that arise in dealing with non-equilibrium transport in mesoscopic diffusive systems, we consider first the classical Boltzmann equation in the diffusion approximation. Consider a one-dimensional diffusive normal metal wire of length L . We assume that all dimensions of the wire are larger than ℓ , the elastic mean free path of the conduction electrons. $f(E, x)$, the distribution function of electrons at energy E and at a point x along

the wire, obeys the diffusion equation

$$D \frac{d^2 f(E, x)}{dx^2} + C(f) = 0. \quad (22.1)$$

Here, $D = (1/3)v_F \ell$ is the three-dimensional electron diffusion constant, with v_F being the Fermi velocity of the electrons. $C(f)$ is the collision integral, which takes into account inelastic scattering of the electrons, and itself depends on the distribution function f . If we consider a mesoscopic wire whose length L is much shorter than any inelastic scattering length L_{in} , this term can be set to zero.

This diffusion equation has a simple solution in some specific geometries. Consider the case of a wire of length L sandwiched between two normal ‘reservoirs,’ which we shall call the left (at $x = 0$) and right (at $x = L$) reservoirs. The reservoirs are defined as having an equilibrium distribution function $f_L(E)$ and $f_R(E)$ respectively. Then the solution of the diffusion equation, Eqn.(22.1), under the condition $L \ll L_{in}$ (so that $C(f) = 0$), and subject to the boundary conditions that $f(E, x) = f_L(E)$ at the left reservoir, and $f(E, x) = f_R(E)$ at the right reservoir, is given by [20]

$$f(E, x) = [f_R(E) - f_L(E)] \frac{x}{L} + f_L(E) \quad (22.2)$$

The electrical current I through the wire in the diffusion approximation is given by

$$I = -eAD \int N(E) \frac{\partial f}{\partial x} dE \quad (22.3)$$

where A is the cross-sectional area of the wire. With $f(E, x)$ given by Eqn.(22.2), this can be written

$$I = -\frac{eAD}{L} \int N(E) [f_R(E) - f_L(E)] dE \quad (22.4)$$

In order to obtain a finite current, we must apply a voltage across the wire. Let us apply a voltage V to the left reservoir, keeping the right reservoir at ground ($V = 0$). This has the effect of shifting the energies of the electrons in the left reservoir by $-eV$, so that the electron distribution function there is given by $f_L(E) = f_0(E - eV)$, where $f_0(E)$ is the usual equilibrium Fermi distribution function

$$f_0(E) = \frac{1}{e^{E/k_B T} + 1} \quad (22.5)$$

and the energy E is measured from the Fermi energy E_F . Figure 1 shows the distribution function $f(E, x)$ in the wire. It has a step-function form, which varies linearly along the length of the wire. Evidence for such a non-equilibrium distribution has recently been seen experimentally in a series of beautiful experiments by the Saclay group.[21]

In the limit of small voltages V , the difference of distribution functions in Eqn.(22.4) can be expanded as

$$f_R(E) - f_L(E) = f_0(E) - f_0(E - eV) \approx f_0(E) - \left[f_0(E) + eV \left(-\frac{\partial f_0}{\partial E} \right) \right] = -eV \left(-\frac{\partial f_0}{\partial E} \right) \quad (22.6)$$

so that

$$I = \frac{e^2 AD}{L} V \int N(E) \left(-\frac{\partial f_0}{\partial E} \right) dE \approx \frac{N(0) e^2 AD}{L} V. \quad (22.7)$$

where $N(0)$ is the density of states at the Fermi energy. In performing the integral, we have assumed that we are at low enough temperatures so that the derivative of the Fermi function $(-\partial f_0 / \partial E)$ can be approximated by $\delta(E)$. The electrical conductance is then given in the Nernst-Einstein form $G = N(0) e^2 D (A/L) = \sigma_0 (A/L)$.

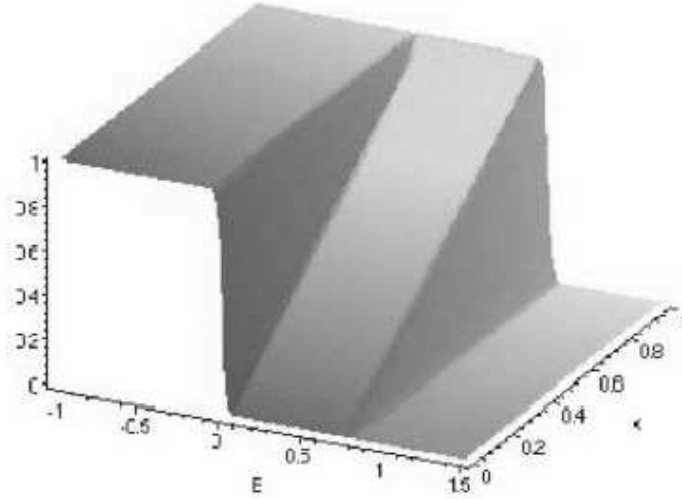


Figure 95: Nonequilibrium electron distribution function f in a normal wire with a potential V applied across it, as a function of energy E and position x . The position x is normalized to the length L of the wire.

The thermal current through the wire is given by an expression similar to Eqn.(22.4):

$$I^T = \frac{AD}{L} \int EN(E)[f_R(E) - f_L(E)]dE \quad (22.8)$$

The critical difference between the expression for the thermal current and the expression for the electrical current is the presence of an additional factor of E in the integrand in Eqn.(22.8), which has important consequences in the calculation of the thermal properties.

To obtain expressions for the thermal coefficients, let the temperature of the left reservoir be T and its voltage V , and let the temperature of the right reservoir $T + \Delta T$ and its corresponding voltage $V = 0$. The distribution function in the left electrode is then $f_L(E) = f_0(E - eV, T)$, and the distribution function in the right electrode is $f_R(E) = f_0(E, T + \Delta T)$. In the limit of small $\Delta T, V$, the difference in the distribution functions in Eqn.(22.4) can be expanded as

$$f_R(E) - f_L(E) = \left[f_0(E, T) + \frac{E}{T} \Delta T \left(-\frac{\partial f_0}{\partial E} \right) \right] - \left[f_0(E, T) + eV \left(-\frac{\partial f_0}{\partial E} \right) \right] = \left[\frac{E}{T} \Delta T - eV \right] \left(-\frac{\partial f_0}{\partial E} \right) \quad (22.9)$$

Putting this expression into the equations for the electrical and thermal currents, Eqn.(22.4) and Eqn.(22.8), one obtains the two transport equations

$$I = \frac{e^2 AD}{L} \left[\int N(E) \left(-\frac{\partial f_0}{\partial E} \right) dE \right] V - \frac{eAD}{TL} \left[\int EN(E) \left(-\frac{\partial f_0}{\partial E} \right) dE \right] \Delta T \quad (22.10a)$$

$$I^T = -\frac{eAD}{L} \left[\int EN(E) \left(-\frac{\partial f_0}{\partial E} \right) dE \right] V + \frac{AD}{TL} \left[\int E^2 N(E) \left(-\frac{\partial f_0}{\partial E} \right) dE \right] \Delta T. \quad (22.10b)$$

These equations are equivalent to the usual form for the transport equations for a metal:

$$I = GV + \eta \Delta T \quad (22.11a)$$

$$I^T = \zeta V + \kappa \Delta T \quad (22.11b)$$

with the linear-response thermoelectric coefficients defined by

$$G = \frac{e^2 AD}{L} \left[\int N(E) \left(-\frac{\partial f_0}{\partial E} \right) dE \right], \quad (22.12a)$$

$$\eta = -\frac{eAD}{TL} \left[\int EN(E) \left(-\frac{\partial f_0}{\partial E} \right) dE \right], \quad (22.12b)$$

$$\zeta = -\frac{eAD}{L} \left[\int EN(E) \left(-\frac{\partial f_0}{\partial E} \right) dE \right], \quad (22.12c)$$

$$\kappa = \frac{AD}{TL} \left[\int E^2 N(E) \left(-\frac{\partial f_0}{\partial E} \right) dE \right]. \quad (22.12d)$$

If we approximate the derivative of the Fermi function by a δ function at low temperatures, as we did for the electrical conductance G , we see that all other coefficients vanish. In order to obtain a finite value, we use the Sommerfeld expansion for the integrals [22]

$$\int \Phi(E) \left(-\frac{\partial f_0}{\partial E} \right) dE = \Phi(0) + \frac{\pi^2 (k_B T)^2}{6} \left[\frac{\partial^2 \Phi(E)}{\partial E^2} \right]_{E=0} + \dots \quad (22.13)$$

Using the fact that $N'(0) = N(0)/2E_F$, we obtain the following expressions for the last three coefficients

$$\eta = -\frac{\pi^2 k_B^2 T}{6} \frac{eAD}{L} \frac{N(0)}{E_F}, \quad (22.14a)$$

$$\zeta = -\frac{\pi^2 (k_B T)^2}{6} \frac{eAD}{L} \frac{N(0)}{E_F}, \quad (22.14b)$$

$$\kappa = \frac{\pi^2 k_B^2 T}{3} \frac{AD}{L} N(0). \quad (22.14c)$$

Experimentally, the quantities often measured are the thermopower S and the thermal conductance G^T . The thermopower is defined as the thermal voltage generated by a temperature differential ΔT , under the condition that no electrical current flows through the wire ($I = 0$). Putting this condition into the first transport equation, Eqn.(22.10a), we obtain

$$S = \frac{V}{\Delta T} = -\frac{\eta}{G} = \frac{\pi^2 k_B}{6} \frac{k_B T}{e E_F} \quad (22.15)$$

Note that since the expression for S contains a factor $k_B T/E_F$, the thermopower of a typical normal metal is very small.

The thermal conductance G^T is defined as the ratio of the thermal current I^T to the temperature differential ΔT , under the same condition of no electrical current flow ($I = 0$). From the equations above

$$G^T = \frac{I^T}{\Delta T} = \kappa + \zeta S \quad (22.16)$$

For typical metals, the second term is much smaller than the first, and is usually ignored, so that $G^T \approx \kappa$. If we take the ratio of the electrical to the thermal conductance, we obtain

$$\frac{G^T}{G} = \frac{\pi^2 k_B^2}{3} \frac{T}{e^2}. \quad (22.17)$$

Consequently, one finds that the Wiedemann-Franz Law holds, even though the scattering lengths for momentum relaxation (ℓ) and energy relaxation (L_{in}) are quite different. This is because the reservoirs are assumed to be perfect, so that any electron entering a reservoir immediately equilibrates with the other electrons in the reservoir.

Before we go on to discuss normal metals in contact with superconductors, it is worthwhile to review some important assumptions in the calculations above. First, in our calculations, we have assumed that the diffusion coefficient D is a constant independent of the energy E and position x . When coherence effects are important, as in the case of the proximity effect, the diffusion coefficient becomes a function of both these parameters, $D(E, x)$. The differential equation for the distribution function, Eqn.(22.1) is then modified. $D(E, x)$ itself in general is determined by the distribution function $f(E, x)$, so one must solve a set of coupled differential equations to obtain a solution. This is difficult to do analytically in all but the simplest of cases, and more often must be done numerically. The remainder of this chapter will be devoted in great part to deriving the appropriate differential equations for the distribution functions and diffusion coefficient in the case of normal metals in contact with superconductors, using the quasiclassical equations for superconductivity.

Second, apart from the electrical conductance G , the thermal coefficients derived above would all vanish if the density of states at the Fermi energy $N(E)$, were assumed to be constant, so that it could be taken out of the integrals. The small variation in the density of states at the Fermi energy is responsible for finite (but small) values of the off-diagonal transport coefficients η and ζ . For example, the thermopower S , which is non-zero only if there is an asymmetry between the properties of particles and holes near the Fermi energy, vanishes if $N(E)$ is assumed constant. The small difference in $N(E)$ above and below E_F gives rise to the small but finite thermopowers of typical normal metals. The conventional quasiclassical theory of superconductivity assumes particle-hole symmetry *a priori* in the definition of the quasiclassical Green's functions, in that $N(E)$ is assumed constant at E_F . Consequently, thermoelectric effects cannot be calculated in the conventional quasiclassical approximation; an extension of the theory is required.

Finally, we have been discussing here *currents* and *conductances*, rather than *current densities* and *conductivities*. These are the more relevant quantities, since we will be discussing mesoscopic samples in which the measured properties are properties of the sample as a whole. This will be particularly important for the proximity effect, where long range phase coherence means that non-local effects are important.

22.2 The Keldysh Green's Functions

The starting point for developing the quantum analog to the classical Boltzmann transport equation is the Keldysh diagrammatic technique. We shall begin our discussion of the Keldysh technique in the notation of Lifshitz and Pitaevskii.[23] As in the equilibrium case, we define a non-equilibrium Green's function

$$G_{\sigma_1\sigma_2}(X_1, X_2) = -i \left\langle n \left| T \hat{\psi}_{\sigma_1}(X_1) \hat{\psi}_{\sigma_2}^+(X_2) \right| n \right\rangle \quad (22.18)$$

Here X_1 and X_2 take into account the three spatial coordinates (denoted by $\vec{r}_1$ and $\vec{r}_2$ respectively), and the time coordinate (t_1 and t_2). The difference between the non-equilibrium Keldysh Green's function above and the usual equilibrium Green's function is that the average is taken over any quantum state $|n\rangle$ of the system, rather than just the ground state $|0\rangle$. The time ordering operator T has the effect

$$G_{\sigma_1\sigma_2}(X_1, X_2) = \begin{cases} -i \left\langle n \left| \hat{\psi}_{\sigma_1}(X_1) \hat{\psi}_{\sigma_2}^+(X_2) \right| n \right\rangle & \text{if } t_1 > t_2, \\ +i \left\langle n \left| \hat{\psi}_{\sigma_2}^+(X_2) \hat{\psi}_{\sigma_1}(X_1) \right| n \right\rangle & \text{if } t_2 > t_1 \end{cases} \quad (22.19)$$

for fermion operators (the case of interest here).

For simplicity, we label the spin indices and the coordinate indices by numbers. We shall define a set of Green's functions that will be useful later:

$$G_{12}^{\alpha\alpha} = -i \langle T \hat{\psi}_1 \hat{\psi}_2^+ \rangle, \quad (22.20a)$$

$$G_{12}^{\beta\beta} = i \langle \tilde{T} \hat{\psi}_1^+ \hat{\psi}_2 \rangle, \quad (22.20b)$$

$$G_{12}^{\alpha\beta} = -i \langle \hat{\psi}_1 \hat{\psi}_2^+ \rangle, \quad (22.20c)$$

$$G_{12}^{\beta\alpha} = i \langle \hat{\psi}_2^+ \hat{\psi}_1 \rangle. \quad (22.20d)$$

The first Green's function is just the one defined above, in this new notation. The second Green's function is similar in definition to $G_{12}^{\alpha\alpha}$, except that it is defined with operator $\tilde{T}$ instead of T , which orders the operators in reverse chronological order. The last two Green's functions are defined without any time-ordering operators. The four Green's functions so defined are not linearly independent, but are related by linear equations of the form

$$G^{\alpha\alpha} + G^{\beta\beta} = G^{\alpha\beta} + G^{\beta\alpha}. \quad (22.21)$$

The retarded and advanced Green's functions G^R and G^A can be defined as in the equilibrium case

$$G_{12}^R = \begin{cases} -i \langle \hat{\psi}_1 \hat{\psi}_2^+ + \hat{\psi}_2^+ \hat{\psi}_1 \rangle & \text{if } t_1 > t_2, \\ 0 & \text{if } t_2 > t_1 \end{cases} \quad (22.22)$$

and

$$G_{12}^A = \begin{cases} 0 & \text{if } t_1 > t_2, \\ i \langle \hat{\psi}_1 \hat{\psi}_2^+ + \hat{\psi}_2^+ \hat{\psi}_1 \rangle & \text{if } t_2 > t_1. \end{cases} \quad (22.23)$$

and are related by

$$G_{12}^A = (G_{21}^R)^*. \quad (22.24)$$

G^R and G^A can be written in terms of the Keldysh Green's functions defined earlier as

$$G^R = G^{\alpha\alpha} - G^{\alpha\beta} = G^{\beta\alpha} - G^{\beta\beta} \quad (22.25a)$$

$$G^A = G^{\alpha\alpha} - G^{\beta\alpha} = G^{\alpha\beta} - G^{\beta\beta} \quad (22.25b)$$

$G^{\alpha\alpha}$ satisfies the equation of motion

$$G_{01}^{-1} G_{12}^{(0)\alpha\alpha} = \delta(X_1 - X_2) \quad (22.26)$$

where G_0^{-1} is the differential operator (in the free electron approximation)¹³

$$G_0^{-1} = i \frac{\partial}{\partial t} + \frac{\nabla^2}{2m} + \mu \quad (22.27)$$

and the second subscript to G_0^{-1} denotes that the differentials in Eqn.(22.27) are with respect to coordinates corresponding to this subscript. The argument of the delta function Eqn.(22.26) includes space, time and spin coordinates, and the notation $G_{12}^{(0)\alpha\alpha}$ (with a superscript 0) signifies a Green's function for an ideal gas.

The δ function in Eqn.(22.26) arises from the discontinuities in $G^{\alpha\alpha}$ at $t_1 = t_2$. G^R and G^A have similar discontinuities, and obey a similar equation. $G^{(0)\beta\beta}$ has a discontinuity of the opposite sign, and hence obeys the equation

$$G_{01}^{-1} G_{12}^{(0)\beta\beta} = -\delta(X_1 - X_2) \quad (22.28)$$

¹³We set $\hbar = 1$.

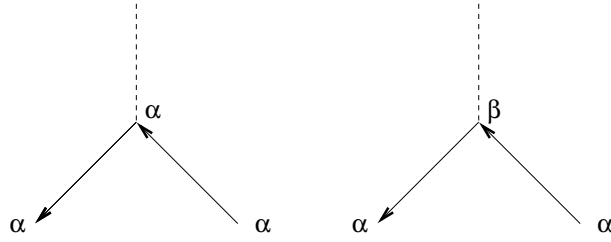


Figure 96: Diagrams corresponding to the first order corrections to the Keldysh Green's function $G_{12}^{\alpha\alpha}$ in the presence of an external potential U , which is represented by a dashed line.

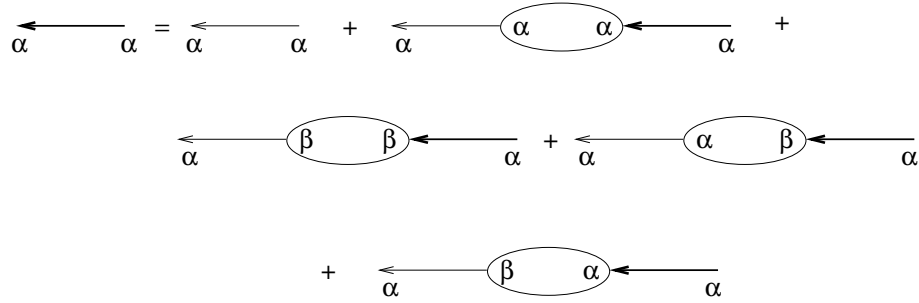


Figure 97: Dyson's equation in diagrammatic form for the Keldysh Green's function $G_{12}^{\alpha\alpha}$. The thick line represents the exact Green's function $G_{12}^{\alpha\alpha}$, the thin line represents the Green's function for an ideal gas $G_{12}^{(0)\alpha\alpha}$, and the ellipse represents the self-energy Σ .

$G^{\alpha\beta}$ and $G^{\beta\alpha}$ have no discontinuities at $t_1 = t_2$, and hence obey the equations

$$G_{01}^{-1} G_{12}^{(0)\beta\alpha} = 0 \quad (22.29a)$$

$$G_{01}^{-1} G_{12}^{(0)\alpha\beta} = 0 \quad (22.29b)$$

The diagram technique for Keldysh Green's functions is similar to that for equilibrium Green's functions, except that one needs to sum over the internal indices α and β , with a corresponding increase in the number of diagrams. For example, Fig.1 shows two diagrams corresponding to the first order corrections to $G_{12}^{\alpha\alpha}$ in the presence of an external potential, which is represented by a dashed line.

In the same manner, Dyson's equation for the Green's function can be represented as shown in Fig. 2, where the ellipse represents the self-energy function $\Sigma^{\alpha\beta}$. In order to make the notation more compact and tractable, it is useful to introduce a matrix Green's function and corresponding self-energy matrix:

$$\check{G} = \begin{pmatrix} G^{\beta\beta} & G^{\alpha\beta} \\ G^{\beta\alpha} & G^{\alpha\alpha} \end{pmatrix}, \quad \check{\Sigma} = \begin{pmatrix} \Sigma^{\beta\beta} & \Sigma^{\alpha\beta} \\ \Sigma^{\beta\alpha} & \Sigma^{\alpha\alpha} \end{pmatrix}. \quad (22.30)$$

Dyson's equation can then be written in matrix form as

$$\check{G}_{12} = \check{G}_{12}^0 + \int \check{G}_{14}^0 \check{\Sigma}_{43} \check{G}_{32} d^4 X_3 d^4 X_4 \quad (22.31)$$

where the usual rules of matrix multiplication are used.

As we saw earlier, the components of the Green's function matrix are not independent, but are linearly related. One can therefore perform a transformation to set one of the matrix components to zero. There are many ways to do this; the one we shall use here is the one employed in most recent literature on the quasiclassical theory of superconductivity.[11] The matrices above are first rotated in Keldysh space using the transformation $\check{G} \rightarrow \tau_3 \check{G}$, where

the τ matrices are identical in form to the Pauli spin matrices

$$\tau^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \tau^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (22.32)$$

and then transformed $\check{G} \rightarrow Q\check{G}Q^\dagger$, where the matrix Q is given by

$$Q = \frac{1}{\sqrt{2}}(\tau^0 - i\tau^2). \quad (22.33)$$

The resulting matrices have the form

$$\check{G} = \begin{pmatrix} G^R & G^K \\ 0 & G^A \end{pmatrix}, \quad \check{\Sigma} = \begin{pmatrix} \Sigma^R & \Sigma^K \\ 0 & \Sigma^A \end{pmatrix} \quad (22.34)$$

where the retarded and advanced Green's functions G^R and G^A have been defined above, and the new Keldysh Green's function G^K is given by

$$G^K = G^{\alpha\alpha} + G^{\beta\beta} = G^{\alpha\beta} + G^{\beta\alpha}. \quad (22.35)$$

Σ^R , Σ^A and Σ^K are defined in a similar manner. In what follows, we shall use the 'check' to denote the Keldysh matrices. In order to avoid writing integrals over the space and time coordinates, we introduce the binary operator $\otimes$, which has the effect of integrating over the free space and time coordinates and performing matrix multiplication when applied between two Keldysh matrices. Thus, the Dyson equation for the Keldysh Green's function written above can be represented as

$$\check{G} = \check{G}^0 + \check{G}^0 \otimes \check{\Sigma} \otimes \check{G} \quad (22.36)$$

$\check{G}^0$ obeys the equation of motion

$$G_{01}^{-1} \check{G}_{12}^0 = \delta(X_1 - X_2). \quad (22.37)$$

Operating G_{01}^{-1} on the Dyson equation Eqn(22.36) from the left, we obtain

$$G_{01}^{-1} \check{G} = \delta(X_1 - X_2) + \check{\Sigma} \otimes \check{G}. \quad (22.38)$$

The conjugate equation is¹⁴

$$\check{G} G_{02}^{-1} = \delta(X_1 - X_2) + \check{G} \otimes \check{\Sigma}. \quad (22.39)$$

Subtracting the two equations, we obtain

$$(G_0^{-1} \otimes \check{G} - \check{G} \otimes G_0^{-1}) - (\check{\Sigma} \otimes \check{G} - \check{G} \otimes \check{\Sigma}) = 0, \quad (22.40)$$

where we have suppressed the subscript '1' of G_{01}^{-1} . This can be written as a commutator

$$[G_0^{-1} \ominus \check{G}] - [\check{\Sigma} \ominus \check{G}] = 0$$

or

$$[G_0^{-1} \ominus \check{G}] = [\check{\Sigma} \ominus \check{G}] \quad (22.41)$$

where the $\ominus$ operator defines the commutator of two operators $[A \ominus B] = A \otimes B - B \otimes A$.

¹⁴From the right, we operate on the coordinate 2, hence we use the operator G_{02}^{-1} . G_{02}^{-1} operated on the right is equivalent to G_{02}^{-1*} operated on the left.

The mixed or Wigner representation

The Green's function $\check{G}_{12}$ oscillates rapidly with the difference $\vec{r}_2 - \vec{r}_1$, on the scale of the inverse Fermi wave vector k_F^{-1} . In the physical systems to be discussed, we are interested in variations on much longer length scales. We therefore perform a transformation to center-of-mass coordinates, $\vec{R}$, T , and difference coordinates, $\vec{r}$, t , defined by the equations [24]

$$\vec{r}_1 = \vec{R} - \vec{r}/2, \quad \vec{r}_2 = \vec{R} + \vec{r}/2 \quad (22.42a)$$

$$t_1 = T - t/2, \quad t_2 = T + t/2 \quad (22.42b)$$

and then define a Fourier transform of $\check{G}_{12}$ with respect to the variables $\vec{r}$ and t :

$$\check{G}(\vec{R}, T; \vec{r}, t) = \int e^{-iEt} e^{i\vec{p}\cdot\vec{r}} \check{G}(\vec{R}, T; \vec{p}, E) d\vec{p} dE. \quad (22.43)$$

In Eqn.(22.41), the part of the commutator involving G_0^{-1} can be written as

$$G_{01}^{-1} \otimes \check{G} - \check{G} \otimes G_{02}^{-1} = \left[i \left(\frac{\partial}{\partial t_1} + \frac{\partial}{\partial t_2} \right) - \frac{1}{2m} (\nabla_1^2 - \nabla_2^2) \right] \check{G} \quad (22.44)$$

using Eqn.(22.27) in the free electron approximation. Transforming to the coordinates, $\vec{R}$, T , $\vec{r}$, t , this can be written

$$\left(i \frac{\partial}{\partial T} + \frac{1}{m} \nabla_{\vec{r}} \cdot \nabla_{\vec{R}} \right) \check{G} \quad (22.45)$$

From Eqn.(22.35), the Keldysh component G^K of the Green's function can be written as the sum of the Green's functions $G^{\alpha\beta}$ and $G^{\beta\alpha}$. Keeping in mind the definitions of these functions given in Eqns(22.20), we can define a non-equilibrium distribution function $f(\vec{R}, T, \vec{r})$, which is related to the function $G^{\beta\alpha}$ by

$$f(\vec{R}, T, \vec{r}) = \langle \hat{\psi}^+(\vec{R} + \vec{r}/2, T + t/2) \hat{\psi}(\vec{R} - \vec{r}/2, T - t/2) \rangle_{t=0} = -iG^{\beta\alpha}(\vec{R}, T, \vec{r}, t)_{t=0}. \quad (22.46)$$

Similarly

$$\langle \hat{\psi}(\vec{R} + \vec{r}/2, T + t/2) \hat{\psi}^+(\vec{R} - \vec{r}/2, T - t/2) \rangle_{t=0} = 1 - f(\vec{R}, T, \vec{r}) = iG^{\alpha\beta}(\vec{R}, T, \vec{r}, t)_{t=0}. \quad (22.47)$$

Subtracting the first equation from the second, we obtain

$$1 - 2f(\vec{R}, T, \vec{r}) = h(\vec{R}, T, \vec{r}) = i(G^{\alpha\beta}(\vec{R}, T, \vec{r}, t)_{t=0} + G^{\beta\alpha}(\vec{R}, T, \vec{r}, t)_{t=0}) = iG^K(\vec{R}, T, \vec{r}, t)_{t=0}, \quad (22.48)$$

where we have defined a new distribution function $h(\vec{R}, T, \vec{r})$. In terms of the mixed Fourier transform, Eqn.(22.43), this can be written as

$$h(\vec{R}, T, \vec{r}) = iG^K(\vec{R}, T, \vec{r}, t)_{t=0} = \frac{i}{2\pi} \int e^{i\vec{p}\cdot\vec{r}} G^K(\vec{R}, T; \vec{p}, E) d\vec{p} dE. \quad (22.49)$$

Taking the Keldysh component of Eqn.(22.44) at $t = 0$, we obtain

$$\left(i \frac{\partial}{\partial T} + \frac{1}{m} \nabla_{\vec{r}} \cdot \nabla_{\vec{R}} \right) \int d\vec{p} \frac{dE}{2\pi} e^{i\vec{p}\cdot\vec{r}} G^K(\vec{R}, T; \vec{p}, E), \quad (22.50)$$

or, in terms of the Fourier components with respect to $\vec{r}$, corresponding to momentum $\vec{p}$,

$$\frac{\partial h(\vec{R}, T, \vec{p})}{\partial T} + \frac{\vec{p}}{m} \cdot \nabla_{\vec{R}} h(\vec{R}, T, \vec{p}) \quad (22.51)$$

where the Fourier transform of the Wigner distribution function $h(\vec{R}, T, \vec{p})$ is given by

$$h(\vec{R}, T, \vec{p}) = \frac{i}{2\pi} \int G^K(\vec{R}, T; \vec{p}, E) dE. \quad (22.52)$$

In equilibrium, h is given by

$$h_0(\epsilon_{\vec{p}}) = 1 - 2f_0(\epsilon_{\vec{p}}) = \tanh(\epsilon_{\vec{p}}/2k_B T). \quad (22.53)$$

Eqn(22.51), which is the Keldysh component of the left hand side of Eqn(22.41), has the form of the one side of the classical Boltzmann equation for the distribution function. (Using the definition of the function $h(\vec{R}, T, \vec{p})$, this can also be written in a more conventional form in terms of $f(\vec{R}, T, \vec{p})$.) The right hand side of the Keldysh component of Eqn.(22.41) must therefore correspond to the collision terms. The right hand side of the Keldysh component can be written as $2(\Sigma^{\beta\alpha} G^{\alpha\beta} - \Sigma^{\alpha\beta} G^{\beta\alpha})$. Taking the limit at $t = 0$, and writing in terms of the distribution function $f(\vec{R}, T, \vec{p})$ using Eqn.(22.45) and Eqn.(22.46), this Keldysh component of the right hand side of Eqn(22.41) can be written as [23]

$$-2i[\Sigma^{\beta\alpha}(R, T, p)(1 - f(\vec{R}, T, \vec{p})) + \Sigma^{\alpha\beta}(R, T, p)f(\vec{R}, T, \vec{p})]. \quad (22.54)$$

The first term in Eqn.(22.54) with the factor $(1 - f(\vec{R}, T, \vec{p}))$ has the usual form of a scattering-in term, corresponding to the gain of particles, while the second term has the form for a scattering-out term, corresponding to a loss of particles. Consequently, we see that the Keldysh component of the right-left subtracted Dyson equation gives the transport equation for the distribution function. From the diagonal components components of the same equation, one can obtain solutions for the other components of $\check{G}$ and $\check{\Sigma}$. More typically, the scattering terms on the right hand side of the Boltzmann equation make it difficult to solve, and some approximations must be employed. If the variation of the system with the center-of-mass coordinates T and $\vec{R}$ is small, then one can expand the Green's functions and self-energies, which are functions of $\vec{R}, \vec{r}, T, t$ about $\vec{R}$ and T in a Taylor's series to first order in $\vec{r}$ and t . This is the gradient expansion discussed by Kadanoff and Baym[24] and Larkin and Ovchinnikov[4], and we shall return to it at the end of this section.

Instead of taking the difference of Eqn(22.38) and its conjugate equation, Eqn(22.39), we take the sum, we obtain the equation

$$[G_0^{-1} \oplus \check{G}] = 2\delta(X_1 - X_2) + [\check{\Sigma} \oplus \check{G}] \quad (22.55)$$

where the operator $\oplus$ defines the Keldysh anti-commutator, in the same way as the operator $\ominus$ defines the Keldysh commutator. The left hand side of Eqn.(22.55) can be written as

$$G_{01}^{-1} \otimes \check{G} + \check{G} \otimes G_{02}^{-1} = \left[i \left(\frac{\partial}{\partial t_1} - \frac{\partial}{\partial t_2} \right) - \frac{1}{2m} (\nabla_1^2 + \nabla_2^2) \right] \check{G}. \quad (22.56)$$

Transforming to the mixed representation, we obtain

$$\left[2i \frac{\partial}{\partial t} + \frac{1}{m} \left(\frac{1}{4} \nabla_{\vec{R}}^2 + \nabla_{\vec{r}}^2 \right) \right] \check{G}(\vec{R}, T, \vec{r}, t) \quad (22.57)$$

Now, the assumption we are making is that the variations of G on the scale of $\vec{R}$ are much slower than the variations on the scale of $\vec{r}$. Hence, the the terms in equation above involving derivatives with $\vec{R}$ contribute much less than the those with $\vec{r}$, and can be neglected in this approximation. If we consider the equation for G_0 , for which the terms involving Σ on the right hand side of Eqn.(22.55) are 0, then we obtain, after transforming to Fourier components[25]

$$(E - \epsilon_{\vec{p}})G_0(\vec{R}, T, \vec{p}, E) = 1 \quad (22.58)$$

or

$$G_0(\vec{R}, T, \vec{p}, E) = \frac{1}{(E - \epsilon_{\vec{p}})}, \quad (22.59)$$

where

$$\epsilon_{\vec{p}} = \frac{p^2}{2m} - \mu \sim v_F(p - p_F). \quad (22.60)$$

So far, we have assumed a free-electron model. If there is a slowly varying potential $U(\vec{R}, T)$, the equation above can be modified to

$$G_0(\vec{R}, T, \vec{p}, E) = \frac{1}{(E - \epsilon_{\vec{p}} - U(\vec{R}, T))} \quad (22.61)$$

This equation has form of a Green's function for a free particle, but in a slowly varying potential $U(\vec{R}, T)$. The operator G_0^{-1} in the mixed representation can therefore be written as

$$G_0^{-1} = (E - \epsilon_{\vec{p}} - U(\vec{R}, T)). \quad (22.62)$$

More accurately, one transforms Eqn.(22.27) using the following equations

$$\frac{\partial}{\partial t_{1,2}} = \frac{1}{2} \frac{\partial}{\partial T} \pm \frac{\partial}{\partial t}, \quad (22.63a)$$

$$\nabla_{1,2} = \frac{1}{2} \nabla_{\vec{R}} \pm \nabla_{\vec{r}} \quad (22.63b)$$

$$\nabla_{1,2}^2 = \frac{1}{4} \nabla_{\vec{R}}^2 \pm \nabla_{\vec{r}} \cdot \nabla_{\vec{R}} + \nabla_{\vec{r}}^2. \quad (22.63c)$$

Again, assuming the functions in the mixed representation are slowly varying functions of $\vec{R}$, we can ignore the second derivatives with respect to $\vec{R}$, to obtain

$$G_0^{-1} = \frac{1}{2} \frac{\partial}{\partial T} + \frac{\partial}{\partial t} + \frac{\nabla_{\vec{r}}^2}{2m} + \frac{\nabla_{\vec{r}} \cdot \nabla_{\vec{R}}}{2m} + \mu \quad (22.64)$$

In most applications of interest here, we also ignore the slow T dependence. Adding the potential $U(\vec{R}, T)$, and writing in terms of the partial Fourier transform with respect to $\vec{p}$ and E , we obtain

$$G_0^{-1} = E - \epsilon_{\vec{p}} + \frac{i}{2} \vec{v}_F \cdot \nabla_{\vec{R}} + \mu - U(\vec{R}, T), \quad (22.65)$$

where we have replaced $\vec{p}/m$ by $\vec{v}_F$, since the important region of interest is at the Fermi surface. Note that we still keep account of the *direction* of the momentum.

To conclude this section, we derive some expressions for the convolution of two operators in the mixed representation, the gradient expansion discussed above. Consider the convolution of two operators defined as an integral over the internal space and time coordinates

$$A \otimes B = \int d\vec{r}_3 \int dt_3 A(\vec{r}_1, \vec{r}_3, t_1, t_3) B(\vec{r}_3, \vec{r}_2, t_3, t_2) \quad (22.66)$$

If we do a partial transform to $\vec{p}, \vec{R}$ (but do not transform the time coordinates), $A \otimes B$ can be represented as

$$A \otimes B = \int dt_3 e^{\frac{i}{2}(\nabla_{\vec{R}}^A \cdot \nabla_{\vec{p}}^B - \nabla_{\vec{p}}^A \cdot \nabla_{\vec{R}}^B)} A(\vec{R}, \vec{p}, t_1, t_3) B(\vec{R}, \vec{p}, t_3, t_2) \quad (22.67)$$

where the superscripts to the derivatives denote that they operate only those functions. If the derivatives are small, we need to take only the first order expansion of this expression

$$(A \otimes B)(\vec{p}, \vec{R}) = \int dt_3 \left[1 + \frac{i}{2} (\nabla_{\vec{R}}^A \cdot \nabla_{\vec{p}}^B - \nabla_{\vec{p}}^A \cdot \nabla_{\vec{R}}^B) \right] A(\vec{R}, \vec{p}, t_1, t_3) B(\vec{R}, \vec{p}, t_3, t_2). \quad (22.68)$$

Similarly

$$(B \otimes A)(\vec{p}, \vec{R}) = \int dt_3 \left[1 + \frac{i}{2} (\nabla_{\vec{R}}^B \cdot \nabla_{\vec{p}}^A - \nabla_{\vec{p}}^B \cdot \nabla_{\vec{R}}^A) \right] B(\vec{R}, \vec{p}, t_1, t_3) A(\vec{R}, \vec{p}, t_3, t_2). \quad (22.69)$$

If we were dealing with functions alone, then the multiplication of the two function A and B is commutative. When A and B are matrices, however, they do not in general commute, so we obtain

$$(A \ominus B)(\vec{p}, \vec{R}) = A \otimes B - B \otimes A = \int dt_3 [A, B] + \frac{i}{2} \int dt_3 [\{\nabla_{\vec{R}} A, \nabla_{\vec{p}} B\} - \{\nabla_{\vec{p}} A, \nabla_{\vec{R}} B\}], \quad (22.70)$$

where $[A, B]$ notation stands for the commutator of the two functions, and $\{A, B\}$ stands for the anticommutator. A similar equation can be obtained if we transform the times to the mixed representation T, E

$$(A \ominus B)(T, E) = A \otimes B - B \otimes A = \int d\vec{r}_3 [A, B] + \frac{i}{2} \int d\vec{r}_3 [\{\partial_T A, \partial_E B\} - \{\partial_E A, \partial_T B\}], \quad (22.71)$$

When we transform both sets of variables, we obtain

$$\begin{aligned} (A \ominus B)(\vec{p}, \vec{R}, T, E) &= A \otimes B - B \otimes A \\ &= [A, B] + \frac{i}{2} [(\{\partial_T A, \partial_E B\} - \{\partial_E A, \partial_T B\}) + (\{\nabla_{\vec{R}} A, \nabla_{\vec{p}} B\} - \{\nabla_{\vec{p}} A, \nabla_{\vec{R}} B\})]. \end{aligned} \quad (22.72)$$

In most cases, we are interested in stationary situations, where there is no T dependence. In this case, the equation above reduces to

$$(A \ominus B)(\vec{p}, \vec{R}, T, E) = A \otimes B - B \otimes A = [A, B] + \frac{i}{2} [\{\nabla_{\vec{R}} A, \nabla_{\vec{p}} B\} - \{\nabla_{\vec{p}} A, \nabla_{\vec{R}} B\}]. \quad (22.73)$$

These expressions will be useful in the calculations to follow.

22.3 The quasiclassical approximation

The non-equilibrium spectral function A is defined in the same way as in the equilibrium case[11, 26]

$$A = \frac{i}{2\pi} (G^R - G^A) = -\frac{1}{\pi} \text{Im} (G^R), \quad (22.74)$$

where $\text{Im} (G^R)$ denotes the imaginary part of the retarded Green's function. In the equilibrium case, A defines the spectrum of energy levels, for stationary quantum states, it has the form of a sum of δ -functions at each state energy. In the quasiparticle approximation, these δ -functions are broadened, but the width Γ of each state, defining its lifetime, is still small compared to its energy. If Γ is large, the quasiparticle approximation breaks down, and one cannot obtain a kinetic equation for a distribution function by integrating over the energy E . However, for most perturbations of interest, the self-energies typically have a weak dependence on the magnitude of the momentum, this dependence being appreciable only near the Fermi energy. On the other hand, while the Dyson equation has a strong dependence on E and $\epsilon_{\vec{p}}$, the subtracted Dyson

equation, Eqn.(22.41), from which we will obtain the equations of motion, has a very weak dependence on both E and $\epsilon_{\vec{p}}$. In this case, it is possible to average over the particle energy to eliminate the dependence on the magnitude of the momentum, but keep the dependence on the direction of the momentum. Hence, one can think of replacing the Green's functions and self-energies by their values on the Fermi surface, multiplied by a δ -function in the form

$$G(\vec{R}, \vec{p}, t_1, t_2) \rightarrow \delta(\epsilon_{\vec{p}}) g(\vec{R}, \hat{p}, t_1, t_2) \quad (22.75)$$

To this end, we define the so-called quasiclassical Green's function

$$g(\vec{R}, \hat{p}, t_1, t_2) = \frac{i}{\pi} \int d\epsilon_{\vec{p}} G(\vec{R}, \vec{p}, t_1, t_2). \quad (22.76)$$

Care must be taken in performing this integral, since the integrand falls off only as $1/\epsilon_{\vec{p}}$ for large $\epsilon_{\vec{p}}$. To avoid this, one can introduce a cut-off in the integral, as done by Serene and Rainer,[7] or following Eilenberger,[1] use a special contour for integration.

We would like to obtain an equation of motion for the quasiclassical Green's functions. If we obtain an equation of motion by operating G_0^{-1} in the form of Eqn(22.64) on Dyson's equation, and then integrating over $\epsilon_{\vec{p}}$, there are terms in the equation that will have large contributions. In order to eliminate these large terms, we start from the left-right subtracted equation of motion, Eqn.(22.41), in which the troublesome terms are cancelled, and then integrate over $\epsilon_{\vec{p}}$. The terms on the right hand side of Eqn.(22.41) are of the form

$$\check{\Sigma} \otimes \check{G} = \int dt_3 dx_3^3 \check{\Sigma}(x_1, t_1, x_3, t_3) \check{G}(x_3, t_3, x_2, t_2) \quad (22.77)$$

We first Fourier transform this term with respect to $\vec{p}$, which takes care of the integral over x_3 . We then average the resulting equation with respect to $\epsilon_{\vec{p}}$. The assumption here is that only $\check{G}$ has a strong dependence on the momentum $\vec{p}$, so the result of this averaging is a term of the form

$$\int dt_3 \check{\Sigma}(\vec{R}, \vec{p}, t_1, t_3) \check{g}(\vec{R}, \hat{p}, t_3, t_2) \quad (22.78)$$

Now, to complete the transformation, the self-energy $\check{\Sigma}$, which is a functional of the Green's functions $\check{G}$, must become a functional only of the quasiclassical Green's functions $\check{g}$, $\check{\Sigma}[\check{G}] \rightarrow \check{\sigma}[\check{g}]$. With this final change, Eqn.(22.41) for the quasiclassical Green's functions becomes

$$[(g_0^{-1} - \check{\sigma}) \circ \check{g}] = 0 \quad (22.79)$$

The $\circ$ operator in the commutator $[A \circ B]$ involves an integral over the internal time coordinates in addition to the usual matrix multiplication for Keldysh matrices. If we transform the time coordinates as well, this integral can be removed, as shown at the end of the last section (Eqn.22.72). In this case, the commutator becomes a simple commutator (but involving matrix multiplication of the Keldysh matrices).

It remains to express the physical quantities of interest in terms of the Green's functions. The particle density is given by

$$\rho(1) = -2iG_{11}^{\beta\alpha} \quad (22.80)$$

and the current density in the absence of external fields is given by

$$\vec{j}(1) = -\frac{e}{m} (\nabla_1 - \nabla_2) G_{12}^{\beta\alpha}|_{2=1} \quad (22.81)$$

From the definitions of G^R , G^A and G^K in terms of $G^{\alpha\beta}$, $G^{\beta\alpha}$ and $G^{\beta\beta}$, we can write the function $G^{\beta\alpha}$ as

$$G^{\beta\alpha} = \frac{1}{2} [G^K + (G^R - G^A)]. \quad (22.82)$$

G^R and G^A depend on the equilibrium properties of the system, and so do not contribute to the current on nonequilibrium density. Consequently, these terms can be dropped in the expression for the particle density and current. Writing in terms of the quasiclassical Green's functions, we can obtain expressions for the density and the current in the mixed representation. Consider the expression for charge density

$$\delta\rho(\vec{R}, T) = -ieG^K(\vec{R}, T, \vec{r} = 0, t = 0), \quad (22.83)$$

where we use $\delta\rho$ instead of ρ to emphasize that this does not include the equilibrium contributions. Expanding G^K in terms of Fourier components in momentum space $\vec{p}$

$$\begin{aligned} \delta\rho(\vec{R}, T) &= -ie \left[\int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot\vec{r}} G^K(\vec{R}, T, \vec{p}, t) \right]_{\vec{r}=0, t=0} \\ &= -ie \int \frac{dE}{2\pi} \int \frac{d^3p}{(2\pi)^3} G^K(\vec{R}, T, \vec{p}, E) \end{aligned} \quad (22.84)$$

where we have used the fact that setting $t = 0$ in $G^K(t)$ is equivalent to integrating $G^K(E)$ over $dE/2\pi$.

The equivalent expression for the current density can also be written in terms of the quasiclassical Green's function. From Eqn(22.81) and Eqn(22.82), we have, in the mixed representation

$$\vec{j}(\vec{R}, T) = -\frac{e}{2m} \left[(\nabla_1 - \nabla_2) G^K(\vec{R}, T, \vec{r}, \vec{p}) \right]_{\vec{r}=0, t=0} \quad (22.85)$$

The integral over the momentum $\vec{p}$ can be rewritten as

$$\int \frac{d^3p}{(2\pi)^3} \rightarrow N_0 \int d\epsilon_{\vec{p}} \int \frac{d\Omega_p}{4\pi} \quad (22.86)$$

where N_0 is the density of states at the Fermi energy, and $d\Omega_p$ is an element of solid angle in momentum space p . Writing $\nabla_1 - \nabla_2 = 2\nabla_{\vec{r}}$ from Eqn.(22.63b), we obtain

$$\vec{j}(\vec{R}, T) = -\frac{e}{m} \left[\nabla_{\vec{r}} \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot\vec{r}} G^K(\vec{R}, T, \vec{p}, t) \right]_{\vec{r}=0, t=0}. \quad (22.87)$$

Operating $\nabla_{\vec{r}}$ on the exponential within the integral gives $i\vec{p}$. In the quasiclassical approximation, we assume that the major contribution comes from near the Fermi surface, so that $\vec{p}/m = \vec{v}_F$, the Fermi velocity. With this approximation, we obtain

$$\vec{j}(\vec{R}, T) = -ie \int \frac{dE}{2\pi} \int \frac{d^3p}{(2\pi)^3} \vec{v}_F G^K(\vec{R}, T, \vec{p}, E). \quad (22.88)$$

If we use the definition of the distribution function h given by Eqn.(22.48), we obtain

$$\vec{j}(\vec{R}, T) = -e \int \frac{d^3p}{(2\pi)^3} \vec{v}_F h(\vec{R}, T, \vec{p}) \quad (22.89)$$

which is the expected classical form for the current. Recalling the definition of the quasiclassical Green's function, Eqn(22.76), we obtain

$$\vec{j}(\vec{R}, T) = -\frac{1}{2}eN_0 \int dE \int \frac{d\Omega_p}{4\pi} \vec{v}_F g^K(E, \hat{p}, \vec{R}, T) \quad (22.90)$$

where we have used Eqn(22.76) and Eqn.(22.86). Note that writing these expressions in terms of the quasiclassical Green's functions necessitates reversing the order of integration over E and $\epsilon_{\vec{p}}$.

The transformation, Eqn.(22.86), assumes that the density of states is constant at the Fermi energy, and hence assumes that particle-hole symmetry holds. Hence, within this approximation, we will not be able to obtain any results on physical phenomenon that depend on particle-hole asymmetry, in particular, thermoelectric effects. This should be contrasted with the derivation of the current in the diffusive limit that we performed in the introduction, where the energy dependence of the density of states was taken into account explicitly.

So far, we have ignored the effects of external fields and potentials. In particular, the effect of a magnetic field is of interest. The magnetic field is introduced in the form of a vector potential $\vec{A}(\vec{R}, T)$. Here, we shall consider only time independent fields. $\vec{A}(\vec{R})$ is introduced by making the change

$$\nabla_{\vec{R}} \rightarrow \nabla_{\vec{R}} - ie\vec{A}(\vec{R}) \equiv \partial_{\vec{R}} \quad (22.91)$$

in all equations involving the spatial derivative. For example, the operator $\tilde{G}_{01}^{-1}$ is now written

$$\tilde{G}_{01}^{-1} = i\frac{\partial}{\partial t_1} + \frac{\partial_{\vec{r}}^2}{2m} + \mu - e\phi. \quad (22.92)$$

In the equation above, we have also added a term $e\phi$, corresponding to the presence of a scalar potential.

In our formulation, we would like all observable quantities to be invariant under a gauge transformation of the vector potential

$$\vec{A} \rightarrow \vec{A} + \nabla\chi(\vec{r}), \quad (22.93)$$

which also transforms the potential

$$\phi \rightarrow \phi - \frac{\partial\chi}{\partial t}. \quad (22.94)$$

Eqn.(22.87) for the electrical current is then modified to

$$\vec{j}(\vec{R}, T) = -\frac{2e}{m} \left[[\nabla_{\vec{r}} - ie\vec{A}(\vec{R})] \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot\vec{r}} G^{\beta\alpha}(\vec{R}, T, \vec{p}, t) \right]_{\vec{r}=0, t=0}, \quad (22.95)$$

where we have written the current in terms of $G^{\beta\alpha}$ rather than G^K . It follows that Eqn.(22.88) can be written as

$$\vec{j}(\vec{R}, T) = -i2e \int \frac{dE}{2\pi} \int \frac{d^3p}{(2\pi)^3} \vec{v}_F G^{\beta\alpha}(\vec{R}, T, \vec{p}, E) + \frac{2ie^2}{m} \vec{A}(\vec{R}) \int \frac{dE}{2\pi} \int \frac{d^3p}{(2\pi)^3} G^{\beta\alpha}(\vec{R}, T, \vec{p}, E). \quad (22.96)$$

The second term in the equation above is called the diamagnetic term. It is cancelled by a contribution from the first term arising from energies far from the Fermi energy, which are not taken into account in the quasiclassical Green's function as defined by Eqn.(22.76).[11] Note that such a cancellation does not occur in the case of a superconductor, and the second term gives rise to the supercurrent, which is proportional to the vector potential and the phase gradient according to Eqn.(22.94). With this high energy contribution cancelled by the diamagnetic term, the equation above transforms into the expression Eqn.(22.90) for the electrical current, written now in terms of the Keldysh component of the quasiclassical Green's function.

Such a cancellation does not occur in transforming $\delta\rho$ in terms of the quasiclassical Green's functions, and the contribution of the integrals in Eqn.(22.84) must be explicitly calculated.[11] The result that was obtained by Eliashberg[3] can be written as

$$\delta\rho(\vec{R}, T) = -\frac{eN_0}{2} \int dE \int \frac{d\Omega_p}{4\pi} g^K(E, \hat{p}, \vec{R}, T) - 2e^2 N_0 \phi(\vec{R}, T), \quad (22.97)$$

where $\phi(\vec{R}, T)$ is now the scalar electrochemical potential. This expression can also be obtained by invoking gauge invariance arguments. Since we must preserve charge neutrality in the system, $\delta\rho = 0$, so that $\phi(\vec{R}, T)$ is given by

$$\phi(\vec{R}, T) = -\frac{1}{4e} \int dE \int \frac{d\Omega_p}{4\pi} g^K(E, \hat{p}, \vec{R}, T) \quad (22.98)$$

For completeness, we also write expressions for the thermal current and the density of states. The expression for the thermal current in terms of the Keldysh Green's function is given by

$$\vec{j}_{th}(\vec{R}, T) = i \int \frac{dE}{2\pi} \int \frac{d^3p}{(2\pi)^3} E \vec{v}_F G^K(\vec{R}, T, \vec{p}, E), \quad (22.99)$$

and the corresponding form in terms of the quasiclassical Green's functions

$$\vec{j}_{th}(\vec{R}, T) = \frac{1}{2} N_0 \int dE \int \frac{d\Omega_p}{4\pi} E \vec{v}_F g^K(E, \hat{p}, \vec{R}, T). \quad (22.100)$$

As in the equilibrium case, the density of states is given directly by the spectral function A defined in Eqn.(22.74), now expressed in the mixed representation

$$\begin{aligned} N(E, \vec{R}, T) = A(E, \vec{R}, T) &= -\frac{1}{\pi} \text{Im} G^R(E, \vec{R}, T) \\ &= -\frac{1}{\pi} \text{Im} \int \frac{d^3p}{(2\pi)^3} G^R(E, \vec{p}, \vec{R}, T). \end{aligned} \quad (22.101)$$

Written in terms of the quasiclassical Green's function, this becomes

$$N(E, \vec{R}, T) = N_0 \text{Re} \int \frac{d\Omega_p}{4\pi} g^R(\vec{R}, E, T), \quad (22.102)$$

where the notations Re and Im stand for the real and imaginary components respectively.

22.4 Non-equilibrium Green's functions for superconducting systems

The formalism that we have developed for non-equilibrium Green's functions for normal systems carries over into the superconducting case, except that the Green's functions for the superconducting case are more complicated. In order to show how the Green's functions are defined, we start by expressing the Green's functions in the superconducting case in terms of field operators in the Nambu-Gorkov formalism. In this formalism, the field operators can be written as two-component column matrices

$$\hat{\Psi}_1 = \begin{pmatrix} \hat{\psi}_{1\uparrow} \\ \hat{\psi}_{1\downarrow}^+ \end{pmatrix} \quad (22.103)$$

where the up and down arrows refer to the spin indices, and the numbers now refer only to the time and space coordinates. The Hermitian adjoint of this operator can be written as a two-component row matrix

$$\hat{\Psi}_1^+ = \left(\hat{\psi}_{1\uparrow}^+ \quad \hat{\psi}_{1\downarrow} \right). \quad (22.104)$$

The multiplication of two such operators is defined as the tensorial product of the two matrices. For example

$$\hat{\Psi}_1 \hat{\Psi}_2^+ = \begin{pmatrix} \hat{\psi}_{1\uparrow} \hat{\psi}_{2\uparrow}^+ & \hat{\psi}_{1\uparrow} \hat{\psi}_{2\downarrow}^+ \\ \hat{\psi}_{1\downarrow}^+ \hat{\psi}_{2\uparrow}^+ & \hat{\psi}_{1\downarrow}^+ \hat{\psi}_{2\downarrow}^+ \end{pmatrix} \quad (22.105)$$

It is natural then to define the Green's functions in terms of these products. For example, the natural definition of the Green's function corresponding to $G^{\alpha\alpha}$, Eqn.(22.20a), would be

$$G_{12}^{\alpha\alpha} = -i \langle T \hat{\Psi}_1 \hat{\Psi}_2^+ \rangle = -i \begin{pmatrix} \langle T \hat{\psi}_{1\uparrow} \hat{\psi}_{2\uparrow}^+ \rangle & \langle T \hat{\psi}_{1\uparrow} \hat{\psi}_{2\downarrow}^+ \rangle \\ \langle T \hat{\psi}_{1\downarrow}^+ \hat{\psi}_{2\uparrow}^+ \rangle & \langle T \hat{\psi}_{1\downarrow}^+ \hat{\psi}_{2\downarrow}^+ \rangle \end{pmatrix}. \quad (22.106)$$

However, we would like to keep the same form for the equations of motion and Dyson's equation as for the normal case, except, of course, the quantities will be matrices in Nambu-Gorkov space. To see if the definition above fits this requirement, let us operate $i\partial/\partial t_1$ on the Green's function for an ideal gas, as defined above. With the definition

$$h_{01} = -\frac{\nabla_1^2}{2m} - \mu, \quad (22.107)$$

we have

$$i \frac{\partial}{\partial t_1} G_{12}^{(0)\alpha\alpha} = h_{01} \begin{pmatrix} \langle T \hat{\psi}_{1\uparrow} \hat{\psi}_{2\uparrow}^+ \rangle & \langle T \hat{\psi}_{1\uparrow} \hat{\psi}_{2\downarrow}^+ \rangle \\ -\langle T \hat{\psi}_{1\downarrow}^+ \hat{\psi}_{2\uparrow}^+ \rangle & -\langle T \hat{\psi}_{1\downarrow}^+ \hat{\psi}_{2\downarrow}^+ \rangle \end{pmatrix} + \begin{pmatrix} \delta(X_1 - X_2) & 0 \\ 0 & \delta(X_1 - X_2) \end{pmatrix} \quad (22.108)$$

This does not have the required form of the analogous equation for the normal case, Eqn.(22.26). To remedy this, we define our Green's functions with an extra factor of τ^3 . The definitions corresponding to Eqns.(22.20) are then

$$\tilde{G}_{12}^{\alpha\alpha} = -i\tau^3 \langle T \hat{\Psi}_1 \hat{\Psi}_2^+ \rangle = -i\tau^3 \begin{pmatrix} \langle T \hat{\psi}_{1\uparrow} \hat{\psi}_{2\uparrow}^+ \rangle & \langle T \hat{\psi}_{1\uparrow} \hat{\psi}_{2\downarrow}^+ \rangle \\ \langle T \hat{\psi}_{1\downarrow}^+ \hat{\psi}_{2\uparrow}^+ \rangle & \langle T \hat{\psi}_{1\downarrow}^+ \hat{\psi}_{2\downarrow}^+ \rangle \end{pmatrix}, \quad (22.109a)$$

$$\tilde{G}_{12}^{\beta\beta} = -i\tau^3 \langle \tilde{T} \hat{\Psi}_1 \hat{\Psi}_2^+ \rangle = -i\tau^3 \begin{pmatrix} \langle \tilde{T} \hat{\psi}_{1\uparrow} \hat{\psi}_{2\uparrow}^+ \rangle & \langle \tilde{T} \hat{\psi}_{1\uparrow} \hat{\psi}_{2\downarrow}^+ \rangle \\ \langle \tilde{T} \hat{\psi}_{1\downarrow}^+ \hat{\psi}_{2\uparrow}^+ \rangle & \langle \tilde{T} \hat{\psi}_{1\downarrow}^+ \hat{\psi}_{2\downarrow}^+ \rangle \end{pmatrix}, \quad (22.109b)$$

$$\tilde{G}_{12}^{\alpha\beta} = -i\tau^3 \langle \hat{\Psi}_1 \hat{\Psi}_2^+ \rangle = -i\tau^3 \begin{pmatrix} \langle \hat{\psi}_{1\uparrow} \hat{\psi}_{2\uparrow}^+ \rangle & \langle \hat{\psi}_{1\uparrow} \hat{\psi}_{2\downarrow}^+ \rangle \\ \langle \hat{\psi}_{1\downarrow}^+ \hat{\psi}_{2\uparrow}^+ \rangle & \langle \hat{\psi}_{1\downarrow}^+ \hat{\psi}_{2\downarrow}^+ \rangle \end{pmatrix}, \quad (22.109c)$$

and

$$\tilde{G}_{12}^{\beta\alpha} = i\tau^3 \langle \hat{\Psi}_2^+ \hat{\Psi}_1 \rangle = i\tau^3 \begin{pmatrix} \langle \hat{\psi}_{2\uparrow}^+ \hat{\psi}_{1\uparrow} \rangle & \langle \hat{\psi}_{2\downarrow}^+ \hat{\psi}_{1\uparrow} \rangle \\ \langle \hat{\psi}_{2\uparrow}^+ \hat{\psi}_{1\downarrow} \rangle & \langle \hat{\psi}_{2\downarrow}^+ \hat{\psi}_{1\downarrow} \rangle \end{pmatrix}, \quad (22.109d)$$

where the 'tilde' over the Green's functions denotes that they are matrices in Nambu-Gorkov space. The operator corresponding to Eqn.(22.26) is then defined as

$$\tilde{G}_{01}^{-1} = i\tau^3 \frac{\partial}{\partial t_1} + \frac{\partial_1^2}{2m} + \mu. \quad (22.110)$$

where it is understood that any 'scalar' quantities in the equation above and in what follows are multiplied by the identity matrix τ^0 . With these modifications, all the equations derived for the Keldysh Green's functions for the normal case can be carried over directly to the superconducting case. In particular, the advanced and retarded Green's functions $\tilde{G}^R$ and $\tilde{G}^A$ are defined in terms of the Green's functions in Eqn.(22.109) in the same manner as before. The Keldysh matrices corresponding to Eqn.(22.34) are then

$$\hat{G} = \begin{pmatrix} \tilde{G}^R & \tilde{G}^K \\ 0 & \tilde{G}^A \end{pmatrix}, \quad \hat{\Sigma} = \begin{pmatrix} \tilde{\Sigma}^R & \tilde{\Sigma}^K \\ 0 & \tilde{\Sigma}^A \end{pmatrix}. \quad (22.111)$$

Since each element of these matrices are themselves 2x2 matrices, the resulting Keldysh matrices for superconductors are 4x4 'supermatrices,' and we shall denote them by a 'hat' symbol ($\hat{}$). The equation of motion equivalent to Eqn.(22.41) is

$$[\hat{G}_0^{-1} \ominus \hat{G}] = [\hat{\Sigma} \ominus \hat{G}]. \quad (22.112)$$

Before we go any further, we need to specify the self-energy $\tilde{\Sigma}$. At the temperatures of interest in the experiments on proximity systems, the important self-energy terms are due to electron-phonon scattering ($\tilde{\Sigma}_{e-p}$), and electron impurity scattering ($\tilde{\Sigma}_{imp}$). For conventional superconductors, the elastic component ($\tilde{\Sigma}_{e-p}^R + \tilde{\Sigma}_{e-p}^A$) of the electron-phonon contribution is the one that leads to the coupling between superconducting electrons[8] (the other contributions of electron-phonon scattering will be ignored, under the assumption that we are at low enough temperatures so that inelastic electron-phonon scattering can be ignored). Perhaps a simpler way of dealing with this component of the electron-phonon interaction is to start directly with the Gorkov equations of motion[26] for the Green's function Eqn.(22.109a), which we write in the form (ignoring impurity scattering for the moment)

$$\begin{pmatrix} i\frac{\partial}{\partial t_1} + \frac{\nabla_1^2}{2m} + \mu & \Delta \\ -\Delta^* & -i\frac{\partial}{\partial t_1} + \frac{\nabla_1^2}{2m} + \mu \end{pmatrix} \tilde{G}_{12}^{\alpha\alpha} = \delta(X_1 - X_2), \quad (22.113)$$

with the pair potential Δ defined as¹⁵

$$\Delta = \lambda \lim_{2 \rightarrow 1+} \langle T \hat{\psi}_{1\uparrow} \hat{\psi}_{2\downarrow} \rangle = i\lambda \lim_{2 \rightarrow 1+} [\tilde{G}_{12}^{\alpha\alpha}]_{12} \quad (22.114)$$

in terms of the upper left component of the Green's function $\tilde{G}_{12}^{\alpha\alpha}$, which is frequently called the anomalous Green's function (or pair amplitude), and denoted by F . Here λ is the coupling constant. For a uniform bulk superconductor Δ is real. For a normal metal, λ vanishes, and hence, although the pair amplitude in a normal metal may be finite, the pair potential vanishes. Eqn.(22.114) defines a self-consistent equation for the pair potential Δ ; it is defined in terms of the anomalous Green's function F , which in turn is determined by an equation that depends on Δ . Note that unlike $[\tilde{G}_{12}^{\alpha\alpha}]_{11}$, $[\tilde{G}_{12}^{\alpha\alpha}]_{12}$ is continuous at $t_2 = t_1$, so that Δ can also be written in terms of $[\tilde{G}_{12}^{\alpha\beta}]_{12}$ or $[\tilde{G}_{12}^{\beta\alpha}]_{12}$ at $t_2 = t_1$.

Still ignoring impurity scattering, the equation of motion Eqn.(22.112) can then be written in compact form as

$$\left[\left(i\hat{\tau}^3 \frac{\partial}{\partial t_1} + \frac{\nabla_1^2}{2m} + \mu + \hat{\Delta} \right) \ominus \hat{G} \right] = 0. \quad (22.115)$$

Here $\hat{\Delta}$ represents a 4x4 matrix

$$\hat{\Delta} = \begin{pmatrix} \tilde{\Delta} & 0 \\ 0 & \tilde{\Delta} \end{pmatrix}, \quad (22.116)$$

where

$$\tilde{\Delta} = \begin{pmatrix} 0 & \Delta \\ -\Delta^* & 0 \end{pmatrix}, \quad (22.117)$$

with Δ defined by Eqn.(22.114), and

$$\hat{\tau}^3 = \begin{pmatrix} \tau^3 & 0 \\ 0 & \tau^3 \end{pmatrix}. \quad (22.118)$$

¹⁵ Δ defined this way differs from the conventional definition by a factor of i . Note that the signs associated with Δ in the matrix above are different from the conventional definition of Gorkov's equation, because of the τ^3 factor in the definition of the Green's functions.

Before we move on to making the quasiclassical approximation, it is useful to obtain expressions for measurable quantities in terms of the Green's functions defined in this section. Following Eqn.(22.49), one can define a distribution function by averaging the Keldysh component of the Green's function defined in Eqn.(22.111)

$$\tilde{h}(\vec{R}, T, \vec{p}) = \frac{i}{2\pi} \int dE \tilde{G}^K(\vec{R}, T; \vec{p}, E). \quad (22.119)$$

However, since $\tilde{G}^K$ is a 2x2 matrix, $\tilde{h}$ is also a 2x2 matrix, so that its interpretation as a simple distribution function (in the flavor of Eqn.(22.49) for the equilibrium case, for example) is not immediately clear. A more physical interpretation can be obtained by diagonalizing the Gorkov equations given in Eqn.(22.113). [12] The matrix on the left hand side of this equation can be diagonalized by a unitary transformation; this transformation, of course, is just the Bogoliubov-Valatin transformation, which results in a diagonal 'energy' matrix with eigenvalues

$$E_{\vec{p}}^2 = \epsilon_{\vec{p}}^2 + \Delta^2. \quad (22.120)$$

The same transformation also diagonalizes the *equilibrium* distribution matrix $\tilde{h}_0$, which now has the form

$$\tilde{h}_0 \simeq \begin{pmatrix} 1 - 2f_{0\uparrow}(E_{\vec{p}}) & 0 \\ 0 & 2f_{0\downarrow}(E_{\vec{p}}) - 1 \end{pmatrix}. \quad (22.121)$$

The top-left component is for electron-like excitations, and the bottom-right component for hole-like excitations. This suggests that for our combined Nambu-Gorkov-Keldysh Green's functions, the equations for the electric current and thermal current should be modified to

$$\vec{j}(\vec{R}, T) = -\frac{ie}{2} \int \frac{dE}{2\pi} \int \frac{d^3p}{(2\pi)^3} \vec{v}_F \text{Tr} \left[\tau^3 \tilde{G}^K(\vec{R}, T, \vec{p}, E) \right], \quad (22.122)$$

and

$$\vec{j}_{th}(\vec{R}, T) = -\frac{i}{2} \int \frac{dE}{2\pi} \int \frac{d^3p}{(2\pi)^3} E \vec{v}_F \text{Tr} \left[\tilde{G}^K(\vec{R}, T, \vec{p}, E) \right]. \quad (22.123)$$

Taking the trace of the matrix takes the contributions of both electrons and holes. Note that the expression for electrical current includes a factor of τ^3 in the argument of the trace, while the thermal current does not.

22.5 Quasiclassical Superconducting Green's Functions

In principle, the properties of the superconducting system may be calculated starting from Gorkov's equations. In practice, however, such calculations are difficult in all but the simplest cases. The problem is that Gorkov's equations contain information at length scales much finer than those of interest. The way around this is to make the quasiclassical approximation as we did for the case of the normal Green's functions, which we proceed to do below.

Taking into account impurity scattering, the equation of motion for the superconducting Green's function can now be written as

$$\left[\left(i\hat{\tau}^3 \frac{\partial}{\partial t_1} + \frac{\partial_1^2}{2m} + \mu + \hat{\Delta} - \hat{\Sigma}_{imp} \right) \ominus \hat{G} \right] = 0, \quad (22.124)$$

where the impurity self-energy $\hat{\Sigma}_{imp}$ includes contributions from both spin-flip and spin-independent elastic scattering.

The elastic contribution to the self-energy can be written as

$$\hat{\Sigma}_0(\vec{p}) = N_i \int \frac{d^3p'}{(2\pi)^3} \left| v(\vec{p} - \vec{p}') \right|^2 \hat{G}(\vec{p}'). \quad (22.125)$$

Here v is the impurity potential, and N_i the number of impurities per unit volume. We assume that $v(\vec{p})$ is independent of the magnitude of $\vec{p}$, so that

$$\hat{\Sigma}_0(p) = N_i N_0 \int d\epsilon_{\vec{p}'} \frac{d\Omega_{\vec{p}'}}{4\pi} |v(\hat{p} \cdot \hat{p}')|^2 \hat{G}(\vec{p}'). \quad (22.126)$$

Defining the elastic scattering rate $1/\tau$ in the Born approximation by

$$\frac{1}{\tau} = 2\pi N_i N_0 \int \frac{d\Omega_{\vec{p}'}}{4\pi} |v(\hat{p} \cdot \hat{p}')|^2, \quad (22.127)$$

we can write

$$\hat{\Sigma}_0 = \frac{1}{2\pi\tau} \int d\epsilon_{\vec{p}} \hat{G}(p) \quad (22.128)$$

Here $\hat{G}(p)$ is a function of the magnitude of $\vec{p}$ alone. Similarly, for the contribution from spin-flip scattering, one obtains

$$\hat{\Sigma}_{sf} = \frac{1}{2\pi\tau_{sf}} \int d\epsilon_{\vec{p}} \hat{\tau}^3 \hat{G}(p) \hat{\tau}^3. \quad (22.129)$$

From Eqn.(22.111), one has

$$\left[\left(i\hat{\tau}^3 \frac{\partial}{\partial t_1} + \frac{\partial_1^2}{2m} + \mu + \hat{\Delta} - \hat{\Sigma}_0 - \hat{\Sigma}_{sf} \right) \ominus \hat{G} \right] = 0. \quad (22.130)$$

Converting the left hand side of this equation to relative coordinates using Eqn.(22.63) (neglecting any derivatives with respect to the center-of-mass time T), and using only the first term of Eqn.(22.73) to lowest order, we obtain

$$[\hat{\tau}^3 E + \hat{\Delta}, \hat{G}] + i\vec{v}_F \cdot \partial_{\vec{R}} \hat{G} - [\hat{\Sigma}_0 + \hat{\Sigma}_{sf}, \hat{G}] = 0. \quad (22.131)$$

Note that there are no integrations over space or time coordinates in the equation above, but only matrix multiplications. Written in terms of the quasiclassical Green's functions, this becomes

$$[\hat{\tau}^3 E + \hat{\Delta}, \hat{g}] + i\vec{v}_F \cdot \partial_{\vec{R}} \hat{g} - [\hat{\sigma}_0 + \hat{\sigma}_{sf}, \hat{g}] = 0, \quad (22.132)$$

where

$$\hat{\sigma}_0 = -i \frac{1}{2\tau} \hat{g}_s, \quad (22.133)$$

and

$$\hat{\sigma}_{sf} = -i \frac{1}{2\tau_{sf}} \hat{\tau}^3 \hat{g}_s \hat{\tau}^3, \quad (22.134)$$

where the subscript to the quasiclassical Green's functions denotes that they are averaged over all angles of $\vec{p}$. Eqn.(22.132) is Eilenberger's equation,[1] and is the starting point for many calculations on superconducting systems.

To complete the transformation to quasiclassical Green's functions, we express the equations for physically observable quantities in terms of the quasiclassical Green's functions. From Eqn.(22.114), the gap parameter Δ can be expressed as

$$\Delta = i\lambda [\tilde{G}_{11+}^{\alpha\beta}]_{12} = \frac{i}{2} \lambda [\tilde{G}_{11+}^K]_{12} \quad (22.135)$$

using the definition Eqn.(22.35) of the Keldysh Green's function. In the mixed representation, this can be written as

$$\begin{aligned} \Delta &= \frac{i\lambda}{2} \int \frac{dE}{2\pi} \int \frac{d^3p}{(2\pi)^3} [G^K(\vec{R}, T, \vec{p}, E)]_{12} \\ &= N_0 \frac{\lambda}{4} \int dE \int \frac{d\Omega_p}{4\pi} [g^K(\vec{R}, t, \hat{p}, E)]_{12} \\ &= N_0 \frac{\lambda}{8} \int dE \int \frac{d\Omega_p}{4\pi} \text{Tr} \{ (\tau^1 - i\tau^2) g^K(\vec{R}, t, \hat{p}, E) \} \end{aligned} \quad (22.136)$$

From Eqns.(22.122) and (22.123), the electrical current and thermal currents are given by

$$\vec{j}(\vec{R}, T) = -\frac{eN_0}{4} \int dE \int \frac{d\Omega_p}{4\pi} \vec{v}_F \text{Tr} \left[\tau^3 \tilde{g}^K(\vec{R}, T, \vec{p}, E) \right], \quad (22.137)$$

and

$$\vec{j}_{th}(\vec{R}, T) = -\frac{N_0}{4} \int dE \int \frac{d\Omega_p}{4\pi} E \vec{v}_F \text{Tr} \left[\tilde{g}^K(\vec{R}, T, \vec{p}, E) \right]. \quad (22.138)$$

One can also define an equation that gives the amount of charge associated with quasiparticle excitations. In a superconductor in equilibrium, the number of particles and holes are equal, so the quasiparticle charge should vanish. If there is an imbalance in the population of electrons and holes, one can obtain a *charge imbalance*, [2, 27] usually denoted by Q^* , which is given by

$$\begin{aligned} Q^* &= -\frac{ie}{2} \int \frac{dE}{2\pi} \int \frac{d^3p}{(2\pi)^3} \text{Tr} \left[\tilde{G}^K(\vec{R}, T, \vec{p}, E) \right] \\ &= -\frac{eN_0}{4} \int dE \int \frac{d\Omega_p}{4\pi} \text{Tr} \left[\tilde{g}^K(\vec{R}, T, \vec{p}, E) \right] \end{aligned} \quad (22.139)$$

Q^* is essentially the first term on the right hand side of Eqn.(22.97). Invoking charge neutrality, this then would result in a electrochemical potential given Eqn.(22.98)

$$\phi(\vec{R}, T) = -\frac{1}{N_0 e} Q^* \quad (22.140)$$

The normalization condition and the distribution function

In obtaining the equation of motion of the Green's functions by subtracting the Dyson equation from its conjugate, some information was lost regarding the norm of the quasiclassical Green's function $\hat{g}$. The norm of $\hat{g}$ can be obtained by the normalization condition for the quasiclassical Green's function obtained by Eilenberger,[1] and Larkin and Ovchinnikov[4]

$$\hat{g}\hat{g} = \hat{\tau}^0, \quad (22.141)$$

where (4x4) matrix multiplication is implied. This normalization condition can be obtained by explicit calculation for a bulk system in equilibrium. Furthermore, it can be shown that the normalization condition is consistent with Eilenberger's equation Eqn.(22.132).

Eqn.(22.141) is equivalent to the three (2x2) matrix equations

$$\tilde{g}^R \tilde{g}^R = \tau^0, \quad (22.142a)$$

$$\tilde{g}^A \tilde{g}^A = \tau^0, \quad (22.142b)$$

and

$$\tilde{g}^R \tilde{g}^K + \tilde{g}^K \tilde{g}^A = 0. \quad (22.142c)$$

As in the case of the E -integrated Green's functions, we would like to obtain an equation of motion for a distribution function, equivalent to the quantum Boltzmann equation derived earlier. We introduce such a distribution function $\tilde{h}$ by the ansatz[11]

$$\tilde{g}^K = \tilde{g}^R \tilde{h} - \tilde{h} \tilde{g}^A. \quad (22.143)$$

This form of the quasiclassical Keldysh Green's function satisfies the normalization condition Eqn.(22.142c), as can be verified by direct substitution

$$\tilde{g}^R (\tilde{g}^R \tilde{h} - \tilde{h} \tilde{g}^A) + (\tilde{g}^R \tilde{h} - \tilde{h} \tilde{g}^A) \tilde{g}^A = 0 \quad (22.144)$$

using Eqns.(22.142a) and(22.142b). Note that the function $\tilde{h}$ is not uniquely defined; if $\tilde{h}$ is replaced by $\tilde{h} + \tilde{h}'$, where

$$\tilde{h}' = \tilde{g}^R \tilde{x} + \tilde{x} \tilde{g}^A \quad (22.145)$$

and $\tilde{x}$ is an arbitrary matrix function, the right hand side of Eqn.(22.143) is unchanged.[7] This arbitrariness allows us some flexibility in choosing the distribution function $\tilde{h}$. At low frequencies, for example, following Schmid and Schön,[2] and Larkin and Ovchinnikov,[5] we may choose $\tilde{h}$ to be diagonal in particle-hole space

$$\tilde{h} = h_L \tau^0 + h_T \tau^3. \quad (22.146)$$

The subscripts refer to the longitudinal (h_L) and transverse (h_T), a terminology introduced by Schmid and Schön to refer to changes that are associated with the magnitude (h_L) or phase (h_T) of the complex order parameter. In equilibrium, $h_L(E) = 1 - 2f_0(E)$ and $h_T(E) = 0$.

Another possible choice is one introduced by Shelankov[28]

$$\tilde{h} = h_1 \tau^0 + \tilde{g}^R h_2. \quad (22.147)$$

This representation has the advantage that the equation for the distribution function reduces to a form very similar to the Boltzmann equation when the quasiparticle approximation is valid. In what follows, we shall use the representation of Schmid and Schön.

The representation of the Keldysh Green's function in terms of a distribution function allows us to obtain an equation of motion for the distribution function from the equation of motion for the Green's function. From Eilenberger's equation (22.132), we obtain three equations of motion for the three components of the quasiclassical Green's function

$$[E\tau^3 + \tilde{\Delta}, \tilde{g}^R] + i\vec{v}_F \cdot \partial_{\vec{R}} \tilde{g}^R - [\tilde{\sigma}^R, \tilde{g}^R] = 0 \quad (22.148a)$$

$$[E\tau^3 + \tilde{\Delta}, \tilde{g}^A] + i\vec{v}_F \cdot \partial_{\vec{A}} \tilde{g}^A - [\tilde{\sigma}^A, \tilde{g}^A] = 0 \quad (22.148b)$$

$$[E\tau^3 + \tilde{\Delta}, \tilde{g}^K] + i\vec{v}_F \cdot \partial_{\vec{R}} \tilde{g}^K - \tilde{\sigma}^R \tilde{g}^K - \tilde{\sigma}^K \tilde{g}^A + \tilde{g}^R \tilde{\sigma}^K + \tilde{g}^K \tilde{\sigma}^A = 0 \quad (22.148c)$$

where $\tilde{\sigma} = \tilde{\sigma}_0 + \tilde{\sigma}_{sf}$. Substituting $\tilde{g}^K$ from Eqn.(22.143) into Eqn.(22.148c), we obtain

$$[E\tau^3 + \tilde{\Delta}, \tilde{g}^R \tilde{h} - \tilde{h} \tilde{g}^A] + i\vec{v}_F \cdot \partial_{\vec{R}} (\tilde{g}^R \tilde{h} - \tilde{h} \tilde{g}^A) - \tilde{\sigma}^R (\tilde{g}^R \tilde{h} - \tilde{h} \tilde{g}^A) - \tilde{\sigma}^K \tilde{g}^A + \tilde{g}^R \tilde{\sigma}^K + (\tilde{g}^R \tilde{h} - \tilde{h} \tilde{g}^A) \tilde{\sigma}^A = 0. \quad (22.149)$$

Subtracting from the equation above Eqn.(22.148a) multiplied by $\tilde{h}$ on the right, and adding Eqn.(22.148b) multiplied by $\tilde{h}$ on the left, we obtain

$$\tilde{g}^R B[\tilde{h}] - B[\tilde{h}] \tilde{g}^A = 0, \quad (22.150)$$

where

$$B[\tilde{h}] = [E\tau^3 + \tilde{\Delta}, \tilde{h}] + \tilde{\sigma}^K - (\tilde{\sigma}^R \tilde{h} - \tilde{h} \tilde{\sigma}^A) + i\vec{v}_F \cdot \partial_{\vec{R}} \tilde{h}. \quad (22.151)$$

Equation(22.150) is the required equation of motion for the distribution function $\tilde{h}$.

As an example, we shall calculate the Green's functions for the equilibrium case for a bulk superconductor, in the limit where $\sigma = 0$. Equation(22.148a) in this limit has the form

$$[E\tau^3 + \tilde{\Delta}, \tilde{g}^R] = 0 \quad (22.152)$$

First, if we represent the retarded Green's function by

$$\tilde{g}^R = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}, \quad (22.153)$$

the normalization condition Eqn.(22.142a) and Eqn.(22.152) together imply that

$$g_{11} = -g_{22}, \quad (22.154a)$$

$$g_{11}^2 + g_{21}g_{12} = 1, \quad (22.154b)$$

$$g_{21} = -g_{12} \frac{\Delta^*}{\Delta} \quad (22.154c)$$

and

$$g_{11} = g_{12} \frac{E}{\Delta}. \quad (22.154d)$$

Solving these equations, we obtain

$$\tilde{g}^R = \begin{pmatrix} \frac{E}{\sqrt{E^2 - |\Delta|^2}} & \frac{\Delta}{\sqrt{E^2 - |\Delta|^2}} \\ -\frac{\Delta^*}{\sqrt{E^2 - |\Delta|^2}} & -\frac{E}{\sqrt{E^2 - |\Delta|^2}} \end{pmatrix}. \quad (22.155)$$

In taking the square roots in the equation above, a factor of i will appear if $E^2 < |\Delta|^2$. This is the expected solution from Gorkov's equations. Note that g_{11} is just the normalized BCS density of states for $E^2 > |\Delta|^2$, so that g_{11} represents a generalized density of states.

From the equations above, it is clear that $g_{21} = -g_{12}^*$. To obtain $\tilde{g}^A$, we can use the relation Eqn.(22.24), which in terms of Nambu-Gorkov matrices reads

$$\tilde{G}^A = -\tau^3 \tilde{G}^{R+} \tau^3 \quad \text{or} \quad \tilde{g}^A = -\tau^3 \tilde{g}^{R+} \tau^3 \quad (22.156)$$

where the $^+$ symbol denotes the Hermitian conjugate.

22.6 The dirty limit: the Usadel equation

In systems where elastic impurity scattering is strong, the motion of quasiparticles is not ballistic, but diffusive. In this case, the effect of the strong impurity scattering is to randomize the momentum of the quasiparticles, so that it makes sense to average the properties of system over the directions of the momentum, keeping in mind that the magnitude of the momentum is still its value at the Fermi surface, p_F .

With this in mind, let us consider a system with strong impurity scattering, which we define as a system in which the impurity scattering rate $1/\tau$, as defined by Eqn.(22.127), is much larger than any energy in the problem (E, Δ). In this case, we expand the quasiclassical Green's function to first order in momentum (essentially an expansion in spherical harmonics)

$$\hat{g} = \hat{g}_s + \hat{p}\hat{g}_p \quad (22.157)$$

where $\hat{p}$ denotes a vector of unit magnitude in the direction of $\vec{p}$, and $\hat{g}_s$ and $\hat{g}_p$ are independent of the direction of $\vec{p}$, and the subscripts stand for s -wave and p -wave components of $\hat{g}$. The assumption is that $\hat{g}_p \ll \hat{g}_s$. The self-energy is expanded in a similar fashion

$$\hat{\sigma} = \hat{\sigma}_0 + \hat{\sigma}_{sf} = \hat{\sigma}_s + \hat{p}\hat{\sigma}_p \quad (22.158)$$

We would like to calculate the components of $\hat{\sigma}$ in terms of the components of $\hat{g}$, for which we shall need the definition of $\hat{\sigma}$ in terms of $\hat{g}$. Ignoring spin-flip scattering for the moment, this relation is given by Eqn.(22.126), which can be rewritten in terms of the quasiclassical Green's functions in the form

$$\hat{\sigma}_0(p) = -i\pi N_i N_0 \int \frac{d\Omega_{p'}}{4\pi} |v(\hat{p} \cdot \hat{p}')|^2 \hat{g}(\vec{p}') = -i\pi N_i N_0 \int \frac{d\Omega_{p'}}{4\pi} |v(\hat{p} \cdot \hat{p}')|^2 (\hat{g}_s + \hat{p}'\hat{g}_p), \quad (22.159)$$

using the expansion Eqn.(22.157) for the Green's function above. In order to do this, we take the dot product of $\hat{p}$ with both sides of the equation above

$$\hat{p} \cdot \hat{\sigma} = \hat{p} \hat{\sigma}_s + (\hat{p} \cdot \hat{p}) \hat{\sigma}_p = \hat{p} \hat{\sigma}_s + \hat{\sigma}_p. \quad (22.160)$$

We perform a similar operation on Eqn.(22.159)

$$\hat{p} \cdot \hat{\sigma}_0(p) = -i\pi N_i N_0 \int \frac{d\Omega_{p'}}{4\pi} |v(\hat{p} \cdot \hat{p}')|^2 (\hat{p} \hat{g}_s + (\hat{p} \cdot \hat{p}') \hat{g}_{p'}). \quad (22.161)$$

Since both $\hat{p}$ and $\hat{g}_s$ are independent of the integration over $d\Omega_{p'}$, the first term under the integral in the equation above can be written as

$$-i\pi \hat{p} \hat{g}_s N_i N_0 \int \frac{d\Omega_{p'}}{4\pi} |v(\hat{p} \cdot \hat{p}')|^2 = -\frac{i}{2\tau} \hat{p} \hat{g}_s, \quad (22.162)$$

using the definition Eqn.(22.127) of $1/\tau$. If we consider non-spin-flip scattering alone, then $\hat{\sigma}$ can be obtained by equating like terms in Eqns.(22.160) and (22.161). If we include spin-flip scattering, one can write

$$\hat{\sigma}_s = -\frac{i}{2\tau} \hat{g}_s - \frac{i}{2\tau_{sf}} \hat{\tau}^3 \hat{g}_s \hat{\tau}^3 \quad (22.163)$$

as expected from Eqns.(22.133) and (22.134). For the p -wave component, we consider only the contribution from non spin-flip impurity scattering, under the assumption that it is much stronger than the spin-flip scattering. The second term under the integral in Eqn.(22.161) can be written as

$$-i\pi N_i N_0 \int \frac{d\Omega_{p'}}{4\pi} |v(\hat{p} \cdot \hat{p}')|^2 (\hat{p} \cdot \hat{p}') \hat{g}_{p'} = -i\pi N_i N_0 \hat{g}_p \int \frac{d\Omega_{p'}}{4\pi} |v(\hat{p} \cdot \hat{p}')|^2 [1 - (1 - \hat{p} \cdot \hat{p}')], \quad (22.164)$$

where $\hat{g}_p$ can be taken out of the integral since it is independent of the direction of $\vec{p}$. The first term in the square bracket can be seen to be related to the elastic scattering rate $1/\tau$. The remaining terms can be written in terms of the transport time, defined by

$$\frac{1}{\tau_{tr}} = 2\pi N_i N_0 \int \frac{d\Omega_{p'}}{4\pi} |v(\hat{p} \cdot \hat{p}')|^2 (1 - \hat{p} \cdot \hat{p}'), \quad (22.165)$$

that is well known in the transport theory of metals. With this definition, $\hat{\sigma}_p$ can be written as

$$\hat{\sigma}_p = -\frac{i}{2} \left(\frac{1}{\tau} - \frac{1}{\tau_{tr}} \right) \hat{g}_p. \quad (22.166)$$

By putting Eqn.(22.157) into the normalization condition for the quasiclassical Green's function, and neglecting terms quadratic in $\hat{g}_p$, we also obtain the two equations

$$\hat{g}_s \hat{g}_s = 1 \quad (22.167a)$$

and

$$\hat{g}_s \hat{g}_p + \hat{g}_p \hat{g}_s = 0. \quad (22.167b)$$

We now proceed to expand the Eilenberger equation, Eqn.(22.132), in terms of the s and p -wave expansions of $\hat{g}$ and $\hat{\sigma}$. Replacing $\vec{v}_F$ by $v_F \hat{p}$, we obtain

$$[\hat{\tau}^3 E - \hat{\Delta}, \hat{g}_s] + \hat{p} [\hat{\tau}^3 E - \hat{\Delta}, \hat{g}_p] + i v_F \hat{p} \cdot \partial_{\vec{R}} (\hat{g}_s + \hat{p} \hat{g}_p) - [\hat{\sigma}_s, \hat{g}_s] - \hat{p} [\hat{\sigma}_s, \hat{g}_p] + [\hat{\sigma}_p, \hat{g}_s] + \hat{p} \cdot \hat{p} [\hat{\sigma}_p, \hat{g}_p] = 0. \quad (22.168)$$

The last term is second order in the small quantity $\hat{g}_p$, and can be neglected. Collecting the terms that are even in $\hat{p}$, we obtain

$$[\hat{\tau}^3 E + \hat{\Delta}, \hat{g}_s] + i v_F (\hat{p} \cdot \hat{p}) \partial_{\vec{R}} \hat{g}_p = 0. \quad (22.169)$$

Averaging this equation over all directions of $\hat{p}$ gives

$$[\hat{\tau}^3 E + \hat{\Delta}, \hat{g}_s] + i \frac{v_F}{3} \partial_{\vec{R}} \hat{g}_p = 0. \quad (22.170)$$

Ignoring spin-dependent scattering for the moment, the terms that are odd in $\hat{p}$ can be written

$$[\hat{\tau}^3 E + \hat{\Delta}, \hat{g}_p] + i v_F \partial_{\vec{R}} \hat{g}_s - \frac{i}{2\tau_{tr}} [\hat{g}_p, \hat{g}_s] = 0, \quad (22.171)$$

where we have used Eqns.(22.163) and (22.166) to write $\hat{\sigma}_s$ and $\hat{\sigma}_p$ in terms of $\hat{g}_s$ and $\hat{g}_p$. If elastic scattering is strong, the first term in the equation above can be neglected compared to the third, so we obtain

$$i v_F \partial_{\vec{R}} \hat{g}_s + \frac{i}{\tau_{tr}} \hat{g}_s \hat{g}_p = 0, \quad (22.172)$$

where we have used Eqn.(22.167b) to simplify the second term. Multiplying this equation by $\hat{g}_s$ on the left, and using Eqn.(22.167a), we obtain

$$v_F \hat{g}_s \partial_{\vec{R}} \hat{g}_s = -\frac{1}{\tau_{tr}} \hat{g}_p, \quad (22.173)$$

or writing $\hat{g}_p$ in terms of $\hat{g}_s$

$$\hat{g}_p = v_F \tau_{tr} \hat{g}_s \partial_{\vec{R}} \hat{g}_s = -\ell \hat{g}_s \partial_{\vec{R}} \hat{g}_s \quad (22.174)$$

where we have introduced the elastic scattering length $\ell = v_F \tau_{tr}$. Putting this into Eqn.(22.171), we obtain

$$[\hat{\tau}^3 E + \hat{\Delta}, \hat{g}_s] - i \frac{v_F \ell}{3} \partial_{\vec{R}} \hat{g}_s \partial_{\vec{R}} \hat{g}_s = 0, \quad (22.175)$$

or writing this in terms of the diffusion coefficient $D = (1/3)v_F \ell$, and reintroducing the spin-flip scattering term,

$$[\hat{\tau}^3 E + \hat{\Delta} - \hat{\sigma}_{sf}, \hat{g}_s] - i D \partial_{\vec{R}} \hat{g}_s \partial_{\vec{R}} \hat{g}_s = 0. \quad (22.176)$$

This is the equation first derived by Usadel,[29] and forms the starting point for most discussions on dirty superconducting systems.

In the remainder of our development, we shall neglect the spin-flip scattering term. Writing $\hat{g}_s$ as a matrix

$$\hat{g}_s = \begin{pmatrix} \tilde{g}_s^R & \tilde{g}_s^K \\ 0 & \tilde{g}_s^A \end{pmatrix}, \quad (22.177)$$

we can write the matrix equation, Eqn.(22.176), as three separate equations

$$[\tau^3 E + \tilde{\Delta}, \tilde{g}_s^R] = i D \partial_{\vec{R}} (\tilde{g}_s^R \partial_{\vec{R}} \tilde{g}_s^R), \quad (22.178a)$$

$$[\tau^3 E + \tilde{\Delta}, \tilde{g}_s^A] = i D \partial_{\vec{R}} (\tilde{g}_s^A \partial_{\vec{R}} \tilde{g}_s^A), \quad (22.178b)$$

and

$$[\tau^3 E + \tilde{\Delta}, \tilde{g}_s^K] = i D \partial_{\vec{R}} [(\tilde{g}_s^R \partial_{\vec{R}} \tilde{g}_s^K) + (\tilde{g}_s^K \partial_{\vec{R}} \tilde{g}_s^A)]. \quad (22.178c)$$

As before, the first two equations come from the diagonal components of the Usadel equation, and describe the equilibrium properties of the system, while the third equation comes from

the off-diagonal or Keldysh component, and represents the kinetic equation for the distribution function. These equations are supplemented by the equations for measurable quantities corresponding to Eqns.(22.137), (22.138) and (22.137). Expanding $\tilde{g}^K$ in Eqn.(22.137) using Eqns.(22.157) and Eqns.(22.137), we have

$$\vec{j}(\vec{R}, T) = -\frac{eN_0}{4} \int dE \int \frac{d\Omega_p}{4\pi} v_F \hat{p} \text{Tr} [\tau^3 (\tilde{g}_s^K + \hat{p} \tilde{g}_p^K)] = \frac{eN_0 D}{4} \int dE \text{Tr} [\tau^3 (\tilde{g}_s \partial_{\vec{R}} \tilde{g}_s)^K], \quad (22.179)$$

where the angular average over the p -wave component give a factor of $(1/3)$, and we have replaced $(1/3)v_F \ell$ by the diffusion coefficient D . Now

$$(\tilde{g}_s \partial_{\vec{R}} \tilde{g}_s)^K = (\tilde{g}_s^R \partial_{\vec{R}} \tilde{g}_s^K) + (\tilde{g}_s^K \partial_{\vec{R}} \tilde{g}_s^A) \quad (22.180)$$

so that Eqn.(22.180) above can be rewritten as

$$\vec{j}(\vec{R}, T) = \frac{eN_0 D}{4} \int dE \text{Tr} [\tau^3 (\tilde{g}_s^R \partial_{\vec{R}} \tilde{g}_s^K + \tilde{g}_s^K \partial_{\vec{R}} \tilde{g}_s^A)]. \quad (22.181)$$

The equation for j_{th} is obtained in a similar way

$$\vec{j}_{th}(\vec{R}, T) = \frac{N_0 D}{4} \int dE E \text{Tr} [\tilde{g}_s^R \partial_{\vec{R}} \tilde{g}_s^K + \tilde{g}_s^K \partial_{\vec{R}} \tilde{g}_s^A], \quad (22.182)$$

The kinetic equation, Eqn.(22.178c), can be recast in terms of a differential equation for a distribution function $\tilde{h}$, in the hopes of separating the equilibrium properties of the system, represented by $\tilde{g}_s^R$ and $\tilde{g}_s^A$, from the non-equilibrium properties of the system, represented by $\tilde{h}$. To this end, as before, we substitute Eqn.(22.143) for $\tilde{g}_s^K$ in Eqn.(22.178c), which then becomes

$$[\tau^3 E + \tilde{\Delta}, \tilde{g}_s^R \tilde{h}] - [\tau^3 E + \tilde{\Delta}, \tilde{h} \tilde{g}_s^A] = iD \partial_{\vec{R}} \left[(\tilde{g}_s^R (\partial_{\vec{R}} \tilde{g}_s^R) \tilde{h} - \tilde{h} \tilde{g}_s^A (\partial_{\vec{R}} \tilde{g}_s^A)) + \partial_{\vec{R}} \tilde{h} - (\tilde{g}_s^R \partial_{\vec{R}} \tilde{h} \tilde{g}_s^A) \right]. \quad (22.183)$$

Taking the trace of both sides of the equation above gives

$$0 = iD \partial_{\vec{R}} \text{Tr} \left\{ \partial_{\vec{R}} \tilde{h} + \tilde{g}_s^R (\partial_{\vec{R}} \tilde{g}_s^R) \tilde{h} - \tilde{h} \tilde{g}_s^A (\partial_{\vec{R}} \tilde{g}_s^A) - \tilde{g}_s^R (\partial_{\vec{R}} \tilde{h}) \tilde{g}_s^A \right\}, \quad (22.184)$$

using Eqns.(22.178a) and (22.178b). In the linear regime, we can use the diagonal representation of $\tilde{h}$ given by Eqn.(22.146) to obtain

$$D \partial_{\vec{R}} \left[(\partial_{\vec{R}} h_L) \text{Tr} \left\{ 1 - \tilde{g}_s^R \tilde{g}_s^A \right\} + h_T \text{Tr} \left\{ \tau^3 (\tilde{g}_s^R (\partial_{\vec{R}} \tilde{g}_s^R) - \tilde{g}_s^A (\partial_{\vec{R}} \tilde{g}_s^A)) \right\} - (\partial_{\vec{R}} h_T) \text{Tr} \left\{ \tilde{g}_s^R \tau^3 \tilde{g}_s^A \right\} \right] = 0, \quad (22.185)$$

where we have used the fact that

$$\text{Tr} \{ \partial_{\vec{R}} (\tilde{g}_s^R \tilde{g}_s^R) \} = \text{Tr} \{ (\partial_{\vec{R}} (\tilde{g}_s^R) \tilde{g}_s^R + \tilde{g}_s^R (\partial_{\vec{R}} (\tilde{g}_s^R)) \} = 2 \text{Tr} \{ \tilde{g}_s^R (\partial_{\vec{R}} (\tilde{g}_s^R)) \} = 0, \quad (22.186)$$

Taking the trace after multiplying both sides of Eqn.(22.183) by τ^3 gives

$$\begin{aligned} D \partial_{\vec{R}} \left[\partial_{\vec{R}} h_T \text{Tr} \{ 1 - \tilde{g}_s^R \tau^3 \tilde{g}_s^A \tau^3 \} + h_L \text{Tr} \{ \tau^3 (\tilde{g}_s^R (\partial_{\vec{R}} \tilde{g}_s^R) - \tilde{g}_s^A (\partial_{\vec{R}} \tilde{g}_s^A)) \} - \partial_{\vec{R}} h_L \text{Tr} \{ \tilde{g}_s^R \tilde{g}_s^A \tau^3 \} \right] \\ = i \left[h_L \text{Tr} \{ \tau^3 [\tilde{g}_s^R - \tilde{g}_s^A, \tilde{\Delta}] \} - 2 h_T \text{Tr} \{ \tilde{\Delta} (\tilde{g}_s^R + \tilde{g}_s^A) \} \right]. \end{aligned} \quad (22.187)$$

Equations (22.185) and (22.187) form a set of coupled differential equations for the distribution functions h_L and h_T . Let us define the quantities

$$Q = \frac{1}{4} \text{Tr} \{ \tau^3 (\tilde{g}_s^R (\partial_{\vec{R}} \tilde{g}_s^R) - \tilde{g}_s^A (\partial_{\vec{R}} \tilde{g}_s^A)) \} \quad (22.188)$$

and

$$M_{ij} = \frac{1}{4} \text{Tr} \{ \delta_{ij} \tau^0 - \tilde{g}_s^R \tau^i \tilde{g}_s^A \tau^j \}. \quad (22.189)$$

Then Eqns.(22.185) and (22.187) can be written in the form

$$\partial_{\vec{R}} [M_{00}(\partial_{\vec{R}} h_L) + Q h_T + M_{30} \partial_{\vec{R}} h_T] = 0, \quad (22.190a)$$

$$\partial_{\vec{R}} [M_{33}(\partial_{\vec{R}} h_T) + Q h_L + M_{03} \partial_{\vec{R}} h_L] = \frac{i}{4D} \left[h_L \text{Tr} \{ \tau^3 [\tilde{g}_s^R - \tilde{g}_s^A, \tilde{\Delta}] \} - 2 h_T \text{Tr} \{ \tilde{\Delta} (\tilde{g}_s^R + \tilde{g}_s^A) \} \right]. \quad (22.190b)$$

These equations are in the form of diffusion equations for the distribution function, more general forms of the diffusion equation discussed for the normal case in the introduction. As we shall see, the quantity Q is related to the *spectral* supercurrent in the system, DM_{ij} are now the energy and position dependent diffusion coefficients, and Eqns.(22.190) are essentially continuity equations for the spectral thermal and electric current. In the normal limit, $g^R = \tau^3$, $g^A = -\tau^3$, and $\tilde{\Delta} = 0$, so that $M_{00} = M_{33} = 1$, and $Q = M_{03} = M_{30} = 0$. Equations (22.190) then reduce to Eqn.(22.1), as expected. Noting that the term in square brackets in Eqn.(22.178c) is the same as the term in parenthesis in Eqn.(22.181) and the term in square brackets in Eqn.(22.182), the electric current can be written as

$$\vec{j}(\vec{R}, T) = e N_0 D \int dE (M_{33}(\partial_{\vec{R}} h_T) + Q h_L + M_{03} \partial_{\vec{R}} h_L). \quad (22.191)$$

The first term corresponds to quasiparticle (or dissipative) current, and the second term to the supercurrent. The third term, which is proportional to the derivative of h_L , is associated with an imbalance between particles and holes. The thermal current can be written in a similar way

$$\vec{j}_{th}(\vec{R}, T) = N_0 D \int dE E [M_{00}(\partial_{\vec{R}} h_L) + Q h_T + M_{30} \partial_{\vec{R}} h_T]. \quad (22.192)$$

For the charge-imbalance Q^* , we note that only the s -wave part of the Keldysh Green's function in the square brackets in Eqn.(22.139) survives after angular averaging. Writing g_s^K in the form given by Eqn.(22.143), with $\tilde{h}$ given by Eqn.(22.146), we have

$$Q^* = -\frac{e N_0}{4} \int dE h_L \text{Tr} \{ g_s^R - g_s^A \} - \frac{e N_0}{4} \int dE h_T \text{Tr} \{ \tau^3 (g_s^R - g_s^A) \} = -e N_0 \int dE h_T N(E), \quad (22.193)$$

since g_s^R and g_s^A are traceless. Here we have defined the normalized superconducting density of states by

$$N(E) = \frac{1}{4} \text{Tr} \{ \tau^3 (g_s^R - g_s^A) \}, \quad (22.194)$$

which reduces to the conventional BCS density of states in the equilibrium case.

We can also recast the kinetic equations Eqn.(22.190) in a slightly different form sometimes used in the literature. To do this, we subtract Eqn.(22.178b) from Eqn.(22.178a), multiply by τ^3 , and take the trace. The result is

$$\text{Tr} \{ \tau^3 [g_s^R - g_s^A, \tilde{\Delta}] \} = -i D \partial_{\vec{R}} (\text{Tr} \{ \tau^3 [\tilde{g}_s^R \partial_{\vec{R}} \tilde{g}_s^R - \tilde{g}_s^A \partial_{\vec{R}} \tilde{g}_s^A] \}). \quad (22.195)$$

Using the definition of Q , we have

$$\partial_{\vec{R}} Q = \frac{i}{4D} \text{Tr} \{ \tau^3 [g_s^R - g_s^A, \tilde{\Delta}] \}. \quad (22.196)$$

If we choose a gauge in which Δ is real, the right hand side of the equation vanishes, so that $\partial_{\vec{R}}Q = 0$. Clearly, $\partial_{\vec{R}}Q = 0$ also for a normal metal (even a proximity-coupled normal metal), where Δ vanishes. In either case, the spectral supercurrent Q is conserved. The right hand side of the equation above multiplied by h_L is the same as the second term of Eqn.(22.190b). Subtracting Eqn.(22.195) multiplied by h_L from Eqn.(22.190b), we obtain

$$\partial_{\vec{R}}[M_{33}(\partial_{\vec{R}}h_T) + M_{03}(\partial_{\vec{R}}h_L)] + Q(\partial_{\vec{R}}h_L) = -\frac{i}{4D} [2h_T \text{Tr}\{\tilde{\Delta}(\tilde{g}_s^R + \tilde{g}_s^A)\}]. \quad (22.197)$$

For real Δ , when $\partial_{\vec{R}}Q = 0$, Eqn.(22.190a) can be written in a similar manner

$$\partial_{\vec{R}}[M_{00}(\partial_{\vec{R}}h_L) + M_{30}(\partial_{\vec{R}}h_T)] + Q(\partial_{\vec{R}}h_T) = 0 \quad (22.198)$$

Finally, we can also write the kinetic equations in a form similar to Eqn.(22.150)

$$\tilde{g}_s^R B[\tilde{h}] - B[\tilde{h}] \tilde{g}_s^A = 0, \quad (22.199)$$

by performing similar manipulations on Eqns.(22.178) as were performed on Eqns.(22.148). For the diffusive case, $B[\tilde{h}]$ is a functional of $\tilde{h}$ now given by

$$B[\tilde{h}] = [\tau^3 E + \tilde{\Delta}, \tilde{h}] - iD \left[(\partial_{\vec{R}}\tilde{g}_s^R)(\partial_{\vec{R}}\tilde{h}) + \frac{1}{2} (\tilde{g}_s^R(\partial_{\vec{R}}^2\tilde{h}) - (\partial_{\vec{R}}^2\tilde{h})\tilde{g}_s^A) - (\partial_{\vec{R}}\tilde{h})(\partial_{\vec{R}}\tilde{g}_s^A) \right]. \quad (22.200)$$

Boundary conditions for the quasiclassical equations of motion

Eqns.(22.178) or their derivatives form a set of coupled differential equations for the quasiclassical Green's functions and the distribution function. In order to obtain a solution, however, we need to specify the boundary conditions for the Green's functions and the distribution function. To this end, we define the concept of a *reservoir*, where the Green's functions and distribution function have well-defined values.

For a normal reservoir, the retarded and advanced quasiclassical Green's functions are given by

$$g_{N0}^R = \tau^3; \quad g_{N0}^A = -\tau^3, \quad (22.201)$$

and for the superconducting case, by Eqn.(22.155), which we reproduce here

$$\tilde{g}^R = \begin{pmatrix} \frac{E}{\sqrt{E^2 - |\Delta|^2}} & \frac{\Delta}{\sqrt{E^2 - |\Delta|^2}} \\ -\frac{\Delta^*}{\sqrt{E^2 - |\Delta|^2}} & -\frac{E}{\sqrt{E^2 - |\Delta|^2}} \end{pmatrix}. \quad (22.202)$$

with g_{S0}^A given by Eqn.(22.156).

The equilibrium distribution function $\tilde{h}$ is given by Eqn.(22.121) in both normal and superconducting reservoirs (since we are dealing only with excitations), where, as we noted earlier the (1,1) component of the matrix applies to particle-like excitations, and the (2,2) component of the matrix applies to hole-like excitations. From this point of view, the Fermi functions in Eqn.(22.121) are given in terms of the usual equilibrium Fermi function Eqn.(22.5) by $f_{0\uparrow}(E) = f_0(E)$ and $f_{0\downarrow}(E) = f_0(-E)$.¹⁶ Looking ahead to where we might have a finite voltage V on a reservoir, the equilibrium form of $\tilde{h}$ can then be written as

$$\tilde{h}_0 = \begin{pmatrix} \tanh\left(\frac{E+eV}{2k_B T}\right) & 0 \\ 0 & \tanh\left(\frac{E-eV}{2k_B T}\right) \end{pmatrix} \quad (22.203)$$

¹⁶Since we are dealing with the static limit in all that follows, we again use the symbol T to refer to the temperature

If we write $\tilde{h}_0$ in the form of Eqn.(22.146) the equilibrium values of h_L and h_T can then be expressed as

$$h_{L,T} = \frac{1}{2} \left[\tanh \left(\frac{E + eV}{2k_B T} \right) \pm \tanh \left(\frac{E - eV}{2k_B T} \right) \right] \quad (22.204)$$

If a finite voltage is put on the superconductor, we will obtain a time evolution of the phase in accordance with the Josephson relations. Since we have restricted ourselves here to the static case, we must assume that the voltage on the superconducting reservoir $V = 0$. In this case, $h_T = 0$ for the superconducting reservoir.

For a system with perfect interfaces between the superconducting and normal parts, the boundary conditions defined by the equations above are sufficient to solve the differential equations, the implicit assumption being that the Green's functions and the distribution functions are continuous across any interface. When the transparency of the interface is less than unity, this is no longer true. Zaitsev[30] derived the boundary conditions for the Green's functions at an interface of arbitrary transparency. Kupriyanov and Lukichev[31] simplified these equations for the diffusive case in the limit of small barrier transparency. Consider then an interface at $x = 0$ between two metals, say one a normal metal in the half-plane $x < 0$ (labeled by the index '1') and one a superconductor in the half-plane $x > 0$, (labeled by the index '2'), although it also could be an interface between two different normal metals or superconductors. The boundary conditions of Kupriyanov and Lukichev are then

$$v_{F1} D_1 \hat{g}_{s1} (\partial_x \hat{g}_{s1}) = v_{F2} D_2 \hat{g}_{s2} (\partial_x \hat{g}_{s2}), \quad (22.205a)$$

$$\hat{g}_{s1} \partial_x \hat{g}_{s1} = \frac{1}{2r} [\hat{g}_{s1}, \hat{g}_{s2}]. \quad (22.205b)$$

Here ∂_x denotes a derivative in the positive x direction, and $r = R_b/R_N$ is a parameter that is nominally the ratio of the barrier resistance R_b to the normal wire resistance per unit length R_N/L , but is inversely proportional to the transmission t of the interface. The first of the equations is clearly related to the conservation of current across the interface. The right hand side of the second equation has been shown to be the first term in an expansion of terms of the transmission coefficient t , [14] and hence, it is only valid for low t .

The diagonal part of Eqn.(22.205b) gives the boundary condition for the Green's functions

$$\tilde{g}_{s1}^R \partial_x \tilde{g}_{s1}^R = \frac{1}{2r} [\tilde{g}_{s1}^R \tilde{g}_{s2}^R - \tilde{g}_{s2}^R \tilde{g}_{s1}^R] \quad (22.206)$$

The off-diagonal part of Eqn.(22.205b) gives the boundary condition for the distribution function

$$\tilde{g}_{s1}^R \partial_x \tilde{g}_{s1}^K + \tilde{g}_{s1}^K \partial_x \tilde{g}_{s1}^A = \frac{1}{2r} [\tilde{g}_{s1}^R \tilde{g}_{s2}^K + \tilde{g}_{s1}^K \tilde{g}_{s2}^A - \tilde{g}_{s2}^R \tilde{g}_{s1}^K - \tilde{g}_{s2}^K \tilde{g}_{s1}^A] \quad (22.207)$$

If we put in Eqn.(22.143) for $\tilde{g}_s^K$, with Eqn.(22.146) for $\tilde{h}$, and then take as before the trace of the resulting equation, and the trace of the equation multiplied by τ^3 , we will obtain boundary conditions for h_T and h_L . Noting that the left hand side of Eqn.(22.207) is simply the term in square brackets on the right hand side of Eqn.(22.178c), we obtain the two equations

$$2r [M_{00}(\partial_{\bar{R}} h_L) + Q h_T + M_{30} \partial_{\bar{R}} h_T] = \alpha_1, \quad (22.208a)$$

and

$$2r [M_{33}(\partial_{\bar{R}} h_T) + Q h_L + M_{03} \partial_{\bar{R}} h_L] = \alpha_2, \quad (22.208b)$$

where

$$\alpha_1 = \text{Tr} \{ \tilde{g}_1^R (\tilde{g}_2^R \tilde{h}_2 - \tilde{h}_2 \tilde{g}_2^A) + (\tilde{g}_1^R \tilde{h}_1 - \tilde{h}_1 \tilde{g}_1^A) \tilde{g}_2^A - \tilde{g}_2^R (\tilde{g}_1^R \tilde{h}_1 - \tilde{h}_1 \tilde{g}_1^A) - (\tilde{g}_2^R \tilde{h}_2 - \tilde{h}_2 \tilde{g}_2^A) \tilde{g}_1^A \}, \quad (22.208c)$$

and

$$\alpha_2 = \text{Tr} \{ (\tau^3 \tilde{g}_1^R (\tilde{g}_2^R \tilde{h}_2 - \tilde{h}_2 \tilde{g}_2^A) + (\tilde{g}_1^R \tilde{h}_1 - \tilde{h}_1 \tilde{g}_1^A) \tilde{g}_2^A - \tilde{g}_2^R (\tilde{g}_1^R \tilde{h}_1 - \tilde{h}_1 \tilde{g}_1^A) - (\tilde{g}_2^R \tilde{h}_2 - \tilde{h}_2 \tilde{g}_2^A) \tilde{g}_1^A) \}. \quad (22.208d)$$

$$(22.208e)$$

We note that although the boundary conditions of Kupriyanov and Lukichev are valid for large r , they also give the correct boundary conditions (that of continuity of the Green's functions and distribution functions) in the limit of $r \rightarrow 0$. For arbitrary transmission of a barrier with n channels, the boundary condition can be represented as^[19]

$$\hat{g}_{s1} \partial_x \hat{g}_{s1} = \gamma \frac{e^2}{\pi} \sum_n 2T_n \frac{[\hat{g}_{s1}, \hat{g}_{s2}]}{4 + T_n(\hat{g}_{s1} \hat{g}_{s2} + \hat{g}_{s2} \hat{g}_{s1} - 2)}, \quad (22.209)$$

where γ is a constant factor, and T_n is the transmission of the n th channel. It can be seen that for a single channel with small transmission T , this equation reduces to the boundary condition of Kupriyanov and Lukichev.

22.7 Parametrization of the quasiclassical Green's function

The normalization Eqn.(22.167a) permits a parametrization of the quasiclassical Green's functions that is very convenient for calculations. Eqn.(22.167a) is a matrix equation that is equivalent to the three equations (22.142) for the s -wave component of the Green's function. To take into account the macroscopic phase of the superconductor, we note that a gauge transformation that transforms the vector and scalar potentials according to Eqns.(22.93) and (22.94) transforms the field operators according to the equations

$$\hat{\psi} \rightarrow \hat{\psi} e^{i\chi} \quad (22.210a)$$

$$\hat{\psi}^+ \rightarrow \hat{\psi}^+ e^{-i\chi} \quad (22.210b)$$

The Nambu-Gorkov Green's functions defined in Eqns.(22.109) are transformed accordingly. For example, two components of $\tilde{G}_{12}^{\beta\alpha}$ would transform according to

$$[\tilde{G}_{12}^{\beta\alpha}]_{11} \rightarrow [\tilde{G}_{12}^{\beta\alpha}]_{11} e^{-i(\chi(\vec{r}_1) - \chi(\vec{r}_2))} \quad (22.211a)$$

$$[\tilde{G}_{12}^{\beta\alpha}]_{12} \rightarrow [\tilde{G}_{12}^{\beta\alpha}]_{12} e^{i(\chi(\vec{r}_1) + \chi(\vec{r}_2))} \quad (22.211b)$$

Making the transformation (22.42a) to mixed coordinates, and taking the limit as $\vec{r} \rightarrow 0$, we see that the off-diagonal components of the Nambu-Gorkov Green's functions are multiplied by a phase factor $e^{i\chi(\vec{R})}$ or $e^{-i\chi(\vec{R})}$, while the diagonal components remain unchanged. Consequently, Δ also transforms as

$$\Delta \rightarrow \Delta e^{i\chi(\vec{R})} \quad (22.212)$$

Keeping this in mind, we can express $\tilde{g}_s^R$ as

$$\tilde{g}_s^R = \begin{pmatrix} \cosh \theta & \sinh \theta e^{i\chi} \\ -\sinh \theta e^{-i\chi} & -\cosh \theta \end{pmatrix} \quad (22.213)$$

where θ and χ are complex functions of the energy E and position $\vec{R}$. This form satisfies the normalization condition $\hat{g}_s^R \hat{g}_s^R = \tau^0$.¹⁷ For completeness, we also give the expression for $\tilde{g}_s^A$

$$\tilde{g}_s^A = \begin{pmatrix} -\cosh \theta^* & -\sinh \theta^* e^{i\chi^*} \\ \sinh \theta^* e^{-i\chi^*} & \cosh \theta^* \end{pmatrix} \quad (22.214)$$

¹⁷One can also express $\hat{g}_s^R$ equivalently in terms of sines and cosines.

We now put this into the Usadel equation for $\tilde{g}^R$, Eqn.(22.178a). Keeping in mind that the matrix for $\tilde{\Delta}$ involves additional factors of $e^{i\chi}$ and $e^{-i\chi}$ due to the gauge transformation, the diagonal (1,1) component of this matrix equation is

$$D \sinh^2 \theta \partial_{\vec{R}}^2 \chi + D \sinh 2\theta \partial_{\vec{R}} \chi \partial_{\vec{R}} \theta - 2i \text{Im}(\Delta) \sinh \theta = 0, \quad (22.215a)$$

and the off-diagonal component (1,2) is

$$D \partial_{\vec{R}}^2 \theta - \frac{D}{2} \sinh 2\theta (\partial_{\vec{R}} \chi)^2 + 2Ei \sinh \theta - 2i \text{Re}(\Delta) \cosh \theta = 0, \quad (22.215b)$$

where we have used Eqn.(22.215a) to simplify Eqn.(22.215b). Defining a current $j_s(E, \vec{R})$ by the equation

$$j_s(E, \vec{R}) = \sinh^2 \theta(E, \vec{R}) \partial_{\vec{R}} \chi(E, \vec{R}), \quad (22.216)$$

we can rewrite Eqn.(22.215a) as

$$D \partial_{\vec{R}} j_s(E, \vec{R}) - 2i \text{Im}(\Delta) \sinh \theta = 0. \quad (22.217)$$

$j_s(E, \vec{R})$ is proportional to $\sinh^2 \theta$, which is proportional to the square of the pair amplitude, and it is also proportional to the gauge-invariant gradient of the phase. Consequently, it is similar in form to the conventional definition of the supercurrent, and is called the spectral supercurrent. Indeed, Eqn.(22.217) is simply another way of writing Eqn.(22.196), and Q and j_s are related by

$$Q(E, \vec{R}) = -\text{Im}(j_s(E, \vec{R})). \quad (22.218)$$

As we noted before, $\partial_{\vec{R}} Q = 0$ if Δ is purely real, and from Eqn.(22.217), it can be seen that $\partial_{\vec{R}} j_s(E, \vec{R}) = 0$ also if Δ is real.

Equations(22.215) form a set of coupled equations that can be solved in principle for $\theta(E, \vec{R})$ and $\chi(E, \vec{R})$. In the case of a negligible spectral supercurrent, the equations decouple, and one needs to solve only Eqn.(22.215b). In the limit of a bulk superconductor with a uniform real order parameter and no phase gradient, we recover the bulk value of the Green's function, Eqn.(22.155). The differential equations must be supplemented by boundary conditions. From Eqn.(22.155), in a superconducting reservoir, we have

$$\cosh \theta_{S0} = \frac{E}{\sqrt{E^2 - |\Delta|^2}}, \quad (22.219)$$

so that the value of θ in the superconducting reservoir is given by

$$\theta_{S0} = \begin{cases} -i\frac{\pi}{2} + \frac{1}{2} \ln \frac{|\Delta|+E}{|\Delta|-E} & \text{if } E < |\Delta|, \\ \frac{1}{2} \ln \frac{E+|\Delta|}{E-|\Delta|} & \text{if } E > |\Delta|. \end{cases} \quad (22.220)$$

The value of χ in a superconducting reservoir is just the macroscopic phase of the superconductor. In a normal reservoir, $\theta = 0$. The value of χ in a normal reservoir is meaningless, of course, and any choice that results in no phase gradient is valid.

In terms of the θ parametrization, the boundary conditions of Kupriyanov and Lukichev can be expressed as

$$r \sinh \theta_1 (\partial_{\vec{R}} \chi_1) = \sinh \theta_2 \sin(\chi_2 - \chi_1), \quad (22.221a)$$

and

$$r [\partial_{\vec{R}} \theta_1 + i \sinh \theta_1 \cosh \theta_1 (\partial_{\vec{R}} \chi_1)] = \cosh \theta_1 \sinh \theta_2 e^{i(\chi_2 - \chi_1)} - \sinh \theta_1 \cosh \theta_2. \quad (22.221b)$$

Note that for $r = 0$, the boundary conditions reduce to $\chi_1 = \chi_2$ and $\theta_1 = \theta_2$. In the absence of a supercurrent, $(\partial_{\vec{R}}\chi_1) = 0$, so that the equations above simplify to

$$\chi_1 = \chi_2, \quad (22.222a)$$

and

$$r(\partial_{\vec{R}}\theta_1) = \sinh(\theta_2 - \theta_1). \quad (22.222b)$$

Finally, we can write expressions for physical quantities in terms of θ and χ . These quantities can be written in terms of the M_{ij}

$$M_{00} = \frac{1}{2} [1 + \cosh \theta \cosh \theta^* - \sinh \theta \sinh \theta^* \cosh(2\text{Im}(\chi))], \quad (22.223a)$$

$$M_{33} = \frac{1}{2} [1 + \cosh \theta \cosh \theta^* + \sinh \theta \sinh \theta^* \cosh(2\text{Im}(\chi))], \quad (22.223b)$$

$$M_{03} = \frac{1}{2} \sinh \theta \sinh \theta^* \sinh(2\text{Im}(\chi)), \quad (22.223c)$$

and

$$M_{30} = -\frac{1}{2} \sinh \theta \sinh \theta^* \sinh(2\text{Im}(\chi)). \quad (22.223d)$$

If χ is real, then the $M_{03} = M_{30} = 0$, and M_{00} and M_{33} simplify to

$$M_{00} = \cos^2(\text{Im}(\theta)) \quad (22.224a)$$

and

$$M_{33} = \cosh^2(\text{Re}(\theta)) \quad (22.224b)$$

These relations will be used in the next section when we discuss applying the Usadel equation to derive the transport properties of some simple device geometries.

To conclude this section, we shall write an expression for the gap in terms θ . Replacing Eqn.(22.143) for $\tilde{g}^K$ into Eqn.(22.136) for the gap, we have

$$\Delta = N_0 \frac{\lambda}{4} \int dE [\tilde{g}_s^R \tilde{h} - \tilde{h} \tilde{g}_s^A]_{12}, \quad (22.225)$$

where performing the angular average in Eqn.(22.136) gives the s -components of the Green's functions. Putting in $\tilde{h}$ in the form of Eqn.(22.146), we obtain

$$\Delta = N_0 \frac{\lambda}{4} \int dE [h_L(\tilde{g}_s^R - \tilde{g}_s^A) + h_T(\tilde{g}_s^R \tau^3 - \tau^3 \tilde{g}_s^A)]_{12} \quad (22.226)$$

With $\tilde{g}_s^R$ and $\tilde{g}_s^A$ given by Eqns.(22.213) and (22.214), we obtain

$$\Delta = N_0 \frac{\lambda}{4} \int dE [h_L(\sinh \theta e^{i\chi} + \sinh \theta^* e^{i\chi^*}) - h_T(\sinh \theta e^{i\chi} - \sinh \theta^* e^{i\chi^*})]. \quad (22.227)$$

As an example, consider the case of a bulk superconductor, where $\chi = 0$ and $h_T = 0$. We then have

$$\begin{aligned} \Delta &= N_0 \frac{\lambda}{2} \int dE h_L \text{Re}(\sinh \theta) \\ &= N_0 \lambda \int_0^\infty dE \tanh(E/2k_B T) \text{Re} \left(\frac{\Delta}{\sqrt{E^2 - \Delta^2}} \right), \end{aligned} \quad (22.228)$$

which is the usual self-consistent equation for the gap.

22.8 Applications of the quasiclassical equations to proximity-coupled systems

We shall conclude our discussion of the quasiclassical theory by applying the equations that we have derived to some simple devices incorporating normal metals in close proximity with superconductors. Since the equations of motion in the diffusive limit are in general nonlinear, solving them usually requires numerical techniques, except in the limit of large resistances between the superconductor and the normal metal, where the Usadel equation can be linearized. We shall restrict ourselves to one-dimensional examples; these are the ones discussed most in the literature.

22.8.1 Proximity-coupled wire

We start with the simplest possible device, a one-dimensional normal-metal wire of length L connected on one end to a superconducting reservoir and at the other end to a normal metal reservoir. For definitiveness, let us take the superconducting reservoir to be at $x = 0$ and the normal-metal reservoir at $x = L$, and let us consider first the case of a perfect SN interface, so that the interface barrier resistance parameter $r = 0$. In this geometry, there can be no supercurrent, so that Q (or alternatively, j_s) is zero. Furthermore, we can take the phase χ to be zero in the superconducting reservoir without loss of generality, and we note again that $\Delta = 0$ in the normal metal. Under these conditions, only M_{00} and M_{33} are non-zero in Eqns.(22.190), which now read

$$M_{00}(\partial_{\vec{R}} h_L) = K_1, \quad (22.229a)$$

and

$$M_{33}(\partial_{\vec{R}} h_T) = K_2, \quad (22.229b)$$

where K_1 and K_2 are constants of integration. On integrating these equations from $x = 0$ to $x = L$, we obtain

$$h_L(x = L) - h_L(x = 0) = K_1 \int_0^L \frac{1}{M_{00}} dx, \quad (22.230a)$$

$$h_T(x = L) - h_T(x = 0) = K_2 \int_0^L \frac{1}{M_{33}} dx. \quad (22.230b)$$

To calculate the conductance of the normal-metal wire in the linear approximation, we apply a small voltage V on the normal-metal reservoir, keeping the superconducting reservoir at $V = 0$. If we consider the second equation, $h_T(x = 0) = 0$, and expanding $h_T(x = L)$ in a Taylor's expansion to first order, we obtain from Eqn.(22.230b)

$$K_2 = \frac{eV}{2k_B T \cosh^2\left(\frac{E}{2k_B T}\right)} \left[\int_0^L \frac{1}{M_{33}(E, x)} dx \right]^{-1}. \quad (22.231)$$

The electric current in the linear response regime can then be obtained from Eqn.(22.191)

$$j = \frac{N_0 e^2 V D}{2k_B T} \int dE \frac{1}{\cosh^2(E/2k_B T)} \left[\int_0^L \frac{1}{M_{33}(E, x)} dx \right]^{-1} \quad (22.232)$$

There are two differences between this equation and the equivalent equation for a normal metal in the classical regime (Eqn.(22.10a)) that we derived earlier. First, the equation above involves

an integral over energy and position. One can define a energy and position dependent electrical diffusion coefficient

$$D_3(E, x) = DM_{33}(E, x) \quad (22.233)$$

instead of the constant diffusion coefficient D for Eqn(22.10a). Second, the current in Eqn.(22.232) does not involve temperature differentials. Indeed, if one assumes that the superconducting reservoir is at a temperature T , and the normal reservoir at a temperature $T + \Delta T$, and expand $h_T(x = L)$ to first order in ΔT , the terms involving ΔT cancel, so that there is no term proportional to ΔT . As a consequence, thermoelectric phenomena cannot be described using the quasiclassical equations, and an extension to the theory is required to take them into account.¹⁸

From Eqn.(22.232), one can define a spectral or energy dependent conductance of the wire

$$G(E) = G_N \left[\int_0^L \frac{1}{M_{33}(E, x)} dx \right]^{-1}, \quad (22.234)$$

where G_N is the normal state conductance of the wire. The total conductance is then

$$G = \int dE \frac{G(E)}{2k_B T \cosh^2(E/2k_B T)} \quad (22.235)$$

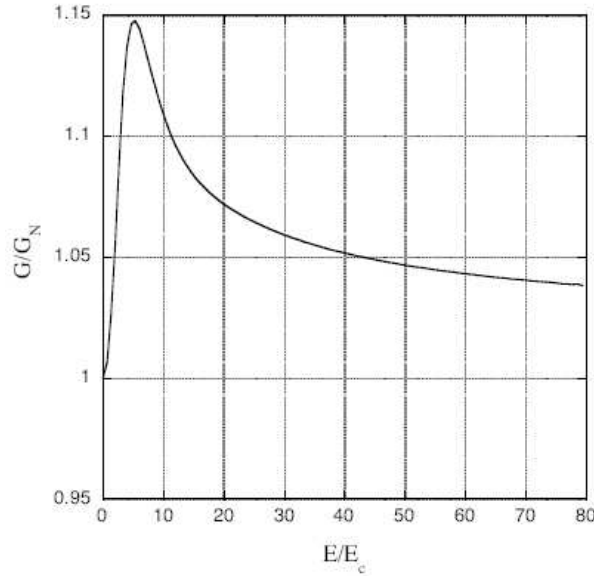


Figure 98: Spectral conductance $G(E)$ of a one-dimensional wire of length L as a function of energy E , normalized to the correlation energy E_c . The barrier interface parameter $r=0$, and the gap is set to be $\Delta = 1000E_c$.

Figure 4 shows the results of a numerical calculation of $G(E)/G_N$ as a function of the normalized energy E/E_c . The normalization factor $E_c = D/L^2$ is called the correlation energy or Thouless energy (from the theory of disordered metals, where it also occurs),^[32] and is dependent on the length L of the wire. At high energies, $G(E)$ approaches its normal state value, as expected. As the energy is lowered, the conductance increases, as might be expected in a proximity-coupled normal metal. However, instead of continually increasing as the energy is reduced, it reaches a maximum of around $1.15 G_N$ at an energy of $E \simeq 5E_c$, and then decreases, reaching its normal state value at $E = 0$. This nonmonotonic behavior of the conductance is

¹⁸See, for example, F. Wilhelm, PhD Thesis, Universität Karlsruhe, 2000

called reentrance, and has been observed in experiments by a number of groups.[33] It should be emphasized that the relevant energy scale where the maximum in conductance is observed is set not by the gap Δ of the superconductor, but by E_c , which itself depends inversely on the square of the length of the sample L . Hence in very long or macroscopic samples, the energy (and correspondingly, the temperature) at which the minimum would occur is far below the experimentally accessible range, and one regains the monotonic behavior expected from the simple Ginzburg-Landau theory of de Gennes.[34]

The temperature dependent conductance $G(T)$ can be obtained from $G(E)$ using Eqn.(22.235); the result of this calculation, plotted in terms of the normalized resistance $R(T)/R_n$, is shown in Fig. 5. In obtaining this plot, we have used a value of $\Delta = 32E_c$, corresponding to a weak-coupling transition temperature of $T_c = 1.764\Delta/k_B$, typical parameters for Al films. We have also assumed that the gap is temperature dependent. Like $G(E)$, $R(T)$ is also non-monotonic, with a minimum at some intermediate temperature. We would expect the minimum in $R(T)$ to be around $T \simeq 5E_c/k_B$, based on the behavior of $G(E)$. However, the temperature dependence of the superconducting gap modifies this behavior, so that the minimum in resistance occurs at a somewhat higher temperature when the interface between the normal metal and the superconductor is perfect. Figure 5 shows additional curves corresponding to progressively increasing values of the interface barrier parameter r . Increasing the resistance of the NS interface decreases the leakage of superconducting correlations from the superconductor into the normal metal, and consequently results in a smaller increase in the conductance of the proximity-coupled normal metal. In addition, the temperature T_{min} at which the minimum in resistance occurs is also shifted down.

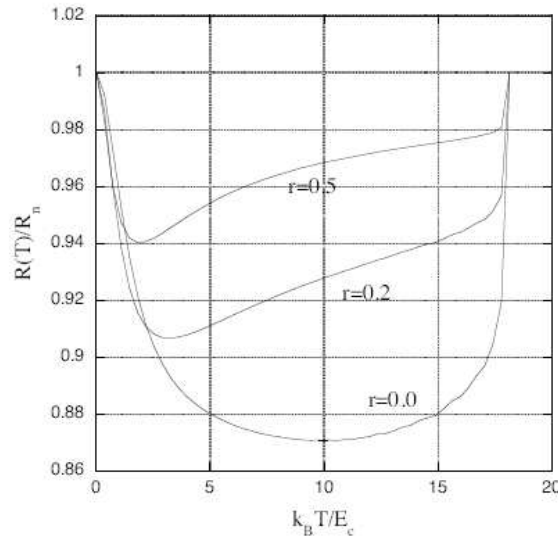


Figure 99: Temperature dependent resistance $R(T)$ normalized to the normal state resistance R_N , as a function of the temperature T normalized to E_c , for different values of the interface resistance parameter r . The gap is set to be $\Delta = 32E_c$.

From Eqn.(22.194), the normalized density of states $N(E)$ can be expressed in terms of θ as

$$N(E) = \cosh(\text{Re}(\theta)) \cos(\text{Im}(\theta)). \quad (22.236)$$

Figure 6 shows the density of states as a function of energy and position along the wire of length L . There is a proximity-induced decrease of $N(E)$ near the superconducting reservoir. In fact, at the NS interface, there is a divergence in $N(E)$ at the gap energy, and it goes to zero at $E = 0$, just as one would expect for a superconductor. However, unlike a superconductor, it is not strictly zero for $E < \Delta$, but still has a finite amplitude. As one moves away from the NS

interface into the proximity-coupled normal wire, both the amplitude of this effective gap and the divergence are smoothly reduced, so that at the normal reservoir, one recovers the normal density of states. This position dependent variation of the density of states has been observed in experiments.[35]

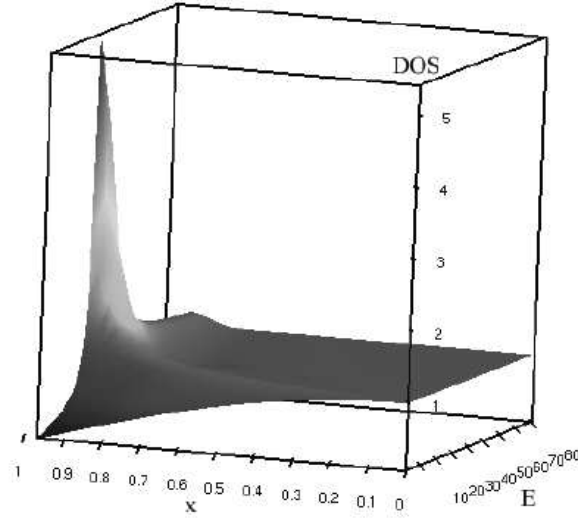


Figure 100: Density of state $N(E, X)$ normalized to the normal state density of states N_0 for a one dimensional wire of length L , as a function of E/E_c and position x along the wire. The superconducting reservoir is at $x = 1.0$, and the normal reservoir is at $x = 0$. The density of states is suppressed at low energies near the superconducting reservoir.

In our analysis above of the proximity effect in a normal metal coupled to a superconductor, we have ignored the effects of electron decoherence on the proximity correction. Phase coherence is essential to observing the proximity effect; if the phase coherence length L_ϕ is less than the length of the sample L , a finite spatial cutoff of the proximity effect is introduced. Phenomenologically, this can be taken into account by saying that the length of the wire is now L_ϕ instead of L . Since L_ϕ is typically of the order of a few microns even at low temperatures, this sets the dimensions of the samples that are required to observe this mesoscopic proximity effect.

While L_ϕ sets the upper cutoff for observing the proximity effect, a second relevant length scale for the problem can be obtained by considering the length at which E_c is equal to $k_B T$. This length is

$$L_T = \sqrt{\frac{\hbar D}{k_B T}}, \quad (22.237)$$

where L_T is called variously the thermal diffusion length or the Thouless length, again from the theory of disordered metals, where it also occurs.[32] (We put in here explicitly the factor of $\hbar$.) In fact, here it is simply the diffusive form of the superconducting coherence length in the normal metal, familiar from the de Gennes/Ginzburg-Landau theory of the proximity effect,[34] which in the clean limit is given by

$$\xi_N = \frac{\hbar v_F}{k_B T}. \quad (22.238)$$

At low temperatures, when L_T is longer than L , the superconducting correlations induced in the normal metal extend throughout its length. At higher temperatures, they are restricted to a region of length L_T near the superconductor. L_T is also on the order of a few microns in typical metallic samples in accessible temperature regimes, so it also sets a limit on the dimensions of the samples in which one can see a proximity effect.

To calculate the thermal conductance of the wire, we proceed from Eqn.(22.230a). We now consider a small temperature differential ΔT applied across the wire. Expanding h_L in a first-order Taylor's expansion, as we did for h_T , we obtain

$$K_1 = -\frac{E\Delta T}{2k_B T^2 \cosh^2\left(\frac{E}{2k_B T}\right)} \left[\int_0^L \frac{1}{M_{00}(E, x)} dx \right]^{-1} \quad (22.239)$$

We then obtain from Eqn.(22.192)

$$j_{th} = -\frac{N_0 D \Delta T}{2k_B T^2} \int dE \frac{E^2}{\cosh^2(E/2k_B T)} \left[\int_0^L \frac{1}{M_{00}(E, x)} dx \right]^{-1} \quad (22.240)$$

As with the electrical conductance, we can define a thermal diffusion coefficient

$$D_0(E, x) = D M_{00}(E, x), \quad (22.241)$$

and a spectral thermal conductance

$$G_{th}(E) = G_{thN} \left[\int_0^L \frac{1}{M_{00}(E, x)} dx \right]^{-1}, \quad (22.242)$$

where G_{thN} is related to the normal state electrical conductance by Eqn.(22.17)

$$G_{thN} = G_N \frac{\pi^2 k_B^2}{3 e^2} T. \quad (22.243)$$

Finally, the thermal conductance itself is given by an integral over energy

$$G_{th} = \frac{3}{\pi^2} \frac{1}{2(k_B T)^3} \int dE \frac{E^2 G_{th}(E)}{\cosh^2(E/2k_B T)}. \quad (22.244)$$

Of course, as noted by Andreev,[36] the thermal conductance of a normal metal wire sandwiched between a normal-metal reservoir on one end and a superconducting reservoir on the other end must vanish, since the superconductor acts as a thermal insulator, so that no thermal current can flow through the device as a whole. However, one may consider a normal-metal wire with the superconducting reservoir connected off to one side, so that it does not block the flow of thermal current through the proximity-coupled wire. One may then consider the thermal conductance of the normal metal wire itself. Figure 7 shows the thermal conductance of this geometry, as a function of temperature, for different transmissivities of the interface barrier. The thermal conductance shows a monotonic decrease as T is lowered below T_c , although there are no distinct features at any particular temperature, unlike for the electrical conductance. In a superconductor, the exponential decrease in the thermal conductivity is associated with the opening of the gap in the quasiparticle density of states, since it is the quasiparticles that carry the thermal current. Noting the decrease in the density of states in the proximity-coupled normal metal wire, shown in Fig. 6, it is not surprising that this system will also show a decrease in the thermal conductance. The thermal conductance of the wire is strongly dependent on the transmission of the NS interface, characterized by the parameter r , and approaches the normal state thermal conductance as r increases.

We note here again that, in our current approximation, a small voltage drop across the S/n-wire/N device will not result in a contribute to a thermal current through the system, since any terms proportional to voltage in the expansion of h_L will cancel. This is the converse of the case for the electrical current, where a small temperature drop did not contribute to the electrical current, emphasizing again that the conventional quasiclassical approximation cannot take into account thermoelectric effects.

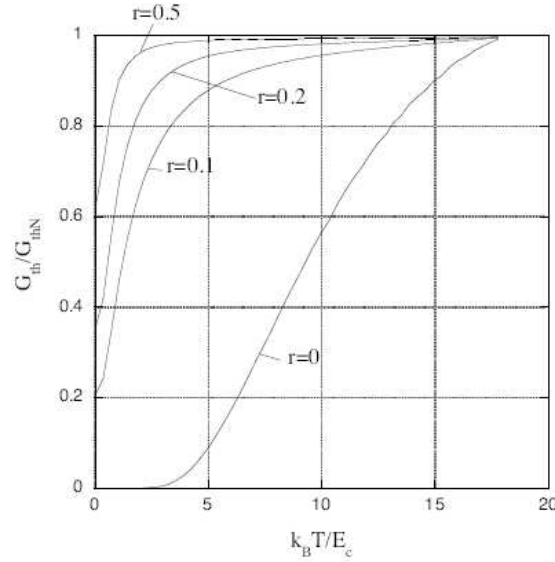


Figure 101: Thermal conductance G_{th} of a normal wire connected to a superconducting reservoir on one end, and a normal metal reservoir on the other, as a function of the normalized temperature T , for a number of different values of the interface barrier parameter r . The gap is set to be $\Delta = 32E_c$, corresponding to a transition temperature of $T = 18.14E_c$.

22.8.2 Superconductor-metal bilayer

Instead of a normal reservoir on one side of the wire, if we consider a wire of length L connected only on one end to a superconducting reservoir (with the other end open), then a true gap in the density of states opens up in the proximity-coupled normal metal. The magnitude of the gap is related to E_c ; hence, one can consider the proximity-coupled normal-metal in this case as a superconductor with a gap of E_c .

If the superconductor is not a reservoir, but a thin layer itself, then one will suppress superconductivity in the superconducting layer due to the proximity of the normal metal, an inverse proximity effect. The suppression of superconductivity is expected to reduce the gap in the superconductor. It is an interesting exercise to calculate the transition temperature of the bilayer in the quasiclassical approximation. We shall loosely follow here the treatment given by Martinis *et al.*[37] and Guéron.[38]

Let the thickness of the superconductor be t_S , and the thickness of the normal metal t_N . We take the origin, $x = 0$, at the NS interface, and the superconductor extends from $x = -t_S$ to $x = 0$, and the normal metal from $x = 0$ to $x = t_N$. Near the transition, the order parameter in the superconductor is small, so we may take the small θ limit of Eqn.(22.215b). Since the phase is not important in this problem, we take $\chi = 0$. The resulting equation is

$$D\partial_x^2\theta + 2Ei\theta - 2i\Delta = 0, \quad (22.245)$$

where we have assumed that the gauge is chosen so that Δ is real. Let us assume that θ at $x = 0$ in the superconductor is θ_{0S} , and that variations of θ about this mean value are small. Under these conditions, we can also assume that the gap Δ in the superconductor is uniform. We can then expand θ_S to second order in x

$$\theta_S = \theta_{0S} + ax + bx^2. \quad (22.246)$$

Now, at $x = -t_S$, (and also at $x = t_N$), we have a vacuum interface, where $\partial_x\theta = 0$. Hence $a = 2bt_S$, and from the differential equation (22.245) taken at $x = 0$, $b = (i/D)(\Delta - E\theta_{0S})$, so that

$$\theta_S = \theta_{0S} + \frac{i}{D}(\Delta - E\theta_{0S})(2t_Sx + x^2). \quad (22.247)$$

Similarly

$$\theta_N = \theta_{0N} + \frac{i}{D} E \theta_{0N} (2t_N x - x^2). \quad (22.248)$$

From the boundary condition Eqn.(22.222b), we have the two equations

$$\frac{2irt_S}{D} (\Delta - E \theta_{0S}) = \theta_{0N} - \theta_{0S}, \quad (22.249a)$$

$$\frac{2irt_N}{D} E \theta_{0N} = \theta_{0N} - \theta_{0S}. \quad (22.249b)$$

Solving this pair of equations for θ_{0S} , we have

$$\theta_{0S} = \frac{\Delta}{E} \frac{D^2 t_S (t_S + t_N) + 4r^2 t_N^2 t_S^2 - 2ir D t_n^2}{D^2 (t_S + t_N)^2 + 4r^2 t_S^2 t_N^2}. \quad (22.250)$$

Putting this into Eqn.(22.228) for the gap, with $T = T_c$, and with the approximation that $\sinh(\theta) \simeq \theta$, we obtain

$$1 = N_0 \frac{\lambda}{2} \int \frac{dE}{E} \tanh(E/2k_B T_c) \left[1 - \frac{t_N (t_S + t_N)}{(t_S + t_N)^2 + (4r^2/D^2) t_S^2 t_N^2} \right]. \quad (22.251)$$

The first term in the square brackets gives the bare transition temperature T_{c0} of the superconducting film, and the second term corresponds to the corrections associated with the inverse proximity effect. For a perfect interface, with $r = 0$, the suppression of T_c is directly proportional to the normal fraction of the bilayer, and $T_c \rightarrow 0$ as t_N increases. For a highly resistive barrier ($r \rightarrow \infty$), the second term in the square brackets goes to zero, so that there is little effect of the normal metal film on T_c of the superconducting film, as expected.

22.8.3 The SNS junction and Andreev interferometers

As our final example of the application of the quasiclassical equations of superconductivity, we consider the case of a dirty SNS junction. The model we consider is a normal metal wire of length L sandwiched between two superconducting reservoirs. As is well known, the application of a phase difference between the two superconducting reservoirs will result in the flow of a supercurrent through the normal wire. The phase difference can be applied, for example, by connecting one of the superconducting reservoirs to the other, thereby forming a loop with two different arms, one superconducting and one normal. This configuration is commonly called an Andreev interferometer. The phase between the two superconducting reservoirs can be varied by applying an Aharonov-Bohm type magnetic flux through the area of the loop; in this respect, we put together all contributions in the gauge-invariant phase χ . Due to the single-valued nature of the wave functions, a phase factor of $2\pi\Phi/\Phi_0$ is picked up in going around the loop, where Φ is the magnetic flux threading the Andreev interferometer, and $\Phi_0 = h/2e$ is the superconducting flux quantum. In a superconductor, the supercurrent I_S that is generated is directly proportional to the phase gradient; if I_S is small compared to the critical current I_c , the phase dropped across the superconductor will be small. Since I_c of the superconducting part of the Andreev interferometer is so much greater than the critical current of the proximity coupled normal-metal wire, most of the phase change will occur across the length L of the normal metal wire.

This fact allows us to map the Andreev interferometer that is coupled with an Aharonov-Bohm flux Φ to a SNS system with a phase difference $\phi = 2\pi\Phi/\Phi_0$ across it. In terms of our model, we consider the superconductors to be reservoir; this means applying a boundary condition for the gauge-invariant phase χ at the superconducting reservoirs. For our purposes,

we apply this boundary condition anti-symmetrically, with a phase $\chi_L = -\pi\Phi/\Phi_0$ at the superconducting reservoir at $x = 0$, and $\chi_R = \pi\Phi/\Phi_0$ at the superconducting reservoir at $x = L$. We then must solve Eqns.(22.215) in the normal-metal wire for θ and χ , with $\Delta = 0$, subject to the boundary condition for χ noted above, and the boundary condition $\theta = \theta_{S0}$ (where θ_{S0} is given by Eqn.(22.220)) at both $x = 0$ and $x = L$. In general, both θ and χ are complex functions of x and E , and the solution of Eqns.(22.215) must be done numerically.

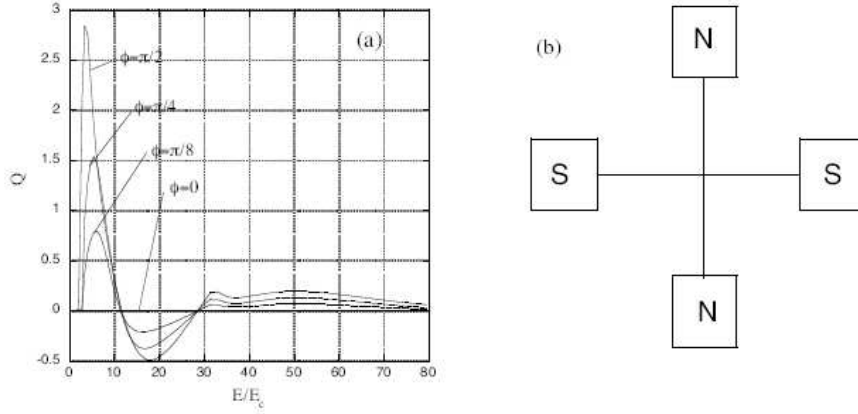


Figure 102: (a) Spectral supercurrent Q in a normal wire between two superconducting reservoirs, as a function of the normalized energy E/E_c , for different values of the phase difference ϕ between the superconducting reservoirs. (b) Geometry of an Andreev interferometer, essentially a cross with one arm connected to superconducting reservoirs, and the other arm connected to normal reservoirs.

Following Yip,[39] we consider first the supercurrent Q , given by Eqn.(22.218). Some insight into the contributions to j_s can be gained by looking again at the case of a bulk superconductor. From Eqn.(22.202), the major contribution to j_s comes from energies near the gap. For a long proximity wire with $E_c < \Delta$, however, the major contribution comes from energies of order E_c . Figure 8(a) shows a plot of Q as a function of energy for various values of the phase difference ϕ between the two superconducting reservoirs, for the case of a perfectly transparent interface, and $\Delta = 32E_c$. Of course, for zero phase difference, the supercurrent vanishes. As ϕ is increased from zero, there is a peak in $Q(E)$ at $E \simeq 6E_c$. This peak moves down in energy as ϕ increases. At larger values of E , Q becomes negative. For shorter wires, this region of negative Q is less prominent. The total supercurrent is given by the second term in Eqn.(22.191)

$$J_s = eN_0D \int dE Q(E)h_L(E), \quad (22.252)$$

and therefore depends also on the distribution of quasiparticles. Any change in this distribution will affect the supercurrent. For example, the distribution can be changed by increasing the temperature, which has the result of decreasing the supercurrent. As we demonstrated in the introduction, a non-equilibrium quasiparticle distribution can also be generated by injecting a normal current into the proximity wire in the SNS geometry by attaching two additional leads to the center of the normal wire, forming a normal cross, with two of the wires attached to superconducting reservoirs, and the two other wires attached to normal reservoirs, as shown in Fig. 8(b). The current in the SNS junction is then a function of the current injected between the normal reservoirs, and the supercurrent can even change sign depending on magnitude of the injected normal current.[39, 40] This effect has been observed in recent experiments.[41]

Due to long-range phase coherence, the Green's function in the arms of the cross attached to the normal reservoirs will also depend on the phase difference between the two superconducting reservoirs in the structure shown in Fig. 8(b). Consequently, the electrical conductance

measured between the two normal reservoirs will also be a periodic function of the phase difference between the two superconducting reservoirs. Experimentally, the conductance of such Andreev interferometers have been found to oscillate periodically with an applied external flux, with a fundamental period of $\phi_0 = h/2e$. [42, 43, 44] Similar oscillations are also expected in the thermal conductance as well. Periodic oscillations are also observed in the thermopower of Andreev interferometers, [45] although these thermopower oscillations cannot be described within the framework of the current quasiclassical theory.

22.9 Summary

The quasiclassical theory of superconductivity has proved to be a powerful tool for the quantitative description of long-range phase coherent phenomena in diffusive proximity coupled systems. As we have shown, the linear electrical and thermal conductance of complicated devices incorporating normal and superconducting elements can be calculated in principle, although the solutions frequently involve numerical techniques. Extension to the nonlinear regime, with finite voltages across the normal reservoirs, is also conceptually straightforward, although numerically challenging.

Application of finite voltages to the superconducting elements is trickier, as it involves time dependent evolution of the phase, and is only beginning to be examined theoretically. Finally, the quasiclassical theory for diffusive systems, in its present form, does not deal at all with thermoelectric phenomena. Extensions to incorporate thermoelectric effects in the theoretical framework have been attempted, [46] but still require further work to be complete.

Part V

Adds

A Дополнения

A.1 Useful mathematical formulas

A.1.1 Useful mathematical formulas for SC rings

$$\sum_n \frac{e^{in\vartheta}}{n+A} = \frac{2\pi i e^{-iA\vartheta}}{1 - e^{-2\pi i A}}, \quad \vartheta \in [0, 2\pi] \quad (\text{A.1})$$

$$\frac{\partial}{\partial \varphi} \ln \det d(z) = \text{tr} \left\{ [d(z)]^{-1} \frac{\partial d(z)}{\partial \varphi} \right\} \quad (\text{A.2})$$

A.2 Tips about mathematical transformations

How not to make expressions big if you work with BGF method?

Introduce new parameters! Do it often, otherwise expressions will be too big to work!

A.3 Acknowledgements

Currently, no one except me has worked on the sections of this note (with the exception of sections taken from books).

A.4 Literature

A.5 Codes for Green's functions

(приведу потом всякие коды)

A.6 Overview of applications (????)

(серьезно, вопрос, зачем это нужно и что с этим можно посчитать? какие устройства можно изобрести??? пока большой минус записи - что она просто обходит этот вопрос стороной, будто вот формулы, и всё. но это большие, сложные и отдельные вопросы, потом их буду продумывать, мб нужно у экспериментаторов спросить тоже про это)

B Bibliography

References

References from Wimmer's PHD

- [1] G. E. Moore. Cramming more components onto integrated circuits. *Electronics* 38 (1965).
- [2] A. L. Robinson. One billion transistors on a chip?. *Science* 223 267-268 (1984).
- [3] S. Thompson P. Packan and M. Bohr. MOS scaling: Transistor challenges for the 21st century. *Intel Tech. J. Q3* (1998).
- [4] M. Schulz. The end of the road for silicon?. *Nature* 399 729-730 (1999).
- [5] P. A. Packan. Pushing the limits. *Science* 285 2079-2081 (1999).
- [6] P. S. Peercy. The drive to miniaturization. *Nature* 406 1023-1026 (2000).
- [7] S. Lloyd. Ultimate physical limits to computation. *Nature* 406 1047-1054 (2000).
- [8] V. V. Zhirnov R. K. Cavin and J. A. Hutchby. Limits to binary logic switch scaling? A gedanken model. *Proc. IEEE* 91 1934-1939 (2003).
- [9] J. D. Meindl Q. Chen and J. A. Davis. Limits on silicon nanoelectronics for terascale integration. *Science* 293 2044-2049 (2001).
- [10] M. Lundstrom. Enhanced: Moore's law forever?. *Science* 299 210-211 (2003).
- [11] S. A. Wolf D. D. Awschalom R. A. Buhrman J. M. Daughton S. von Molnar M. L. Roukes A. Y. Chtchelkanova and D. M. Treger. Spintronics: A spin-based electronics vision for the future. *Science* 294 1488-1495 (2001).
- [12] I. Uté J. Fabian and S. Das Sarma. Spintronics: Fundamentals and applications. *Rev. Mod. Phys.* 76 323-410 (2004).
- [13] J. Fabian A. Matos-Abiad C. Ertler P. Stano and I. Uté. Semiconductor spintronics. *Acta Phys. Slovaca* 4&5 565-907 (2007).
- [14] G. Binash P. Grnberg F. Saurenbach and W. Zinn. Enhanced magnetoresistance in layered magnetic structures with antiferromagnetic interlayer exchange. *Phys. Rev. B* 39 4828-4830 (1989).
- [15] M. N. Baibich J. M. Broto A. Fert F. N. Van Dau F. Petroff P. Eitenne G. Creuzet A. Friederich and J. Chazelas. Giant magnetoresistance of (001)Fe/(001)Cr magnetic superlattices. *Phys. Rev. Lett.* 61 2472-2475 (1988).
- [16] M. Jullière. Tunneling between ferromagnetic films. *Phys. Lett. A* 54 225-226 (1975).
- [17] J. S. Moodera L. R. Kinder T. M. Wong and R. Meservey. Large magnetoresistance at room temperature in ferromagnetic thin film tunnel junctions. *Phys. Rev. Lett.* 74 3273-3276 (1995).
- [18] T. Miyazaki and N. Tezuka. Giant magnetic tunneling effect in Fe/Al₂O₃/Fe junction. *J. Mag. Mag. Mater.* 139 L231-L234 (1995).
- [19] C. Chappert A. Fert and F. N. Van Dau. The emergence of spin electronics in data storage. *Nat. Mater.* 6 813-823 (2007).
- [20] S. Datta and B. Das. Electronic analog of the electro-optic modulator. *Appl. Phys. Lett.* 56 665-667 (1990).
- [21] Y. A. Bychkov and E. I. Rashba. Oscillatory effects and the magnetic susceptibility of carriers in inversion layers. *J. Phys. C: Solid State Phys.* 17 6039-6045 (1984).
- [22] Y. A. Bychkov and E. I. Rashba. Properties of a 2d electron gas with lifted spectral degeneracy. *JETP Lett.* 39 78-81 (1984).
- [23] J. Schliemann J. C. Egues and D. Loss. Nonballistic spin-field-effect transistor. *Phys. Rev. Lett.* 90 146801 (2003).
- [24] S. Bandyopadhyay and M. Cahay. Alternate spintronic analog of the electrooptic

modulator. Appl. Phys. Lett. 85 1814-1816 (2004).

[25] M. I. D'yakonov and V. I. Perel'. Possibility of orienting electron spins. JETP Lett. 13 467-469 (1971).

[26] J. E. Hirsch. Spin Hall effect. Phys. Rev. Lett. 83 1834-1837 (1999).

[27] S. Murakami N. Nagaosa and S.-C. Zhang. Dissipationless quantum spin current at room temperature. Science 301 1348-1351 (2003).

[28] J. Sinova D. Culcer Q. Niu N. A. Sinitsyn T. Jungwirth and A. H. MacDonald. Universal intrinsic spin Hall effect. Phys. Rev. Lett. 92 126603 (2004).

[29] K. S. Novoselov A. K. Geim S. V. Morozov D. Jiang Y. Zhang S. V. Dubonos I. V. Grigorieva and A. A. Firsov. Electric field effect in atomically thin carbon films. Science 306 666-669 (2004).

[30] N. D. Mermin. Crystalline order in two dimensions. Phys. Rev. 176 250-254 (1968).

[31] A. K. Geim and K. S. Novoselov. The rise of graphene. Nat. Mater. 6 183-191 (2007).

[32] K. S. Novoselov A. K. Geim S. V. Morozov D. Jiang M. I. Katsnelson I. V. Grigorieva S. V. Dubonos and A. A. Firsov. Two-dimensional gas of massless dirac fermions in graphene. Nature 438 197-200 (2005).

[33] K. S. Novoselov Z. Jiang Y. Zhang S. V. Morozov H. L. Stormer U. Zeitler J. C. Maan G. S. Boebinger P. Kim and A. K. Geim. Room-temperature quantum Hall effect in graphene. Science 315 1379 (2007).

[34] M. I. Katsnelson K. S. Novoselov and A. K. Geim. Chiral tunnelling and the Klein paradox in graphene. Nat. Phys. 2 620-625 (2006).

[35] A. H. Castro Neto F. Guinea N. M. R. Peres K. S. Novoselov and A. K. Geim. The electronic properties of graphene. To be published in Rev. Mod. Phys. arXiv:0709.1163v2 (2008).

[36] K. Bolotin K. Sikes Z. Jiang M. Klima G. Fudenberg J. Hone P. Kim and H. Stormer. Ultrahigh electron mobility in suspended graphene. Solid State Commun. 146 351-355 (2008).

[37] K. I. Bolotin K. J. Sikes J. Hone H. L. Stormer and P. Kim. Temperature dependent transport in suspended graphene. arXiv:0805.1830v1 (2008).

[38] B. Trauzettel D. V. Bulaev D. Loss and G. Burkard. Spin qubits in graphene quantum dots. Nature Phys. 3 192-196 (2007).

[39] D. Huertas-Hernando F. Guinea and A. BrataAs. Spin-orbit coupling in curved graphene fullerenes nanotubes and nanotube caps. Phys. Rev. B 74 155426 (2006).

[40] H. Min J. E. Hill N. A. Sinitsyn B. R. Sahu L. Kleinman and A. H. MacDonald. Intrinsic and Rashba spin-orbit interactions in graphene sheets. Phys. Rev. B 74 165310 (2006).

[41] D. Huertas-Hernando F. Guinea and A. Brataas. Spin relaxation times in disordered graphene. Eur. Phys. J. Spec. Top. 148 177-181 (2007).

[42] E. W. Hill A. K. Geim K. S. Novoselov F. Schedin and P. Blake. Graphene spin valve devices. IEEE Trans. Magn. 42 2694-2696 (2006).

[43] N. Tombros C. Jozsa M. Popinciuc H. T. Jonkman and B. J. van WeEs. Electronic spin transport and spin precession in single graphene layers at room temperature. Nature 448 571-574 (2007).

[44] S. Cho Y.-F. Chen and M. S. Fuhrer. Gate-tunable graphene spin valve. Appl. Phys. Lett. 91 123105 (2007).

[45] N. Tombros S. Tanabe A. Veligura C. Jozsa M. Popinciuc H. T. Jonkman and B. J. van Wees. Anisotropic spin relaxation in graphene. Phys. Rev. Lett. 101 046601 (2008).

[46] C. Jzsa M. Popinciuc N. Tombros H. T. Jonkman and B. J. van WeEs. Electronic spin drift in graphene field-effect transistors. Phys. Rev. Lett. 100 236603 (2008).

[47] J. Moser A. Matos-Abiague D. Schuh W. Wegscheider J. Fabian and D. Weiss. Tunneling anisotropic magnetoresistance and spin-orbit coupling in Fe/GaAs/Au tunnel junctions. Phys. Rev. Lett. 99 056601 (2007).

- [48] M. Lobenhofer. Unpublished.
- [49] D. A. Grier. Human computers: the first pioneers of the information age. *Endvr.* 25 28-32 (2001).
- [50] J. Wurm A. Rycerz I. Adagideli M. Wimmer K. Richter and H. U. BARANGER. Symmetry classes in graphene quantum dots: Universal spectral statistics weak localization and conductance fluctuations. *arXiv:0808.1008v1* (2008).
- [51] M. Scheid M. Wimmer D. Bercioux and K. Richter. Zeeman ratchets for ballistic spin currents. *Phys. Status Solidi (c)* 3 4235-4238 (2006).
- [52] I. Adagideli M. Scheid M. Wimmer G. E. W. Bauer and K. Richter. Extracting current-induced spins: spin boundary conditions at narrow Hall contacts. *New J. Phys.* 9 382 (2007).
- [53] J. Wurm. Quantum Transport in Graphene-Based Nanosystems. Diplomarbeit Universitt Regensburg (2008).
- [54] M. Scheid A. Pfund D. Bercioux and K. Richter. Coherent spin ratchets: A spin-orbit based quantum ratchet mechanism for spin-polarized currents in ballistic conductors. *Phys. Rev. B* 76 195303 (2007).
- [55] M. Scheid D. Bercioux and K. Richter. Zeeman ratchets: pure spin current generation in mesoscopic conductors with non-uniform magnetic fields. *New J. Phys.* 9 401 (2007).
- [56] V. Krckl. Mesoskopischer Transport in zweidimensionalen Lochgasen. Diplomarbeit Universitt Regensburg (2008).
- [57] M. Wimmer. Unpublished.
- [58] A. L. Fetter and J. Walecka. Quantum Theory of Many-Particle Systems (Dover Mineola 2003).
- [59] S. DatTa. Electronic Transport in Mesoscopic Transport (Cambridge University Press Cambridge 2002).
- [60] D. K. Ferry and S. M. Goodnick. Transport in Nanostructures (Cambridge University Press Cambridge 2001).
- [61] H. Haug and A.-P. Jauno. Quantum Kinetics in Transport and Optics of Semiconductors (Springer Berlin Heidelberg 1998).
- [62] H. Bruus and K. Flensberg. Many-body Quantum Theory in Condensed Matter Physics: An Introduction (Oxford University Press Oxford 2004).
- [63] J. Rammer. Quantum Field Theory of Non-equilibrium States (Cambridge University Press Cambridge 2007).
- [64] L. P. Kadanoff and G. Baym. Quantum Statistical Mechanics (Benjamin New York 1962).
- [65] L. V. Keldysh. Diagram technique for nonequilibrium processes. *Sov. Phys. JETP* 20 1018-1026 (1965).
- [66] J. Rammer and H. Smith. Quantum field-theoretical methods in transport theory of metals. *Rev. Mod. Phys.* 58 323-359 (1986).
- [67] P. Danielewicz. Quantum theory of nonequilibrium processes I. *Ann. Phys. (NY)* 152 239-304 (1984).
- [68] M. Wagner. Expansions of nonequilibrium Green's functions. *Phys. Rev. B* 44 6104-6117 (1991).
- [69] S. Hershfield. Reformulation of steady state nonequilibrium quantum statistical mechanics. *Phys. Rev. Lett.* 70 2134-2137 (1993).
- [70] F. Schwabl. Advanced Quantum Mechanics (Springer Berlin 1999).
- [71] R. A. Craig. Perturbation expansion for real-time Green's functions. *J. Math. Phys.* 9 605-611 (1968).
- [72] D. C. LangReth. Linear and nonlinear response theory with applications. In: Linear and Nonlinear Electron Transport edited by J. T. Pevreese and E. Van Doren pp. 3-32 (Plenum New York 1976).

- [73] S. Datta. Electrical resistance: an atomistic view. *Nanotech.* 15 S433-S451 (2004).
- [74] G. D. Mahan. *Many-Particle Physics* (Kluwer New York NY 2005).
- [75] D. C. Langreth and J. W. Wilkins. Theory of spin resonance in dilute magnetic alloys. *Phys. Rev. B* 6 3189-3227 (1972).
- [76] J. H. Davies S. Hershfield P. Hyldgaard and J. W. Wilkins. Current and rate equation for resonant tunneling. *Phys. Rev. B* 47 4603-4618 (1993).
- [77] G. E. Kimball and G. H. Shortley. The numerical solution of Schrödinger's equation. *Phys. Rev.* 45 815-820 (1934).
- [78] L. Pauling and E. B. Wilson. *Introduction to Quantum Mechanics* (Dover New York 1935).
- [79] D. Frustaglia M. Hentschel and K. Richter. Aharonov-Bohm physics with spin. II. Spin-flip effects in two-dimensional ballistic systems. *Phys. Rev. B* 69 155327 (2004).
- [80] P. Havu V. Havu M. J. Puska and R. M. Nieminen. Nonequilibrium electron transport in two-dimensional nanostructures modeled using Green's functions and the finite-element method. *Phys. Rev. B* 69 115325 (2004).
- [81] R. C. Bowen G. Klimeck R. K. Lake W. R. Frensley and T. Moise. Quantitative simulation of a resonant tunneling diode. *J. Appl. Phys.* 81 32073213 (1997).
- [82] S. Sanvito C. J. Lambert J. H. Jefferson and A. M. Bratkovsky. General Green's-function formalism for transport calculations with spd Hamiltonians and giant magnetoresistance in Co- and Ni-based magnetic multilayers. *Phys. Rev. B* 59 11936-11948 (1999).
- [83] M. Luisier A. Schenk W. Fichtner and G. Klimeck. Atomistic simulation of nanowires in the $sp^3d^5s^*$ tight-binding formalism: From boundary conditions to strain calculations. *Phys. Rev. B* 74 205323 (2006).
- [84] M. Brandbyge J.-L. Mozos P. Ordejón J. Taylor and K. Stokbro. Density-functional method for nonequilibrium electron transport. *Phys. Rev. B* 65 165401 (2002).
- [85] A. Di Carlo A. Pecchia L. Latessa T. Frauenheim and G. Seifert. Tight-binding DFT for molecular electronics (gDFTB). In: *Introducing Molecular Electronics* edited by G. Cuniberti G. Fagas and K. Richter pp. 153-184 (Springer Berlin Heidelberg 2006).
- [86] A. R. Rocha V. M. García-Suárez S. Bailey C. Lambert J. Ferrer and S. Sanvito. Spin and molecular electronics in atomically generated orbital landscapes. *Phys. Rev. B* 73 085414 (2006).
- [87] Y. Meir and N. S. Wingreen. Landauer formula for the current through an interacting electron region. *Phys. Rev. Lett.* 68 2512-2515 (1992).
- [88] R. Landauer. Spatial variation of currents and fields due to localized scatterers in metallic conduction. *IBM J. Res. Dev.* 1 223-231 (1957).
- [89] M. Büttiker Y. Imry R. Landauer and S. Pinhas. Generalized manychannel conductance formula with application to small rings. *Phys. Rev. B* 31 6207-6215 (1985).
- [90] D. S. Fisher and P. A. Lee. Relation between conductivity and transmission matrix. *Phys. Rev. B* 23 6851-6854 (1981).
- [91] A. Lassl. *Spin-dependent Transport of Interacting Electrons in Mesoscopic Systems*. Dissertation Universität Regensburg (2007).
- [92] T. Usuki M. Takatsu R. A. Kiehl and N. Yokoyama. Numerical analysis of electron-wave detection by a wedge-shaped point contact. *Phys. Rev. B* 50 7615-7625 (1994).
- [93] N. W. Ashcroft and N. D. Mermin. *Solid State Physics* (Harcourt College Publishers Orlando FL 1976).
- [94] R. E. Allen. Green's function and generalized phase shift for surface and interface problems. *Phys. Rev. B* 19 917-924 (1979).
- [95] R. E. Allen. Green's functions for surface physics. *Phys. Rev. B* 20 1454-1472 (1979).
- [96] D. H. Lee and J. D. Joannopoulos. Simple scheme for surface-band calculations. I. *Phys. Rev. B* 23 4988-4996 (1981).

- [97] D. H. Lee and J. D. Joannopoulos. Simple scheme for surface-band calculations. II. The Green's function. *Phys. Rev. B* 23 4997-5004 (1981).
- [98] M. P. Lopez Sancho J. M. Lopez Sancho and J. Rubio. Quick iterative scheme for the calculation of transfer matrices: application to Mo (100). *J. Phys. F: Met. Phys.* 14 1205-1215 (1984).
- [99] M. P. Lopez Sancho J. M. Lopez Sancho and J. Rubio. Highly convergent schemes for the calculation of bulk and surface Green functions. *J. Phys. F: Met. Phys.* 15 851-858 (1985).
- [100] J. VeleV and W. Butler. On the equivalence of different techniques for evaluating the Green function for a semi-infinite system using a localized basis. *J. Phys.: Condens. Matter* 16 R637-R657 (2004).
- [101] A. Umerski. Closed-form solutions to surface Green's functions. *Phys. Rev. B* 55 5266-5275 (1997).
- [102] P. S. Krsti X.-G. Zhang and W. H. Butler. Generalized conductance formula for the multiband tight-binding model. *Phys. Rev. B* 66 205319 (2002).
- [103] A. MacKinnon. The calculation of transport properties and density of states of disordered solids. *Z. Phys. B* 59 385-390 (1985).
- [104] T. Ando. Quantum point contacts in magnetic fields. *Phys. Rev. B* 44 80178027 (1991).
- [105] P. A. Khomyakov G. Brocks V. Karpan M. Zwierzycki and P. J. Kelly. Conductance calculations for quantum wires and interfaces: Mode matching and Green's functions. *Phys. Rev. B* 72 035450 (2005).
- [106] F. Tisseur and K. Meerbergen. The quadratic eigenvalue problem. *SIAM Rev.* 43 235-286 (2001).
- [107] E. Anderson Z. Bai C. Bischof S. Blackford J. Demmel J. Dongarra J. Du Croz A. Greenbaum S. Hammarling A. McKenney and D. Sorensen. *LAPACK Users' Guide* 3rd edition (Society for Industrial and Applied Mathematics Philadelphia PA 1999).
- [108] T. B. Boykin. Generalized eigenproblem method for surface and interface states: The complex bands of GaAs and AlAs. *Phys. Rev. B* 54 8107-8115 (1996).
- [109] R. P. Feynman. Forces in molecules. *Phys. Rev.* 56 340-343 (1939).
- [110] H. Hellmann. *Einfhrung in die Quantenchemie* (Deuticke Leipzig 1937).
- [111] C. W. J. Beenakker and H. van Houten. Quantum transport in semiconductor nanostructures. *Solid State Phys.* 44 1-228 (1991).
- [112] G. H. Golub and C. F. Van Loan. *Matrix Computations* 3rd edition (The John Hopkins University Press 1996).
- [113] A. R. Williams P. J. Feibelman and N. D. Lang. Green's-function methods for electronic-structure calculations. *Phys. Rev. B* 26 5433-5444 (1982).
- [114] T. Usuki M. Saito M. Takatsu R. A. Kiehl and N. Yokoyama. Numerical analysis of ballistic-electron transport in magnetic fields by using a quantum point contact and a quantum wire. *Phys. Rev. B* 52 8244-8255 (1995).
- [115] C. J. Lambert and D. Weaire. Decimation and Anderson localization. *Phys. Status Solidi (b)* 101 591-595 (1980).
- [116] M. Leadbeater and C. J. Lambert. A decimation method for studying transport properties of disordered systems. *Ann. Phys.* 7 498-502 (1998).
- [117] D. Mamaluy D. Vasileska M. Sabathil T. Zibold and P. Vogl. Contact block reduction method for ballistic transport and carrier densities of open nanostructures. *Phys. Rev. B* 71 245321 (2005).
- [118] D. J. Thouless and S. Kirkpatrick. Conductivity of the disordered linear chain. *J. Phys. C: Solid State Phys.* 14 235-245 (1981).
- [119] P. A. Lee and D. S. Fisher. Anderson localization in two dimensions. *Phys. Rev. Lett.* 47 882-885 (1981).
- [120] H. U. Baranger D. P. DiVincenzo R. A. Jalabert and A. D. Stone. Classical and

- quantum ballistic-transport anomalies in microjunctions. *Phys. Rev. B* 44 10637-10675 (1991).
- [121] R. Lake G. Klimeck R. C. Bowen and D. Jovanovic. Single and multiband modeling of quantum electron transport through layered semiconductor devices. *J. Appl. Phys.* 81 7845-7869 (1997).
- [122] A. Lassi P. Schlagheck and K. Richter. Effects of short-range interactions on transport through quantum point contacts: A numerical approach. *Phys. Rev. B* 75 045346 (2007).
- [123] P. Drouvelis P. Schmelcher and P. Bastian. Parallel implementation of the recursive Green's function method. *J. Comp. Phys.* 215 741-756 (2006).
- [124] T. Kramer E. J. Heller and R. E. Parrott. An efficient and accurate method to obtain the energy-dependent Green function for general potentials. *J. Phys.: Conf. Ser.* 99 012010 (2008).
- [125] S. Rotter J.-Z. Tang L. Wirtz J. Trost and J. BurgdOrfer. Modular recursive Green's function method for ballistic quantum transport. *Phys. Rev. B* 62 1950-1960 (2000).
- [126] S. Rotter B. Weingartner N. Rohringer and J. BurgdOrfer. Ballistic quantum transport at high energies and high magnetic fields. *Phys. Rev. B* 68 165302 (2003).
- [127] N. E. Gibbs J. William G. Poole and P. K. Stockmeyer. An algorithm for reducing the bandwidth and profile of a sparse matrix. *SIAM J. Num. Anal.* 13 236-250 (1976).
- [128] E. Cuthill and J. McKee. Reducing the bandwidth of sparse symmetric matrices. In: *Proceedings of the 1969 24th national conference* pp. 157-172 (ACM New York 1969).
- [129] A. GEORGE. Computer implementation of the finite element method. Technical Report STAN-CS-71-208 Computer Sci. Dept. Stanford Univ. Stanford CA (1971).
- [130] W.-H. Liu and A. H. Sherman. Comparative analysis of the Cuthill-McKee and the reverse Cuthill-McKee ordering algorithms for sparse matrices. *SIAM J. Num. Anal.* 13 198-213 (1976).
- [131] B. Kernighan and S. Lin. An efficient heuristic procedure for partitioning graphs. *Bell Syst. Tech. J.* 49 291-308 (1970).
- [132] C. M. Fiduccia and R. M. Mattheyses. A linear-time heuristic for improving network partitions. In: *DAC '82: Proceedings of the 19th conference on Design automation* pp. 175-181 (IEEE Press Piscataway NJ USA 1982).
- [133] G. Karypis and V. Kumar. Multilevel k -way hypergraph partitioning. *VLSI Design* 11 285-300 (2000).
- [134] B. Hendrickson and E. Rothberg. Improving the run time and quality of nested dissection ordering. *SIAM J. Sci. Comput.* 20 468-489 (1998).
- [135] J. O'Neil and D. B. Szyld. A block ordering method for sparse matrices. *SIAM J. Sci. Stat. Comput.* 11 811-823 (1990).
- [136] A. Coon and M. Stadtherr. Generalized block-tridiagonal matrix orderings for parallel computation in process flowsheeting. *Comput. Chem. Eng.* 19 787805 (1995).
- [137] K. V. Camarda and M. A. Stadtherr. Matrix ordering strategies for process engineering: graph partitioning algorithms for parallel computation. *Comput. Chem. Eng.* 23 1063-1073 (1999).
- [138] B. Hendrickson. Graph partitioning and parallel solvers: Has the emperor no clothes?. In: *Workshop on Parallel Algorithms for Irregularly Structured Problems* pp. 218-225 (1998).
- [139] B. Hendrickson and T. G. Kolda. Partitioning rectangular and structurally unsymmetric sparse matrices for parallel processing. *SIAM J. Sci. Comput.* 21 2048-2072 (2000).
- [140] C. Aykanat A. Pinar and mit V. atalyrek. Permuting sparse rectangular matrices into block-diagonal form. *SIAM J. Sci. Comput.* 25 1860-1879 (2004).
- [141] R. Sedgewick. *Algorithms in C++* 2nd edition (Addison Wesley 1992).
- [142] A. Gupta. Fast and effective algorithms for graph partitioning and sparse-matrix ordering. *IBM J. Res. Dev.* 41 171-184 (1997).

- [143] H. D. Simon and S.-H. Teng. How good is recursive bisection?. SIAM J. Sci. Comput. 18 1436-1445 (1997).
- [144] M. R. Garey and D. S. Johnson. Computers and Intractability: A Guide to the Theory of NP-Completeness (W. H. Freeman & Co. New York NY USA 1990).
- [145] R. C. Whaley A. Petitet and J. J. Dongarra. Automated empirical optimization of software and the ATLAS project. Parallel Comput. 27 3-35 (2001).
- [146] K. Kazymyrenko and X. Waintal. Knitting algorithm for calculating Green functions in quantum systems. Phys. Rev. B 77 115119 (2008).
- [147] Z. H. Qiao and J. Wang. A variant transfer matrix method suitable for transport through multi-probe systems. Nanotech. 18 435402 (2007).
- [148] K. v. Klitzing G. Dorda and M. Pepper. New method for high-accuracy determination of the fine-structure constant based on quantized hall resistance. Phys. Rev. Lett. 45 494-497 (1980).
- [149] C. Gould C. Rster T. Jungwirth E. Girgis G. M. Schott R. Giraud K. Brunner G. Schmidt and L. W. Molenkamp. Tunneling anisotropic magnetoresistance: A spin-valve-like tunnel magnetoresistance using a single magnetic layer. Phys. Rev. Lett. 93 117203 (2004).
- [150] M. Lobenhofer. TAMR-Effekt in epitaktischen Fe/GaAs-Systemen. Diplomarbeit Universitt Regensburg (2007).
- [151] A. Matos-Abiague and J. Fabian. Spin-orbit induced anisotropy in the tunneling magnetoresistance of magnetic tunnel junctions. arXiv:condmat/0702387v2 (2007).
- [152] A. Matos-Abiague and J. Fabian. Anisotropic tunneling magnetoresistance and tunneling anisotropic magnetoresistance: spin-orbit coupling in magnetic tunnel junctions. arXiv:0807.3564v1 (2008).
- [153] C. Cohen-Tannoudji B. Diu and F. Lalo. Qunatum Mechanics (John Wiley & Sons New York NY 1977).
- [154] H. U. Baranger and A. D. Stone. Electrical linear-response theory in an arbitrary magnetic field: A new Fermi-surface formation. Phys. Rev. B 40 8169-8193 (1989).
- [155] M. Wimmer. Spinpolarisierter Transport in Ferromagnet/HalbleiterTunnelsystemen. Diplomarbeit Universitt Regensburg (2004).
- [156] L. Eaves K. W. H. Stevens and F. W. Sheard. Tunnel currents and electron tunnelling times in semiconductor heterostructure barriers in the presence of an applied magnetic field. In: The physics and fabrication of microstructures and microdevices edited by M. Kelly and C. Weisbuch volume 13 of Springer Proceedings in Physics (Springer Berlin 1986).
- [157] M. I. Katsnelson. Graphene: carbon in two dimensions. Mater. Today 10 20-27 (2007).
- [158] M. Y. Han B. Ozyilmaz Y. Zhang and P. Kim. Energy band-gap engineering of graphene nanoribbons. Phys. Rev. Lett. 98 206805 (2007).
- [159] Z. Chen Y.-M. Lin M. J. Rooks and P. Avouris. Graphene nano-ribbon electronics. Physica E 40 228-232 (2007).
- [160] X. Li X. Wang L. Zhang S. Lee and H. Dai. Chemically derived ultrasmooth graphene nanoribbon semiconductors. Science 319 1229-1232 (2008).
- [161] X. Wang Y. Ouyang X. Li H. Wang J. Guo and H. Dai. Room temperature all semiconducting sub-10nm graphene nanoribbon field-effect transistors. arXiv:0803.3464v1 to appear in Phys. Rev. Lett. (2008).
- [162] I. Martin and Y. M. Blanter. Transport in disordered graphene nanoribbons. arXiv:0705.0532v2 (2007).
- [163] F. Sols F. Guinea and A. H. C. Neto. Coulomb blockade in graphene nanoribbons. Phys. Rev. Lett. 99 166803 (2007).
- [164] Y.-W. Son M. L. Cohen and S. G. Louie. Energy gaps in graphene nanoribbons. Phys. Rev. Lett. 97 216803 (2006).
- [165] M. Evaldsson I. V. Zozoulenko H. Xu and T. Heinzl. Edge disorder induced Anderson

- localization and conduction gap in graphene nanoribbons. arXiv:0805.4326v1 (2008).
- [166] E. R. Mucciolo A. H. C. Neto and C. H. Lewenkopf. Conductance quantization and transport gap in disordered graphene nanoribbons. arXiv:0806.3777v1 (2008).
- [167] L. Tapasztó G. Dobrik P. Lambin and L. P. Biro. Tailoring the atomic structure of graphene nanoribbons by scanning tunnelling microscope lithography. *Nature Nanotech.* 3 397-401 (2008).
- [168] S. S. Datta D. R. Strachan S. M. Khamis and A. T. C. Johnson. Crystallographic etching of few-layer graphene. *Nano Lett.* 8 1912-1915 (2008).
- [169] K. Nakada M. Fujita G. Dresselhaus and M. S. Dresselhaus. Edge state in graphene ribbons: Nanometer size effect and edge shape dependence. *Phys. Rev. B* 54 17954-17961 (1996).
- [170] M. Fujita K. Wakabayashi K. Nakada and K. Kusakabe. Peculiar localized state at zigzag graphite edge. *J. Phys. Soc. Jpn.* 65 1920-1923 (1996).
- [171] A. R. Akhmerov and C. W. J. Beenakker. Boundary conditions for Dirac fermions on a terminated honeycomb lattice. *Phys. Rev. B* 77 085423 (2008).
- [172] Y. Niimi T. Matsui H. Kambara K. Tagami M. Tsukada and H. Fukuyama. Scanning tunneling microscopy and spectroscopy studies of graphite edges. *App. Surf. Sci.* 241 43-48 (2005).
- [173] S. Banerjee M. Sardar N. Gayathri A. K. Tyagi and B. Raj. Conductivity landscape of highly oriented pyrolytic graphite surfaces containing ribbons and edges. *Phys. Rev. B* 72 075418 (2005).
- [174] Y. Niimi T. Matsui H. Kambara K. Tagami M. Tsukada and H. Fukuyama. Scanning tunneling microscopy and spectroscopy of the electronic local density of states of graphite surfaces near monoatomic step edges. *Phys. Rev. B* 73 085421 (2006).
- [175] S. Okada and A. Oshiyama. Magnetic ordering in hexagonally bonded sheets with first-row elements. *Phys. Rev. Lett.* 87 146803 (2001).
- [176] Y.-W. Son M. L. Cohen and S. G. Louie. Half-metallic graphene nanoribbons. *Nature* 444 347-349 (2006).
- [177] F. Muñoz Rojas D. Jacob J. Fernández-Rossier and J. J. Palacios. Coherent transport in graphene nanoconstrictions. *Phys. Rev. B* 74 195417 (2006).
- [178] T. C. Li and S.-P. Lu. Quantum conductance of graphene nanoribbons with edge defects. *Phys. Rev. B* 77 085408 (2008).
- [179] L. P. ZARbo and B. K. Nikolic. Spatial distribution of local currents of massless Dirac fermions in quantum transport through graphene nanoribbons. *Europhys. Lett.* 80 47001 (2007).
- [180] D. Areshkin D. Gunlycke and C. White. Ballistic transport in graphene nanostrips in the presence of disorder: Importance of edge effects. *Nano Lett.* 7 204-210 (2007).
- [181] K. Wakabayashi Y. Takane and M. Sigrist. Perfectly conducting channel and universality crossover in disordered graphene nanoribbons. *Phys. Rev. Lett.* 99 036601 (2007).
- [182] N. M. R. Peres F. Guinea and A. H. C. Neto. Electronic properties of disordered two-dimensional carbon. *Phys. Rev. B* 73 125411 (2006).
- [183] K. Sasaki S. Murakami and R. Saito. Stabilization mechanism of edge states in graphene. *Appl. Phys. Lett.* 88 113110 (2006).
- [184] D. Porezag T. Frauenheim T. Köhler G. Seifert and R. Kaschner. Construction of tight-binding-like potentials on the basis of density-functional theory: Application to carbon. *Phys. Rev. B* 51 12947-12957 (1995).
- [185] L. Brey and H. A. Fertig. Electronic states of graphene nanoribbons studied with the Dirac equation. *Phys. Rev. B* 73 235411 (2006).
- [186] M. V. Berry and R. J. Mondragon. Neutrino billiards: time-reversal symmetry-breaking without magnetic fields. *Proc. R. Soc. Lond. A* 412 53-74 (1987).

- [187] E.-N. Foo M. F. Thorpe and D. Weaire. Effective surface potential method for calculating surface states. *Surf. Sci.* 57 323-347 (1976).
- [188] A. R. Akhmerov J. H. Bardarson A. Rycerz and C. W. J. Beenakker. Theory of the valley-valve effect in graphene nanoribbons. *Phys. Rev. B* 77 205416 (2008).
- [189] A. Rycerz J. Tworzydło and C. W. J. Beenakker. Valley filter and valley valve in graphene. *Nat. Phys.* 3 172-175 (2007).
- [190] S. Reich J. Maultzsch C. Thomsen and P. Ordejón. Tight-binding description of graphene. *Phys. Rev. B* 66 035412 (2002).
- [191] Y. Yao F. Ye X.-L. Qi S.-C. Zhang and Z. Fang. Spin-orbit gap of graphene: First-principles calculations. *Phys. Rev. B* 75 041401 (2007).
- [192] G. Schmidt D. Ferrand L. W. Molenkamp A. T. Filip and B. J. van Wees. Fundamental obstacle for electrical spin injection from a ferromagnetic metal into a diffusive semiconductor. *Phys. Rev. B* 62 R4790-R4793 (2000).
- [193] E. I. Rashba. Theory of electrical spin injection: Tunnel contacts as a solution of the conductivity mismatch problem. *Phys. Rev. B* 62 R16267-R16270 (2000).
- [194] A. Fert and H. Jaffrès. Conditions for efficient spin injection from a ferromagnetic metal into a semiconductor. *Phys. Rev. B* 64 184420 (2001).
- [195] M. Wimmer I. Adagideli S. Berber D. Tomnek and K. Richter. Spin currents in rough graphene nanoribbons: Universal fluctuations and spin injection. *Phys. Rev. Lett.* 100 177207 (2008).
- [196] S. Berber. Unpublished.
- [197] L. Pisani J. A. Chan B. Montanari and N. M. Harrison. Electronic structure and magnetic properties of graphitic ribbons. *Phys. Rev. B* 75 064418 (2007).
- [198] E. H. Lieb. Two theorems on the hubbard model. *Phys. Rev. Lett.* 62 1201-1204 (1989).
- [199] E. C. Stoner. Collective electron ferromagnetism. II. Energy and specific heat. *Proc. R. Soc. London. Ser. A* 169 339-371 (1939).
- [200] N. D. Mermin and H. Wagner. Absence of ferromagnetism or antiferromagnetism in one- or two-dimensional isotropic Heisenberg models. *Phys. Rev. Lett.* 17 1133-1136 (1966).
- [201] O. V. Yazyev and M. I. Katsnelson. Magnetic correlations at graphene edges: Basis for novel spintronics devices. *Phys. Rev. Lett.* 100 047209 (2008).
- [202] D.-E. Jiang B. G. Sumpter and S. Dai. First principles study of magnetism in nanographenes. *J. Chem. Phys.* 127 124703 (2007).
- [203] B. Huang F. Liu J. Wu B.-L. Gu and W. Duan. Suppression of spin polarization in graphene nanoribbons by edge defects and impurities. *Phys. Rev. B* 77 153411 (2008).
- [204] S. Bhowmick and V. B. Shenoy. Edge state magnetism of single layer graphene nanostructures. *J. Chem. Phys.* 128 244717 (2008).
- [205] O. Hod V. Barone and G. E. Scuseria. Half-metallic graphene nanodots: A comprehensive first-principles theoretical study. *Phys. Rev. B* 77 035411 (2008).
- [206] P. W. Anderson D. J. Thouless E. Abrahams and D. S. Fisher. New method for a scaling theory of localization. *Phys. Rev. B* 22 3519-3526 (1980).
- [207] C. W. J. Beenakker. Random-matrix theory of quantum transport. *Rev. Mod. Phys.* 69 731-808 (1997).
- [208] O. N. Dorokhov. Transmission coefficient and the localization length of an electron in N bound disordered chains. *JETP Lett.* 36 318-321 (1982).
- [209] P. A. Mello P. Pereyra and N. Kumar. Macroscopic approach to multichannel disordered conductors. *Ann. Phys. (N.Y.)* 181 290-317 (1988).
- [210] M. E. Gertsenshtein and V. B. Vasil'ev. Waveguides with random inhomogeneities and brownian motion in the lobachevsky plane. *Theory Probab. Appl.* 4 391-398 (1959).
- [211] C. W. J. Beenakker and J. A. Melsen. Conductance fluctuations weak localization and shot noise for a ballistic constriction in a disordered wire. *Phys. Rev. B* 50 2450-2457 (1994).

- [212] J. Repp. Private communication.
- [213] I. Adagideli J. H. Bardarson and P. Jacquod. Electrical probing of the spin conductance of mesoscopic cavities. arXiv:0808.2756v1 (2008).
- [214] I. Adagideli G. E. W. Bauer and B. I. Halperin. Detection of currentinduced spins by ferromagnetic contacts. Phys. Rev. Lett. 97 256601 (2006).
- [215] M. Johnson and R. H. Silsbee. Coupling of electronic charge and spin at a ferromagnetic-paramagnetic metal interface. Phys. Rev. B 37 5312-5325 (1988).
- [216] J. Tworzydeo B. Trauzettel M. Titov A. Rycerz and C. W. J. BeEnAKKEr. Sub-poissonian shot noise in graphene. Phys. Rev. Lett. 96 246802 (2006).
- [217] D. XiaO W. YaO and Q. Niu. Valley-contrasting physics in graphene: Magnetic moment and topological transport. Phys. Rev. Lett. 99 236809 (2007).
- [218] J. Schliemann. Spin Hall effect. Int. J. Mod. Phys. B 20 1015-1036 (2006).
- [219] B. K. Nikoli L. P. ZArbo and S. Souma. Mesoscopic spin Hall effect in multiprobe ballistic spin-orbit-coupled semiconductor bridges. Phys. Rev. B 72 075361 (2005).
- [220] J. H. Bardarson I. Adagideli and P. Jacquod. Mesoscopic spin Hall effect. Phys. Rev. Lett. 98 196601 (2007).
- [221] J. M. Luttinger and W. Kohn. Motion of electrons and holes in perturbed periodic fields. Phys. Rev. 97 869-883 (1955).
- [222] M. Head-Gordon P. E. Maslen and C. A. White. A tensor formulation of many-electron theory in a nonorthogonal single-particle basis. J. Chem. Phys. 108 616-625 (1998).
- [223] L. A. Ponomarenko F. Schedin M. I. Katsnelson R. Yang E. W. Hill K. S. Novoselov and A. K. Geim. Chaotic Dirac billiard in graphene quantum dots. Science 320 356-358 (2008).
- [224] E. McCann K. Kechedzhi V. I. Fal'ko H. Suzuura T. Ando and B. L. Altshuler. Weak-localization magnetoresistance and valley symmetry in graphene. Physical Review Letters 97 146805 (2006).
- [225] S. V. Morozov K. S. Novoselov M. I. Katsnelson F. Schedin L. A. Ponomarenko D. Jiang and A. K. Geim. Strong suppression of weak localization in graphene. Phys. Rev. Lett. 97 016801 (2006).
- [226] X. Wu X. li Z. Song C. Berger and W. A. de Heer. Weak antilocalization in epitaxial graphene: Evidence for chiral electrons. Phys. Rev. Lett. 98 136801 (2007).
- [227] F. V. Tikhonenko D. W. Horsell R. V. Gorbachev and A. K. Savchenko. Weak localization in graphene flakes. Phys. Rev. Lett. 100 056802 (2008).
- [228] I. Adagideli. Private communication.
- [229] F. Schwabl. Quantum Mechanics (Springer Berlin 2003).
- [230] A. Cresti R. Farchioni G. Grosso and G. P. Parravicini. KeldyshGreen function formalism for current profiles in mesoscopic systems. Phys. Rev. B 68 075306 (2003).
- [231] A. D. Stone and A. Szafer. What is measured when you measure a resistance? - the Landauer forumula revisited. IBM J. Res. Dev. 32 384-413 (1988).
- [232] A. Garcia. Numerical Methods for Physics (Prentice Hall Upper Saddle River NJ 1999).
- [233] R. Peierls. Zur Theorie des Diamagnetismus von Leitungselektronen. Z. f. Phys. A 80 763-791 (1933).
- [234] T. L. Li and K. J. Kunn. Effects of mass discontinuity on the numerical solutions to quantum wells using the effective mass equation. J. Comp. Phys. 110 292-300 (1994).
- [235] D. J. BenDaniel and C. B. Duke. Space-charge effects on electron tunneling. Phys. Rev. 152 683-692 (1966).
- [236] G. Dresselhaus. Spin-orbit coupling effects in zinc blende structures. Phys. Rev. 100 580-586 (1955).
- [237] U. ROssler. Solid State Theory (Springer Berlin Heidelberg New York 2004).
- [238] P. R. Wallace. The band theory of graphite. Phys. Rev. 71 622-634 (1947).
- [239] J. C. Slater and G. F. Koster. Simplified LCAO method for the periodic potential

problem. Phys. Rev. 94 1498-1524 (1954).

[240] D. Finkenstadt G. Pennington and M. J. Mehl. From graphene to graphite: A general tight-binding approach for nanoribbon carrier transport. Phys. Rev. B 76 121405 (2007).

[241] H. Raza and E. Kan. An extended Hckel theory based atomistic model for graphene nanoelectronics. To appear in J. Comput. Electron. (2008).

[242] F. D. M. Haldane. Model for a quantum Hall effect without Landau levels: Condensed-matter realization of the "parity anomaly". Phys. Rev. Lett. 61 20152018 (1988).

[243] R. Saito G. Dresselhaus and M. S. Dresselhaus. Physical Properties of Carbon Nanotubes (Imperial College Press London 1998).

[244] J. W. McClure. Diamagnetism of graphite. Phys. Rev. 104 666-671 (1956).

[245] D. P. DiVincenzo and E. J. Mele. Self-consistent effective-mass theory for intralayer screening in graphite intercalation compounds. Phys. Rev. B 29 16851694 (1984).

[246] M. E. Peskin and D. V. Schroeder. An Introduction to Quantum Field Theory (Perseus Books Cambridge MA 1995).

From Diagram technique for nonequilibrium processes by Keldysh

¹ T. Matsubara Progr. Theoret. Phys. (Kyoto) 14 351 (1955).

² Abrikosov Gor'kov and Dzyaloshinskii JETP 36 900 (1959) Soviet Phys. JETP 9 636 (1959); Metody kvantovo teorii polya v statisticheskoy fizike (Methods of quantum field theory in statistical physics) Fizmatgiz 1962.

³ E. S. Fradkin JETP 36 1286 (1959) Soviet Phys. JETP 9 912 (1959); Nuclear Phys. 12 465 (1959).

⁴ J. M. Luttinger and J. C. Ward Phys. Rev. 118 1417 (1960).

⁵ P. C. Martin and J. Schwinger Phys. Rev. 115 1342 (1959).

⁶ V. L. Bonch-Bruевич and Sch. M. Kogan Ann. Phys. 9 125 (1960) V. L. Bonch-Bruевич and S. V. Tyablikov Metod funktsii Grina v statisticheskoy fizike (The method of Green's functions in statistical physics) Fizmatgiz 1961. ⁷ A. Salam Progr. Theoret. Phys. (Kyoto) 9 550 (1953).

⁸ V. L. Bonch-Bruевич JETP 28 121 (1955); 30 343 (1956) Soviet Phys. JETP 1 169 (1955); 3 278 (1956).

⁹ V. M. Galitski and A. B. Migdal JETP 34 139 (1957) Soviet Phys. JETP 7 96 (1958).

¹⁰ O. V. Konstantinov and V. I. Perel' JETP 39 197 (1960) Soviet Phys. JETP 12 142 (1961).

¹¹ I. E. Dzyaloshinski JETP 42 1126 (1962) Soviet Phys. JETP 15 778 (1962).

¹² R. Mills Preprint 1962.

¹³ N. N. Bogolyubov Lektsii po kvantovo statistike (Lectures on Quantum Statistics) Kiev Publishing House "Sovetskaya Shkola" (Soviet School) 1949.

¹⁴ M. Gell-Mann and F. Low Phys. Rev. 84 350 (1951).

¹⁵ D. A. Kirzhnits Polevye metody teorii mnogikh chastits (Field Methods in Many Body Theory) Atomizdat 1963.

¹⁶ J. E. Moyal Proc. Cambridge Phil. Soc. 45 99 (1949).

¹⁷ V. P. Silin and A. A. Rukhadze lektromagnitnye svostva plazmy i plazmopodobnykh sred (Electromagnetic Properties of Plasma and Plasmalike Media) Atomizdat 1961.

¹⁸ L. Van Hove Physica 23 441 (1957).

¹⁹ J. Prigogine and P. Resibois Physica 27 629

(1961). P. Resibois Physica 29 721 (1963).

²⁰ S. Fujita Physica 28 281 (1962).

Translated by G. Volkoff 211

References from The Non-Equilibrium Green Function (NEGF) Method from Kerem Y. Camsari Shuvro Chowdhury and Supriyo Datta

References

- [1] P.C. Martin J. Schwinger Physical Review **115**(6) 1342 (1959)
- [2] L. Kadanoff G. Baym *Quantum Statistical Mechanics Benjamin New York (1962)*
- [3] L.V. Keldysh et al. Sov. Phys. JETP **20**(4) 1018 (1965)
- [4] P. Danielewicz Annals of Physics **152**(2) 239 (1984)
- [5] G.D. Mahan Physics Reports **145**(5) 251 (1987)
- [6] S. Datta Physical Review B **40**(8) 5830 (1989)
- [7] M.J. McLennan Y. Lee S. Datta Phys. Rev. B **43** 13846 (1991)
- [8] Y. Meir N.S. Wingreen Physical review letters **68**(16) 2512 (1992)
- [9] G. Klimeck R. Lake R.C. Bowen W.R. Frensley T.S. Moise Applied Physics Letters **67**(17) 2539 (1995)
- [10] R.C. Bowen G. Klimeck R. Lake W.R. Frensley T. Moise J. Appl. Phys **81** 3207 (1997)
- [11] R. Lake G. Klimeck R.C. Bowen D. Jovanovic Journal of Applied Physics **81**(12) 7845 (1997)
- [12] A. Martinez M. Bescond J.R. Barker A. Svizhenko M. Anantram C. Millar A. Asenov IEEE Transactions on Electron Devices **54**(9) 2213 (2007)
- [13] S. Datta in *Digest. International Electron Devices Meeting (IEEE 2002)* pp. 703–706
- [14] N. Seoane A. Martinez A.R. Brown J.R. Barker A. Asenov IEEE Transactions on electron devices **56**(7) 1388 (2009)
- [15] J. Wang E. Polizzi M. Lundstrom Journal of Applied Physics **96**(4) 2192 (2004)
- [16] N. Mingo L. Yang Phys. Rev. B **68** 245406 (2003)
- [17] B.K. Nikolić L.P. Zârbo S. Souma Physical Review B **73**(7) 075303 (2006)
- [18] T. Kubis P. Vogl Physical Review B **83**(19) 195304 (2011)
- [19] J.S. Wang J. Wang N. Zeng Physical Review B **74**(3) 033408 (2006)
- [20] S.O. Koswatta S. Hasan M.S. Lundstrom M. Anantram D.E. Nikonov IEEE Transactions on Electron Devices **54**(9) 2339 (2007)
- [21] M. Pourfath *The Non-Equilibrium Green's Function Method for Nanoscale Device Simulation* (Springer 2014)
- [22] J. Guo S. Datta M. Anantram M. Lundstrom Journal of Computational Electronics **3**(3-4) 373 (2004)
- [23] A. Martinez N. Seoane A.R. Brown J.R. Barker A. Asenov IEEE Transactions on Nanotechnology **8**(5) 603 (2009)

- [24] M. Di Ventra *Electrical transport in nanoscale systems* (Cambridge University Press Cambridge UK 2008)
- [25] T.T. Heikkilä *The physics of nanoelectronics: transport and fluctuation phenomena at low temperatures* vol. 21 (Oxford University Press 2013)
- [26] S. Datta in *2015 International Workshop on Computational Electronics (IWCE)* (IEEE 2015) pp. 1–6
- [27] S. Datta *Electronic transport in mesoscopic systems* (Cambridge University Press 1995)
- [28] S. Datta *Quantum transport: atom to transistor* (Cambridge university press 2005)
- [29] S. Datta *Lessons from nanoelectronics: a new perspective on transport* vol. 1 (World Scientific publishing company 2012)
- [30] A. Ghosh *Nanoelectronics: a molecular view* vol. 13 (World Scientific Publishing Company 2016)
- [31] B. Van Wees H. Van Houten C. Beenakker J.G. Williamson L. Kouwenhoven D. Van der Marel C. Foxon *Physical Review Letters* **60**(9) 848 (1988)
- [32] B. Xu N.J. Tao *science* **301**(5637) 1221 (2003)
- [33] M.V. Fischetti S.E. Laux *Phys. Rev. B* **38** 9721 (1988). doi:[10.1103/PhysRevB.38.9721](https://doi.org/10.1103/PhysRevB.38.9721). URL <https://link.aps.org/doi/10.1103/PhysRevB.38.9721>
- [34] K.Y. Camsari S. Ganguly D. Datta S. Datta *IEEE Transactions on Nanotechnology* **18** 858 (2019)
- [35] B. Altshuller A. Aronov D. Khmelnitsky *Solid State Communications* **39**(5) 619 (1981). doi:[https://doi.org/10.1016/0038-1098\(81\)90336-7](https://doi.org/10.1016/0038-1098(81)90336-7). URL <http://www.sciencedirect.com/science/article/pii/0038109881903367>
- [36] R. Golizadeh-Mojarad S. Datta *Physical Review B* **75**(8) 081301 (2007)
- [37] S. Datta arXiv preprint arXiv:0809.4460 (2008)
- [38] S. Hershfield *Physical Review B* **43**(14) 11586 (1991)
- [39] I. Žutić J. Fabian S.D. Sarma *Reviews of modern physics* **76**(2) 323 (2004)
- [40] J.R. Sánchez L. Vila G. Desfonds S. Gambarelli J. Attané J. De Teresa C. Magén A. Fert *Nature communications* **4**(1) 1 (2013)
- [41] A. Caviglia M. Gabay S. Gariglio N. Reyren C. Cancellieri J.M. Triscone *Physical review letters* **104**(12) 126803 (2010)
- [42] E. Lesne Y. Fu S. Oyarzun J. Rojas-Sánchez D. Vaz H. Naganuma G. Sicoli J.P. Attané M. Jamet E. Jacquet et al. *Nature materials* **15**(12) 1261 (2016)
- [43] M.B. Jungfleisch Q. Zhang W. Zhang J.E. Pearson R.D. Schaller H. Wen A. Hoffmann *Physical review letters* **120**(20) 207207 (2018)
- [44] J. Tian I. Miotkowski S. Hong Y.P. Chen *Scientific reports* **5** 14293 (2015)

- [45] I.M. Miron G. Gaudin S. Auffret B. Rodmacq A. Schuhl S. Pizzini J. Vogel P. Gambardella Nature materials **9**(3) 230 (2010)
- [46] L. Liu C.F. Pai Y. Li H. Tseng D. Ralph R. Buhrman Science **336**(6081) 555 (2012)
- [47] S. Datta B. Das Applied Physics Letters **56**(7) 665 (1990)
- [48] J. Nitta T. Akazaki H. Takayanagi T. Enoki Physical Review Letters **78**(7) 1335 (1997)
- [49] H.C. Koo J.H. Kwon J. Eom J. Chang S.H. Han M. Johnson Science **325**(5947) 1515 (2009)
- [50] J. Wunderlich B.G. Park A.C. Irvine L.P. Zârbo E. Rozkotová P. Nemec V. Novák J. Sinova T. Jungwirth Science **330**(6012) 1801 (2010)
- [51] W.Y. Choi H.j. Kim J. Chang S.H. Han H.C. Koo M. Johnson Nature nanotechnology **10**(8) 666 (2015)
- [52] D. Steiauf M. Fähnle Physical Review B **79**(14) 140401 (2009)
- [53] V. Mourik K. Zuo S.M. Frolov S.R. Plissard E.P.A.M. Bakkers L.P. Kouwenhoven Science **336**(6084) 1003 (2012). doi:[10.1126/science.1222360](https://doi.org/10.1126/science.1222360)
- [54] R.M. Lutchyn J.D. Sau S. Das Sarma Phys. Rev. Lett. **105** 077001 (2010). doi:[10.1103/PhysRevLett.105.077001](https://doi.org/10.1103/PhysRevLett.105.077001)
- [55] H. Zhang C.X. Liu S. Gazibegovic D. Xu J.A. Logan G. Wang N. van Loo J.D.S. Bommer M.W.A. de Moor D. Car R.L.M. Op het Veld P.J. van Veldhoven S. Koelling M.A. Verheijen M. Pendharkar D.J. Pennachio B. Shojaei J.S. Lee C.J. Palmstrøm E.P.A.M. Bakkers S.D. Sarma L.P. Kouwenhoven Nature **556**(7699) 74 (2018). doi:[10.1038/nature26142](https://doi.org/10.1038/nature26142)
- [56] H. Zhang D.E. Liu M. Wimmer L.P. Kouwenhoven Nature Communications **10**(1) 5128 (2019). doi:[10.1038/s41467-019-13133-1](https://doi.org/10.1038/s41467-019-13133-1)
- [57] H. Ren F. Pientka S. Hart A.T. Pierce M. Kosowsky L. Lunczer R. Schlereth B. Scharf E.M. Hankiewicz L.W. Molenkamp B.I. Halperin A. Yacoby Nature **569**(7754) 93 (2019)
- [58] M.P. Samanta S. Datta Phys. Rev. B **57** 10972 (1998). doi:[10.1103/PhysRevB.57.10972](https://doi.org/10.1103/PhysRevB.57.10972)
- [59] S. Datta P.F. Bagwell Superlattices and Microstructures **25**(5) 1233 (1999). doi:<https://doi.org/10.1006/spmi.1999.0747>
- [60] P. San-Jose J. Cayao E. Prada R. Aguado New Journal of Physics **15**(7) 075019 (2013)
- [61] A. Levy Yeyati A. Martín-Rodero F.J. García-Vidal Phys. Rev. B **51** 3743 (1995). doi:[10.1103/PhysRevB.51.3743](https://doi.org/10.1103/PhysRevB.51.3743)
- [62] P. Sriram S.S. Kalantre K. Gharavi J. Baugh B. Muralidharan Phys. Rev. B **100** 155431 (2019). doi:[10.1103/PhysRevB.100.155431](https://doi.org/10.1103/PhysRevB.100.155431)
- [63] M. Fischetti L. Wang B. Yu C. Sachs P. Asbeck Y. Taur M. Rodwell in *2007 IEEE International Electron Devices Meeting* (IEEE 2007) pp. 109–112
- [64] S. Salahuddin M. Lundstrom S. Datta IEEE Transactions on Electron Devices **52**(8) 1734 (2005)
- [65] S. Sayed S. Hong S. Datta Physical Review Applied **10**(5) 054044 (2018)

- [66] R.A. Webb S. Washburn C. Umbach R. Laibowitz Physical Review Letters **54**(25) 2696 (1985)
- [67] R.K. Kumar D. Bandurin F. Pellegrino Y. Cao A. Principi H. Guo G. Auton M.B. Shalom L.A. Ponomarenko G. Falkovich et al. Nature Physics **13**(12) 1182 (2017)
- [68] Y. Kajiwar K. Harii S. Takahashi J.i. Ohe K. Uchida M. Mizuguchi H. Umezawa H. Kawai K. Ando K. Takanashi et al. Nature **464**(7286) 262 (2010)
- [69] S. Sayed V.Q. Diep K.Y. Camsari S. Datta Scientific reports **6** 28868 (2016)
- [70] N.A. Zimbovskaya Phys. Rev. B **78** 035331 (2008). doi:[10.1103/PhysRevB.78.035331](https://doi.org/10.1103/PhysRevB.78.035331)

References from “Green’s functions of quasi-one-dimensional layered systems and their application to Josephson junctions” by Piasotski Pletyukhov Shnirman

References

- [1] S. Datta *Electronic transport in mesoscopic systems* (Cambridge university press 1997).
- [2] J. K. Freericks *Transport in multilayered nanostructures: the dynamical mean-field theory approach* (World Scientific 2016).
- [3] C. W. Groth M. Wimmer A. R. Akhmerov and X. Waintal *New J. Phys.* **16** 063065 (2014) .
- [4] T. Ando *Phys. Rev. B* **44** 8017 (1991) .
- [5] M. Cheng and R. M. Lutchyn *Phys. Rev. B* **86** 134522 (2012) .
- [6] T. Yokoyama M. Eto and Y. V. Nazarov *Phys. Rev. B* **89** 195407 (2014) .
- [7] R. Landauer *IBM J. Res. Dev.* **1** 223 (1957) .
- [8] R. Landauer *Philos. Mag.* **21** 863 (1970) .
- [9] C. Caroli R. Combescot D. Lederer P. Nozieres and D. Saint-James *J. Phys. C: Solid State Phys.* **4** 2598 (1971) .
- [10] P. A. Khomyakov G. Brocks V. Karpan M. Zwierzycki and P. J. Kelly *Phys. Rev. B* **72** 035450 (2005) .
- [11] D. S. Fisher and P. A. Lee *Phys. Rev. B* **23** 6851 (1981) .
- [12] H. Boumrar M. Hamidi H. Zenia and S. Lounis *J. Phys. Condens. Matter* **32** 355302 (2020) .
- [13] G. E. Blonder M. Tinkham and T. M. Klapwijk *Phys. Rev. B* **25** 4515 (1982) .
- [14] C. W. J. Beenakker *Phys. Rev. Lett.* **67** 3836 (1991) .
- [15] A. Andreev *Sov. Phys. JETP* **19** 1228 (1964) .
- [16] A. Andreev *Sov. Phys. JETP* **22** 18 (1966) .
- [17] G. B. Arnold *Phys. Rev. B* **18** 1076 (1978) .

- [18] G. B. Arnold *J. Low Temp. Phys.* **59** 143 (1985) .
- [19] T. E. Feuchtwang *Phys. Rev. B* **10** 4121 (1974) .
- [20] T. E. Feuchtwang *Phys. Rev. B* **10** 4135 (1974) .
- [21] T. E. Feuchtwang *Phys. Rev. B* **12** 3979 (1975) .
- [22] A. A. Golubov M. Y. Kupriyanov and E. Il'ichev *Rev. Mod. Phys.* **76** 411 (2004) .
- [23] M. Eschrig *Philosophical Transactions of the Royal Society A* **376** 20150149 (2018) .
- [24] F. García-Moliner and J. Rubio *J. Phys. C: Solid State Phys.* **2** 1789 (1969) .
- [25] F. García-Moliner V. Heine and J. Rubio *J. Phys. C: Solid State Phys.* **2** 1797 (1969) .
- [26] F. García-Moliner and V. Velasco *Phys. Rep.* **200** 83 (1991) .
- [27] F. García-Moliner and V. Velasco *Surf. Sci.* **299-300** 332 (1994) .
- [28] H. Rodriguez-Coppola V. R. Velasco F. Garcia-Moliner and R. Pérez-Alvarez *Phys. Scr.* **42** 115 (1990) .
- [29] J. E. Inglesfield *J. Phys. C: Solid State Phys.* **4** L14 (1971) .
- [30] J. E. Inglesfield *J. Phys. C: Solid State Phys.* **14** 3795 (1981) .
- [31] D. Wortmann H. Ishida and S. Blügel *Phys. Rev. B* **65** 165103 (2002) .
- [32] B. van Heck J. I. Väyrynen and L. I. Glazman *Phys. Rev. B* **96** 075404 (2017) .
- [33] C. Murthy V. D. Kurilovich P. D. Kurilovich B. van Heck L. I. Glazman and C. Nayak *Phys. Rev. B* **101** 224501 (2020) .
- [34] R. M. Lutchyn J. D. Sau and S. Das Sarma *Phys. Rev. Lett.* **105** 077001 (2010) .
- [35] Y. Oreg G. Refael and F. von Oppen *Phys. Rev. Lett.* **105** 177002 (2010) .
- [36] E. Akkermans A. Auerbach J. E. Avron and B. Shapiro *Phys. Rev. Lett.* **66** 76 (1991) .
- [37] S. Souma and A. Suzuki *Phys. Rev. B* **65** 115307 (2002) .
- [38] Note that the heterostructural Green's function $G(xx')$ is a highly non-trivial object containing information about all possible multi-barrier scattering phenomena and interface-localized bound states.
- [39] K. Piasotski M. Pletyukhov and A. Shnirman In preparation .
- [40] J.-M. Lévy-Leblond *Phys. Rev. A* **52** 1845 (1995) .
- [41] K. Piasotski N. Müller D. M. Kennes H. Schoeller and M. Pletyukhov *Phys. Rev. B* **106** 165405 (2022) .
- [42] In real experiments the phase difference is typically varied by closing the system into a circular geometry far away from the tunneling region and varying the magnetic flux threading the resulting ring.

- [43] A. A. Abrikosov *Fundamentals of the Theory of Metals* (Courier Dover Publications 2017).
- [44] C. Chamon R. Jackiw Y. Nishida S.-Y. Pi and L. Santos *Phys. Rev. B* **81** 224515 (2010) .
- [45] W. Haberkorn H. Knauer and J. Richter *Phys. Status Solidi (a)* **47** K161 (1978) .
- [46] M. Žonda V. Pokorný V. Janiš and T. Novotný *Sci. Rep.* **5** 8821 (2015) .
- [47] A. Altland and B. D. Simons *Condensed matter field theory* (Cambridge university press 2010).
- [48] V. Meden *J. Phys. Condens. Matter* **31** 163001 (2019) .
- [49] J. W. A. Robinson S. Piano G. Burnell C. Bell and M. G. Blamire *Phys. Rev. B* **76** 094522 (2007) .
- [50] N. M. Chtchelkatchev W. Belzig Y. V. Nazarov and C. Bruder *JETP Letters* **74** 323 (2001) .
- [51] Note that one can envisage the situation in which the two bands touch one another at exactly zero energy. In this case however the states around zero energy are expected to disperse quadratically as opposed to the linear dispersion found in the topological regime.
- [52] F. J. Matute-Cañadas C. Metzger S. Park L. Tosi P. Krogstrup J. Nygård M. F. Goffman C. Urbina H. Pothier and A. L. Yeyati *Phys. Rev. Lett.* **128** 197702 (2022) .
- [53] C. R. Harris K. J. Millman S. J. van der Walt R. Gommers P. Virtanen D. Cournapeau E. Wieser J. Taylor S. Berg N. J. Smith R. Kern M. Picus S. Hoyer M. H. van Kerkwijk M. Brett A. Haldane J. F. del Río M. Wiebe P. Peterson P. Gérard-Marchant K. Sheppard T. Reddy W. Weckesser H. Abbasi C. Gohlke and T. E. Oliphant *Nature* **585** 357 (2020) .
- [54] M. Wimmer *Quantum transport in nanostructures: From computational concepts to spintronics in graphene and magnetic tunnel junctions* *Ph.D. thesis* University of Regensburg (2009).
- [55] K. Piasotski *On the universal properties of excess charges in one dimension* *Ph.D. thesis* RWTH Aachen University (2023).
- [56] A. Zazunov A. Iks M. Alvarado A. L. Yeyati and R. Egger *Beilstein J. Nanotechnol.* **9** 1659 (2018) .
- [57] M. Alvarado A. Iks A. Zazunov R. Egger and A. L. Yeyati *Phys. Rev. B* **101** 094511 (2020) .
- [58] V. Kaladzhyan and C. Bena *Phys. Rev. B* **100** 081106 (2019) .
- [59] S. Pinon V. Kaladzhyan and C. Bena *Phys. Rev. B* **101** 115405 (2020) .
- [60] A. Komnik and S. Heinze *Phys. Rev. B* **96** 155103 (2017) .
- [61] A. Zazunov R. Egger M. Alvarado and A. L. Yeyati *Phys. Rev. B* **96** 024516 (2017) .
- [62] Y. Peng Y. Bao and F. von Oppen *Phys. Rev. B* **95** 235143 (2017) .

- [63] A. Zazunov R. Egger and A. Levy Yeyati *Phys. Rev. B* **94** 014502 (2016) .
- [64] L. Arrachea G. S. Lozano and A. A. Aligia *Phys. Rev. B* **80** 014425 (2009) .
- [65] N. Müller K. Piasotski D. M. Kennes H. Schoeller and M. Pletyukhov *Phys. Rev. B* **104** 125447 (2021) .

References from Boundary Green's function approach for spinful single-channel and multichannel Majorana nanowires

References

- [1] J. Alicea Rep. Prog. Phys. **75** 076501 (2012).
- [2] M. Leijnse and K. Flensberg Semicond. Sci. Techn. **27** 124003 (2012).
- [3] C. W. J. Beenakker Annu. Rev. Con. Mat. Phys. **4** 113 (2013).
- [4] S. Das Sarma M. Freedman and C. Nayak npj Quantum Inf. **1** 15001 (2015).
- [5] R. Aguado Riv. Nuovo Cim **40** 523 (2017).
- [6] M. Sato and Y. Ando Rep. Prog. Phys. **80** 076501 (2017).
- [7] R. M. Lutchyn E. P. A. M. Bakkers L. P. Kouwenhoven P. Krogstrup C. M. Marcus and Y. Oreg Nat. Rev. Mater. **3** 52 (2018).
- [8] H. Zhang D. E. Liu M. Wimmer and L. P. Kouwenhoven Nat. Comm. **10** 5128 (2019).
- [9] V. Mourik K. Zuo S. M. Frolov S. R. Plissard E. P. A. Bakkers and L. P. Kouwenhoven Science **336** 1003 (2012).
- [10] S. M. Albrecht A. P. Higginbotham M. Madsen F. Kuemmeth T. S. Jespersen J. Nygård P. Krogstrup and C. M. Marcus Nature **531** 206 (2016).
- [11] M. T. Deng S. Vaitiekenas E. B. Hansen J. Danon M. Leijnse K. Flensberg J. Nygård P. Krogstrup and C. M. Marcus Science **354** 1557 (2016).
- [12] F. Nichele A. C. C. Drachmann A. M. Whiticar E. C. T. O'Farrell H. J. Suominen A. Fornieri T. Wang G. C. Gardner C. Thomas A. T. Hatke P. Krogstrup M. J. Manfra K. Flensberg and C. M. Marcus Phys. Rev. Lett. **119** 136803 (2017).
- [13] H. J. Suominen M. Kjaergaard A. R. Hamilton J. Shabani C. J. Palmstrøm C. M. Marcus and F. Nichele Phys. Rev. Lett. **119** 176805 (2017).
- [14] S. Gazibegovich D. Car H. Zhang S. C. Balk J. A. Logan M. W. A. de Moor M. C. Cassidy R. Schmits D. Xu G. Wang P. Krogstrup R. L. M. Op het Veld J. Shen D. Bouman B. Shojaei D. Pennachio J. S. Lee P. J. van Veldhoven S. Koelling M. A. Verheijen L. P. Kouwenhoven C. J. Palmstrøm and E. P. A. M. Bakkers Nature **548** 434 (2017).
- [15] H. Zhang C. X. Liu S. Gazibegovic D. Xu J. A. Logan G. Wang N. van Loo J. D. S. Bommer M. W. A. de Moor D. Car R. L. M. Op het Veld P. J. van Veldhoven S. Koelling M. A. Verheijen M. Pendharkar D. J. Pennachio B. Shojaei J. S. Lee C. J. Palmstrom E. P. A. M. Bakkers S. Das Sarma and L. P. Kouwenhoven Nature **556** 74 (2018).

-
- [16] A. Fornieri A. M. Whiticar F. Setiawan E. P. Marín A. C. C. Drachmann A. Keselman S. Gronin C. Thomas T. Wang R. Kallaher G. C. Gardner E. Berg M. J. Manfra A. Stern C. M. Marcus and F. Nichele *Nature* **569** 89 (2019).
- [17] S. Vaitiekėnas M. T. Deng P. Krogstrup and C. M. Marcus arXiv:1809.05513.
- [18] C. L. M. Wong and K. T. Law *Phys. Rev. B* **86** 184516 (2012).
- [19] S. Nakosai Y. Tanaka and N. Nagaosa *Phys. Rev. Lett.* **108** 147003 (2012).
- [20] S. Deng L. Viola and G. Ortiz *Phys. Rev. Lett.* **108** 036803 (2012).
- [21] F. Zhang C. L. Kane and E. J. Mele *Phys. Rev. Lett.* **111** 056402 (2013).
- [22] A. Keselman L. Fu A. Stern and E. Berg *Phys. Rev. Lett.* **111** 116402 (2013).
- [23] E. Dumitrescu and S. Tewari *Phys. Rev. B* **88** 220505(R) (2013).
- [24] S. B. Chung J. Horowitz and X.-L. Qi *Phys. Rev. B* **88** 214514 (2013).
- [25] S. Nakosai J. C. Budich Y. Tanaka B. Trauzettel and N. Nagaosa *Phys. Rev. Lett.* **110** 117002 (2013).
- [26] A. Haim A. Keselman E. Berg and Y. Oreg *Phys. Rev. B* **89** 220504(R) (2014).
- [27] E. Gaidamauskas J. Paaske and K. Flensberg *Phys. Rev. Lett.* **112** 126402 (2014).
- [28] J. Klinovaja A. Yacoby and D. Loss *Phys. Rev. B* **90** 155447 (2014).
- [29] C. Schrade A. A. Zyuzin J. Klinovaja and D. Loss *Phys. Rev. Lett.* **115** 237001 (2015).
- [30] A. Haim E. Berg K. Flensberg and Y. Oreg *Phys. Rev. B* **94** 161110(R) (2016).
- [31] Y. Kim D. E. Liu E. Gaidamauskas J. Paaske K. Flensberg and R. M. Lutchyn *Phys. Rev. B* **94** 075439 (2016).
- [32] E. Mellars and B. Béri *Phys. Rev. B* **94** 174508 (2016).
- [33] A. Camjayi L. Arrachea A. Aligia and F. von Oppen *Phys. Rev. Lett.* **119** 046801 (2017).
- [34] C. Schrade and L. Fu *Phys. Rev. Lett.* **120** 267002 (2018).
- [35] A. Haim and Y. Oreg *Phys. Rep.* **825** 1 (2019).
- [36] L. Tosi C. Metzger M. F. Goffman C. Urbina H. Pothier Sunghun Park A. L. Yeyati J. Nygård and P. Krogstrup *Phys. Rev. X* **9** 011010 (2019).
- [37] K. Flensberg *Phys. Rev. B* **82** 180516(R) (2010).
- [38] M. Leijnse and K. Flensberg *Phys. Rev. B* **84** 140501(R) (2011).
- [39] S. Das Sarma J. D. Sau and T. D. Stanescu *Phys. Rev. B* **86** 220506(R) (2012).
- [40] E. Prada P. San-Jose and R. Aguado *Phys. Rev. B* **86** 180503(R) (2012).
- [41] D. Rainis L. Trifunovic J. Klinovaja and D. Loss *Phys. Rev. B* **87** 024515 (2013).
- [42] T. D. Stanescu and S. Tewari *J. Phys.: Condens. Matter* **25** 233201 (2013).

-
- [43] J. Cayao E. Prada P. San-Jose and R. Aguado Phys. Rev. B **91** 024514 (2015).
- [44] T.D. Stanescu and S. Das Sarma Phys. Rev. B **97** 045410 (2018).
- [45] Y. Huang J. D. Sau T. D. Stanescu and S. Das Sarma Phys. Rev. B **98** 224512 (2018).
- [46] C. Reeg O. Dmytruk D. Chevallier D. Loss and J. Klinovaja Phys. Rev. B **98** 245407 (2018).
- [47] A. E. G. Mikkelsen P. Kotetes P. Krogstrup and K. Flensberg Phys. Rev. X **8** 031040 (2018).
- [48] A. E. Antipov A. Bargerbos G. W. Winkler B. Bauer E. Rossi and R. M. Lutchyn Phys. Rev. X **8** 031041 (2018).
- [49] S. D. Escribano A. Levy Yeyati Y. Oreg and E. Prada Phys. Rev. B **100** 045301 (2019).
- [50] G. W. Winkler A. E. Antipov B. van Heck A. A. Soluyanov L. I. Glazman M. Wimmer and R. M. Lutchyn Phys. Rev. B **99** 245408 (2019).
- [51] A. Haim and A. Stern Phys. Rev. Lett. **122** 126801 (2019).
- [52] A. R. Akhmerov J. P. Dahlhaus F. Hassler M. Wimmer and C. W. J. Beenakker Phys. Rev. Lett. **106** 057001 (2011).
- [53] F. Pientka G. Kells A. Romito P. W. Brouwer and F. von Oppen Phys. Rev. Lett. **109** 227006 (2012).
- [54] M. Houzet J. S. Meyer D. M. Badiane and L. I. Glazman Phys. Rev. Lett. **111** 046401 (2013).
- [55] P. A. Ioselevich P. M. Ostrovsky and M. V. Feigel'man Phys. Rev. B **93** 125435 (2016).
- [56] F. Setiawan W. S. Cole J. D. Sau and S. Das Sarma Phys. Rev. B **95** 174515 (2017).
- [57] C. Reeg and D. L. Maslov Phys. Rev. B **95** 205439 (2017).
- [58] E. Nakhmedov B. D. Suleymanli O. Z. Alekperov F. Tatardar H. Mammadov A. A. Konovko A. M. Saletsky Yu. M. Shukrinov K. Sengupta and B. Tanatar arXiv:1911.09289.
- [59] A. Zazunov R. Egger and A. Levy Yeyati Phys. Rev. B **94** 014502 (2016).
- [60] Y. Peng Y. Bao and F. von Oppen Phys. Rev. B **95** 235143 (2017).
- [61] A. Zazunov R. Egger M. Alvarado and A. L. Yeyati Phys. Rev. B **96** 024516 (2017).
- [62] A. Komnik and S. Heinze Phys. Rev. B **96** 155103 (2017).
- [63] A. Zazunov A. Iks M. Alvarado A. Levy Yeyati and R. Egger Beilstein J. Nanotechn. **9** 1659 (2018).
- [64] M. Istas C. Groth and X. Waintal Phys. Rev. Res. **1** 033188 (2019).
- [65] A. M. Essin and V. Gurarie Phys. Rev. B **84** 125132 (2011).
- [66] X.-L. Qi T. L. Hughes and S.-C. Zhang Phys. Rev. B **81** 134508 (2010).
- [67] S. Tewari and J. D. Sau Phys. Rev. Lett. **109** 150408 (2012).

- [68] R.M. Lutchyn J.D. Sau and S. Das Sarma Phys. Rev. Lett. **105** 077001 (2010).
- [69] Y. Oreg G. Refael and F. von Oppen Phys. Rev. Lett. **105** 177002 (2010).
- [70] S. Park and A. L. Yeyati Phys. Rev. B **96** 125416 (2017).
- [71] L. Arrachea G. S. Lozano and A. A. Aligia Phys. Rev. B **80** 014425 (2009).
- [72] I. C. Fulga F. Hassler and A. R. Akhmerov Phys. Rev. B **85** 165409 (2012).
- [73] See Supplemental Material at <http://link.aps.org/supplemental/DOI> for two movies that illustrate the motion of the roots in the complex plane. The first movie refers to a single-channel wire with varying ratio V_x/V_c for the parameters in Figs. 96(a) and 96(b). The second movie is for the two-channel case with varying chemical potential μ using the parameters in Fig. 98.
- [74] R. M. Lutchyn T. D. Stanescu and S. Das Sarma Phys. Rev. Lett. **106** 127001 (2011).
- [75] B. D. Woods J. Chen S. M. Frolov and T. D. Stanescu Phys. Rev. B **100** 125407 (2019).
- [76] R. S. K. Mong and V. Shivamoggi Phys. Rev. B **83** 125109 (2011).
- [77] A. A. Aligia and L. Arrachea Phys. Rev. B **98** 174507 (2018).
- [78] J. Alicea Y. Oreg G. Refael F. von Oppen and M. P. A. Fisher Nat. Phys. **7** 412 (2011).
- [79] O. Deb K. Sengupta and D. Sen Phys. Rev. B **97** 174518 (2018).
- [80] M. Trif and P. Simon Phys. Rev. Lett. **122** 236803 (2019).
- [81] L. Peralta Gavensky G. Usaj and C. A. Balseiro Phys. Rev. B **100** 014514 (2019).
- [82] J. S. Meyer and M. Houzet arXiv:1911.07705.
- [83] T. Jonckheere J. Rech A. Zazunov R. Egger A. L. Yeyati and T. Martin Phys. Rev. Lett. **122** 097003 (2019).
- [84] P. Wei S. Manna M. Eich P. Lee and J. Moodera Phys. Rev. Lett. **122** 247002 (2019).
- [85] In the limit $\mu \rightarrow 0$ there is an appreciable canting angle in the TS lead [61] which transforms $\hat{\Psi}_R$ and requires one to include higher-order processes. For $V_x \rightarrow V_c$ we then find that φ_0 -junction behavior is possible with a finite supercurrent at $\phi = 0$. Moreover higher-order processes may also cause an effective hybridization of the Majorana-Kramers pair in the TRITOPS wire. This fact explains the sinusoidal dispersion of the Andreev bound states for $\theta = \pi/2$ in Fig. 80(c).

References from Conductance calculations for quantum wires and interfaces: mode matching and Green functions

References

- [1] M. N. Baibich J. M. Broto A. Fert F. N. van Dau F. Petroff P. Etienne G. Creuzet A. Friedrich and J. Chazelas Phys. Rev. Lett. **61** 2472 (1988).
- [2] G. Binasch P. Grünberg F. Saurenbach and W. Zinn Phys. Rev. B **39** 4828 (1989).

- [3] N. Agraït A. L. Yeyati and J. M. van Ruitenbeek Phys. Rep. **377** 81 (2003).
- [4] M. Büttiker Y. Imry R. Landauer and S. Pinhas Phys. Rev. B **31** 6207 (1985).
- [5] K. M. Schep J. B. A. N. van Hoof P. J. Kelly G. E. W. Bauer and J. E. Inglesfield Phys. Rev. B **56** 10805 (1997).
- [6] J. B. A. N. van Hoof K. M. Schep A. Brataas G. E. W. Bauer and P. J. Kelly Phys. Rev. B **59** 138 (1999).
- [7] J. Kudrnovský V. Drchal C. Blaas P. Weinberger I. Turek and P. Bruno Phys. Rev. B **62** 15084 (2000).
- [8] K. Xia P. J. Kelly G. Bauer I. Turek J. Kudrnovský and V. Drchal Phys. Rev. B **63** 064407 (2001).
- [9] I. Riedel P. Zahn and I. Mertig Phys. Rev. B **63** 195403 (2001).
- [10] J. Taylor H. Guo and J. Wang Phys. Rev. B **63** 245407 (2001).
- [11] M. Brandbyge J. L. Mozos P. Ordejón J. Taylor and K. Stokbro Phys. Rev. B **65** 165401 (2002).
- [12] D. Wortmann H. Ishida and S. Blügel Phys. Rev. B **66** 075113 (2002).
- [13] K. S. Thygesen M. V. Bollinger and K. W. Jacobsen Phys. Rev. B **67** 115404 (2003).
- [14] P. Mavropoulos N. Papanikolaou and P. Dederichs Phys. Rev. B **69** 125104 (2004).
- [15] C. Caroli R. Combescot P. Nozières and D. Saint-James J. Phys. C: Sol. State Phys. **4** 916 (1971).
- [16] S. Datta *Electronic Transport in Mesoscopic Systems* (Cambridge University Press Cambridge 1995).
- [17] T. Ando Phys. Rev. B **44** 8017 (1991).
- [18] K. Nicolíć and A. MacKinnon Phys. Rev. B **50** 11008 (1994).
- [19] K. Xia P. J. Kelly G. E. W. Bauer A. Brataas and I. Turek Phys. Rev. B **65** 220401 (2002).
- [20] K. Xia P. J. Kelly G. E. W. Bauer and I. Turek Phys. Rev. Lett. **89** 166603 (2002).
- [21] M. Zwierzycki K. Xia P. J. Kelly and G. E. W. Bauer Phys. Rev. B **67** 092401 (2003).
- [22] P. S. Krstić X.-G. Zhang and W. H. Butler Phys. Rev. B **66** 205319 (2002).
- [23] P. Sautet and C. Joachim Phys. Rev. B **38** 12238 (1988).
- [24] We restrict ourselves to a representation on orthogonal basis sets but the extension to non-orthogonal basis sets is straightforward.
- [25] K. Xia M. Zwierzycki M. Talanana P. J. Kelly and G. E. W. Bauer to be published.
- [26] P. A. Khomyakov and G. Brocks Phys. Rev. B **70** 195402 (2004).
- [27] A. MacKinnon Z. Phys. B **59** 385 (1985).

- [28] In some representations the hopping matrices can be singular. In all cases where the inverse of such a matrix would be needed one can use the well-defined pseudo-inverse matrix instead cf. Ref. [30].
- [29] F. Tisseur and K. Meerbergen *SIAM Review* **43** 235 (2001).
- [30] G. Golub and C. F. van Loan *Matrix Computations* (Johns Hopkins University Press Baltimore 1996).
- [31] L. Molinari *J. Phys. A: Math. Gen.* **30** 983 (1997).
- [32] The current per mode is given by velocity $\times$ density. Since we normalize a mode in a slice its density is given by 1/thickness of the slice.
- [33] E. Godfrin *J. Phys.: Condens. Matter.* **3** 7843 (1991).
- [34] A. R. Williams P. J. Feibelman and N. D. Lang *Phys. Rev. B* **26** 5433 (1982).
- [35] I. Turek V. Drchal J. Kudrnovský M. Šob and P. Weinberger *Electronic Structure of Disordered Alloys Surfaces and Interfaces* (Kluwer Boston-London-Dordrecht 1997).
- [36] F. Guinea C. Tejedor F. Flores and E. Louis *Phys. Rev. B* **28** 4397 (1983).
- [37] D. S. Fisher and P. A. Lee *Phys. Rev. B* **23** 6851 (1981).
- [38] A. Messiah *Quantum Mechanics* (North-Holland Amsterdam 1961).
- [39] P. A. Khomyakov (unpublished).
- [40] T. Kostyrko *Phys. Rev. B* **62** 2458 (2000).
- [41] L. D. Landau and E. M. Lifschitz *Quantum Mechanics: Non-relativistic Theory* (Pergamon Press Oxford 1981) §128.
- [42] J. P. Perdew and A. Zunger *Phys. Rev. B* **23** 5048 (1981).
- [43] O. K. Andersen O. Jepsen and D. Glötzl in *Highlights of Condensed Matter Theory* edited by F. Bassani F. Fumi and M. P. Tosi (North-Holland Amsterdam 1985) pp. 59–176.
- [44] C. S. Wang and A. J. Freeman *Phys. Rev. B* **24** 4364 (1981).
- [45] J. A. Strosio D. T. Pierce A. Davies R. J. Celotta and M. Weinert *Phys. Rev. Lett.* **75** 2960 (1995).
- [46] C. Uiberacker and P. M. Levy *Phys. Rev. B* **64** 193404 (2001).
- [47] C. Uiberacker and P. M. Levy *Phys. Rev. B* **65** 169904(E) (2002).
- [48] O. Wunnicke N. Papanikolaou R. Zeller P. H. Dederichs V. Drchal and J. Kudrnovský *Phys. Rev. B* **65** 064425 (2002).
- [49] W. H. Butler X.-G. Zhang T. C. Schulthess and J. M. MacLaren *Phys. Rev. B* **63** 054416 (2001).
- [50] J. Mathon and A. Umerski *Phys. Rev. B* **63** 220403(R) (2001).

References from Caroli's articles

Appelbaum J A and Brinkman W F 1969 Phys. Rev. 186464 – 70

Caroli C Combescot R Nozieres P and Saint-James D 1971 J. Phys. C: Solid St. Phys. 491629

Duke C B 1969 Tunnelling in Solids (New York: Academic Press)

Leger A 1971 PhD thesis Université de Paris

Schwinger J 1961 J. math. Phys. 2 407-32

(some others are also there)

References Boundary Green functions of topological insulators and superconductors

References

- [1] M. Z. Hasan and C. L. Kane *Rev. Mod. Phys.* **82** 3045 (2010).
- [2] X.-L. Qi and S.-C. Zhang *Rev. Mod. Phys.* **83** 1057 (2011).
- [3] B. A. Bernevig and T. L. Hughes *Topological insulators and topological superconductors* (Princeton University Press 2013).
- [4] A. P. Schnyder S. Ryu A. Furusaki and A. W. W. Ludwig *Phys. Rev. B* **78** 195125 (2008).
- [5] A. Kitaev AIP Conf. Proc. **1134** 22 (2009).
- [6] S. Ryu A. P. Schnyder A. Furusaki and A. W. Ludwig New Journal of Physics **12** 065010 (2010).
- [7] J. C. Y. Teo and C. L. Kane *Phys. Rev. B* **82** 115120 (2010).
- [8] A. M. Essin and V. Gurarie *Phys. Rev. B* **84** 125132 (2011).
- [9] I. C. Fulga F. Hassler and A. R. Akhmerov *Phys. Rev. B* **85** 165409 (2012).
- [10] C.-K. Chiu J. C. Y. Teo A. P. Schnyder and S. Ryu *Rev. Mod. Phys.* **88** 035005 (2016).
- [42] M. L. Sancho J. L. Sancho J. L. Sancho and J. Rubio J. Phys. F: Met. Phys. **15** 851 (1985).
- [12] A. Umerski *Phys. Rev. B* **55** 5266 (1997).
- [13] W. P. Su J. R. Schrieffer and A. J. Heeger *Phys. Rev. Lett.* **42** 1698 (1979).
- [14] J. Pichard and G. Sarma Journal of Physics C: Solid State Physics **14** L127 (1981).
- [15] D. H. Lee and J. D. Joannopoulos *Phys. Rev. B* **23** 4988 (1981).
- [16] Y. Hatsugai *Phys. Rev. Lett.* **71** 3697 (1993).
- [17] Y. Hatsugai *Phys. Rev. B* **48** 11851 (1993).
- [18] A. Kitaev *Phys. Usp.* **44** 131 (2001).

- [19] W. A. Benalcazar B. A. Bernevig and T. L. Hughes arXiv:1611.07987 (2016).
- [20] X. Dai T. L. Hughes X.-L. Qi Z. Fang and S.-C. Zhang *Phys. Rev. B* **77** 125319 (2008).
- [21] K. W. Kim R. S. K. Mong M. Franz and G. Refael *Phys. Rev. B* **92** 075110 (2015).
- [22] C. Mahaux and H. Weidenmüller *Shell-model approach to nuclear reactions* (North-Holland Pub. Co. 1969).
- [23] I. L. Aleiner P. W. Brouwer and L. I. Glazman *Phys. Rep.* **358** 309 (2002).
- [24] N. M. Boffi J. C. Hill and M. G. Reuter *Comput. Sci. Discov.* **8** 015001 (2014).
- [25] C. W. J. Beenakker *Rev. Mod. Phys.* **69** 731 (1997).
- [26] A. Zazunov R. Egger and A. Levy Yeyati *Phys. Rev. B* **94** 014502 (2016).
- [27] R. B. Laughlin *Phys. Rev. B* **23** 5632 (1981).
- [28] P. W. Brouwer *Phys. Rev. B* **58** R10135 (1998).
- [29] Y. V. Nazarov and Y. M. Blanter *Quantum transport: introduction to nanoscience* (Cambridge University Press 2009).
- [30] A. Altland and M. R. Zirnbauer *Phys. Rev. B* **55** 1142 (1997).
- [31] C. W. J. Beenakker *Rev. Mod. Phys.* **87** 1037 (2015).
- [32] J. Zak *Phys. Rev. Lett.* **62** 2747 (1989).
- [33] S. Ryu and Y. Hatsugai *Phys. Rev. Lett.* **89** 077002 (2002).
- [34] X.-L. Qi T. L. Hughes and S.-C. Zhang *Phys. Rev. B* **81** 134508 (2010).
- [35] L. Fu and C. L. Kane *Phys. Rev. B* **74** 195312 (2006).

From Thermal transport in one-dimensional spin heterostructures

References

- [1] For a recent review see for instance A. V. Sologubenko T. Lorenz H. R. Ott A. Freimuth J. Low Temp. Phys. **147** 387 (2007). See also X. Zotos P. Prelovsek in *"Interacting Electrons in Low Dimensions" book series "Physics and Chemistry of Materials with Low-Dimensional Structures"* Kluwer Academic Publishers (2003)
- [2] Abhishek Dhar *Advances In Physics* **57** 457 (2008).
- [3] A. V. Sologubenko E. Felder K. Giannò H. R. Ott A. Vietkine and A. Revcolevschi *Phys. Rev. B* **62** R6108 (2000); A. V. Sologubenko K. Giannò H. R. Ott A. Vietkine and A. Revcolevschi *ibid.* **64** 054412 (2001); A. V. Sologubenko H. R. Ott G. Dhalenne and A. Revcolevschi *Europhys. Lett.* **62** 540 (2003).

- [4] For some recent works on spin chain and ladders see for instance F.Heidrich-Meisner A. Honecker and W. Brenig Phys. Rev. **B 71** 184415 (2005); A. V. Rozhkov and A. L. Chernyshev Phys. Rev. Lett. **94** 087201 (2005); P. Jung R. W. Helmes and A. Rosch Phys. Rev. Lett. **96** 067202 (2006); K. Louis P. Prelovsek and X. Zotos Phys. Rev. **B 74** 235118 (2006); P. Jung and A. Rosch Phys. Rev. **B 75** 245104 (2007); E. Boulat P. Mehta N. Andrei E. Shimshoni and A. Rosch Phys. Rev. **B 76** 214411 (2007); A. V. Savin G. P. Tsironis and X. Zotos Phys. Rev. **B 75** 214305 (2007)
- [5] See also H. Castella X. Zotos and P. Prelovsek Phys. Rev. Lett. **74** 972 (1995); S. Fujimoto and N. Kawakami J. Phys. **A 31** 465 (1998); X. Zotos Phys. Rev. Lett. **82** 1764 (1999); T. Prosen and D. K. Campbell Phys. Rev. Lett. **84** 2857 (2000); J. V. Alvarez and C. Gros Phys. Rev. Lett. **88** 077203 (2002); E. Shimshoni N. Andrei and A. Rosch Phys. Rev. **B 68** 104401 (2003); K. Saito Phys. Rev. B **67** 064410 (2003); F. Heidrich-Meisner A. Honecker D.C. Cabra and W. Brenig Phys. Rev. B **66** 140406(R) (2002) Phys. Rev. **B 68** 134436 (2003) Phys. Rev. Lett. **92** 069703 (2004) E. Orignac R. Chitra R. Citro Phys. Rev. B **67** 134426 (2003); X. Zotos Phys. Rev. Lett. **92** 067202 (2004).
- [6] H. M. Pastawski P. R. Levstein and G. Usaj Phys. Rev. Lett. **75** 4310 (1995); D. Loss and D. P. DiVincenzo Phys. Rev. A **57** 120 (1998); T. S. Cubitt and J. I. Cirac Phys. Rev. Lett. **100** 180406 (2008) and references therein.
- [7] G.D. Mahan *Many Particle Physics* (Kluver/Plenum New York 2000).
- [8] W. Aschbacher and C.A. Pillet J. Stat. Phys. **112** 1153 (2003) Y. Ogata Phys. Rev. **66** 066123 (2002) Y. Ogata Phys. Rev. **66** 016135 (2002).
- [9] M. Michel G. Mahler and J. Gemmer Phys. Rev. Lett. **95** 180602 (2005); J. Gemmer R. Steinigeweg and M. Michel Phys. Rev. B **73** 104302 (2006).
- [10] M. Michel O. Hess H. Wichterich and J. Gemmer Phys. Rev. B **77** 104303 (2008).
- [11] For other type of baths see D. Karevski and T. Platini unpublished.
- [12] S. Datta *Electronic Transport in Mesoscopic Systems* (Cambridge University press 1995); Y. Imry *Introduction to Mesoscopic Physics* (Oxford University Press 1997).
- [13] J. Rammer and H. Smith Rev. Mod. Phys. **58** 323 (1986).
- [14] L. Arrachea M. Moskalets and L. Martin-Moreno Phys. Rev. **B 75** 245420 (2007); L. Arrachea and M. Moskalets cond-mat/0903.1153.
- [15] L. Arrachea Phys. Rev. B in press cond-mat/0811.1648.
- [16] D.C. Langreth in *Linear and Non-Linear Transport in Solids* J.T. Devreese E. Van Dofen (Eds.) (Plenum Press New York 1076).
- [17] A. Hernández V. M. Apel F. A. Pinheiro and C. H. Lewenkopf Physica A **385** 148 (2007).
- [18] G. E. Blonder M. Tinkham and T. M. Klapwijk Phys. Rev. B **25** 4515 (1982).
- [19] N.W. Ashcroft and N. D. Mermin "Solid State Physics" Holt Rinehart and Winston. 1976.
- [20] J. C. Cuevas A. Martín-Rodero and A. Levy Yeyati Phys. Rev. B **54** 7366 (1996).
- [21] I.S. Gradshteyn and I.M. Ryzhik *Tables of Integrals Series and Products* (Academic Press New York 1980)

- [22] A. E. Feiguin S. R. White and D. J. Scalapino *Phys. Rev. B.* **75** 024505 (2007); A. E. Feiguin S. R. White D. J. Scalapino and I. Affleck; *Phys. Rev. Lett.* **101** 217001 (2008);

From Equivalence of wave function matching and Green's functions methods for quantum transport: generalized Fisher-Lee relation Hocine by Boumrar Hamidi Zenia Lounis

References

- [1] Ando T 1991 *Phys. Rev. B* **44** 8017.
- [2] Khomyakov P A Brocks G Karpan V Zwierzycki M and Kelly P J 2005 *Phys. Rev. B* **72** 035450.
- [3] Datta S 1997 *Electronic Transport in Mesoscopic Systems* (Cambridge University Press).
- [4] Di Ventra M 2008 *Electrical Transport in Nanoscale Systems* (Cambridge University Press).
- [5] Evers F Korytár R Tewari S and van Ruitenbeek J M Advances and challenges in single-molecule electron transport arXiv:1906.10449 [cond-mat.mes-hall].
- [6] Bell L E 2008 *Science* **321** 1457.
- [7] Dubi Y and Di Ventra M 2011 *Rev. Mod. Phys.* **83** 131.
- [8] Cahill D G Braun P V Chen G Clarke D R Fan S Goodson K E Keblinski P King W P Mahan G D Majumdar A Maris H J Phillpot S R Pop E and Shi L 2014 *Appl. Phys. Rev.* **1** 011305.
- [9] Freericks J K 2006 *Transport in multilayered nanostructures: the dynamical mean-field theory approach* (Imperial College Press London).
- [10] Landauer R 1957 *IBM J. Res. Dev.* **1** 233.
- [11] Landauer R 1970 *Philos. Mag.* **21** 863.
- [12] Khomyakov P A and Brocks G 2004 *Phys. Rev. B* **70** 195402.
- [13] Xia K Zwierzycki M Talanana M Kelly P J and Bauer G E W 2006 *Phys. Rev. B* **73** 064420.
- [14] Farmanbar M Amlaki T and Brocks G 2016 *Phys. Rev. B* **93** 205444.
- [15] Chen X Xu Y Wang J and Guo H 2019 *Phys. Rev. B* **99** 064302.
- [16] Mingo N and Yang L 2003 *Phys. Rev. B* **68** 245406.
- [17] Wang J S Wang J and Lü J T 2008 *Eur. Phys. J. B* **62** 381.
- [18] Zhang W Fisher T and Mingo N 2007 *ASME J. Heat Transfer* **129** 483.
- [19] Li X and Yang R 2012 *Phys. Rev. B* **86** 054305.
- [20] Gu X Li X and Yang R 2015 *Phys. Rev. B* **91** 205313.
- [21] Tian Z Esfarjani K and Chen G 2012 *Phys. Rev. B* **86** 235304.

- [22] Caroli C Combescot R Nozieres P and Saint-James D 1971 *J. Phys. C* **4** 916.
- [23] Keldysh L V 1965 *Sov. Phys. JETP* **20** 1018.
- [24] Zhu Y and Liu L 2016 *Atomistic Simulation of Quantum Transport in Nanoelectronic Devices* (World Scientific Singapore).
- [25] Krstić P S Zhang X G and Butler W H 2002 *Phys. Rev. B* **66** 205319
- [26] Giannozzi P Baroni S Bonini N *et al.* *J.Phys.:Condens.Matter* 21 395502 (2009); Giannozzi P Andreussi O Brumme T *et al.* *J.Phys.:Condens.Matter* 29 465901 (2017)
- [27] H. Ebert D. KÖdderitzsch and J. Minár *Rep. Prog. Phys.* 74 096501 (2011).
- [28] Huang Z Murthy J Y and Fisher T S 2011 *J. Heat Transfer* **133** 114502.
- [29] Sadasivam S Waghmare U V and Fisher T S 2017 *Phys. Rev. B* **96** 174302.
- [30] Ong Z Y and Zhang G 2015 *Phys. Rev. B* **91** 174302.
- [31] Latour B Shulumba N and Minnich A J 2017 *Phys. Rev. B* **96** 104310.
- [32] Ong Z Y 2018 *J. Appl. Phys.* **124** 151101.
- [33] Klöckner J C Cuevas J C and Pauly F 2018 *Phys. Rev. B* **97** 155432.
- [34] Fisher D S and Lee P A 1981 *Phys. Rev. B* **23** 6851.
- [35] Ferry D K and Goodnick S M 2001 *Transport in Nanostructures* (Cambridge University Press).
- [36] Sanvito S Lambert C J Jefferson J H and Bratkovsky A M 1999 *Phys. Rev. B* **59** 11936.
- [37] Wimmer M 2009 Quantum transport in nanostructures: From computational concepts to spintronics in graphene and magnetic tunnel junctions *PhD Thesis* Regensburg University.
- [38] Zhang S H Yang W and Chang K 2017 *Phys. Rev. B* **95** 075421.
- [39] Cohen-Tannoudji C Diu B and Lalöe F 1977 *Quantum Mechanics* (John Wiley & Sons New York).
- [40] Mingo N 2009 *Green's function methods for phonon transport through nano-contacts* (Springer Heidelberg).
- [41] Sadasivam S Che Y Huang Z Chen L Kumar S and Fisher T S 2014 *Annu. Rev. Heat Transfer* **17** 89.
- [42] Lopez Sancho M P Lopez Sancho J M and Rubio J 1985 *J. Phys. F: Met. Phys.* **15** 851.
- [43] Velev J and Butler W 2004 *J. Phys.: Condens. Matter* **16** R637
- [44] Cuevas J C and Scheer E 2017 *Molecular Electronics: An Introduction to Theory and Experiment* (World Scientific Singapore 2nd edition edition).
- [45] Cuevas J C Yeyati A L and Martín-Rodero A 1998 *Phys. Rev. Lett.* **80** 1066.
- [46] Häfner M Viljas J K Frustaglia D Pauly F Dreher M Nielaba P and Cuevas J C 2008 *Phys. Rev. B* **77** 104409.

From Hamiltonian approach to the transport properties of superconducting quantum point contacts by Cuevas Martín-Rodero Levy Yeyati

- ¹ A.F. Andreev Zh. Éksp. Teor. Fiz. 46 1823 (1964) [Sov. Phys. JETP 19 1228 (1964)]; 49 655 (1966) [22 455 (1966)]; B.D. Josephson Phys. Lett. 1 251 (1962); Rev. Mod. Phys. 36 216 (1964).
- ² K.K. Likharev Rev. Mod. Phys. 51 101 (1979).
- ³ G.E. Blonder M. Tinkham and T.M. Klapwijk Phys. Rev. B 25 4515 (1982).
- ⁴ N. van der Post E.T. Peters I.K. Yanson and J.M. van Ruitenbeek Phys. Rev. Lett. 73 2611 (1994); B.J. Vleeming C.J. Muller M.C. Koops and R. de Bruyn Ouboter Phys. Rev. B 50 16741 (1994).
- ⁵ H. Takayanagi T. Akazaki and J. Nitta Phys. Rev. Lett. 75 3533 (1995).
- ⁶ E.N. Bratus V.S. Shumeiko and G. Wendin Phys. Rev. Lett. 74 2110 (1995).
- ⁷ A. Martín-Rodero A. Levy Yeyati and J.C. Cuevas Physica B 218 126 (1996); A. Levy Yeyati A. Martín-Rodero and J.C. Cuevas J. Phys.: Condens. Matter 8 449 (1996).
- ⁸ D. Averin and A. Bardas Phys. Rev. Lett. 75 1831 (1995).
- ⁹ R. Landauer Philos. Mag. 21 863 (1970).
- ¹⁰ J. Bardeen Phys. Rev. Lett. 6 57 (1961).
- ¹¹ V. Ambegaokar and A. Baratoff Phys. Rev. Lett. 10 486 (1963); 11 104 (1963).
- ¹² J.R. Schrieffer and J.W. Wilkins Phys. Rev. Lett. 10 17 (1963); J.W. Wilkins Tunneling Phenomena in Solids (Plenum New York 1963) p. 333.
- ¹³ A. Martín-Rodero F. J. García-Vidal and A. Levy Yeyati Phys. Rev. Lett. 72 554 (1994); A. Levy Yeyati A. Martín-Rodero and F. J. García-Vidal Phys. Rev. B 51 3743 (1995).
- ¹⁴ A. Levy Yeyati A. Martín-Rodero and F. Flores Phys. Rev. Lett. 71 2991 (1993).
- ¹⁵ P.G. de Gennes Superconductivity of Metals and Alloys (Benjamin New York 1966).
- ¹⁶ D. Rogovin and D.J. Scalapino Ann. Phys. 86 1 (1974).
- ¹⁷ L.V. Keldysh Sov. Phys. JETP 20 1018 (1965).
- ¹⁸ A typical estimate of η for a traditional superconductor is $\eta/\Delta \sim 10^{-2}$. See for instance S.B. Kaplan et al. Phys. Rev. B 14 4854 (1976).
- ¹⁹ C. Caroli R. Combescot P. Nozières and D. Saint-James J. Phys. C 4 916 (1971).
- ²⁰ J. Ferrer A. Martín-Rodero and F. Flores Phys. Rev. B 38 10113 (1988).
- ²¹ A different derivation of Eq. (16) can be found in Y.A. Genenko and Y.A. Ivanchenko Theor. Math. Phys. 69 1056 (1986).
- ²² G.D. Mahan Many-Particle Physics (Plenum New York 1981).
- ²³ A finite order perturbative calculation in t yields a conductance which diverges at $eV = \Delta$ due to the divergencies of the uncoupled BCS Green functions [Eq. (7)].
- ²⁴ L.E. Hasselberg J. Phys. F 3 1438 (1973); L.E. Hasselberg M.T. Levinsen and M.R. Samuelsen Phys. Rev. B 9 3757 (1974).
- ²⁵ A.V. Zaitsev Zh. Éksp. Teor. Fiz. 78 221 (1980) [Sov. Phys. JETP 51 111 (1980)].
- ²⁶ G. B. Arnold J. Low Temp. Phys. 59 143 (1985); 68 1 (1987).
- ²⁷ A. Levy Yeyati and F. Flores J. Phys.: Condens. Matter 4 7341 (1992); A. Levy Yeyati J. C. Cuevas and A. Martín-Rodero in Photons and Local Probes edited by O. Marti and R. Moller (Kluwer Academic Dordrecht 1995).
- ²⁸ M. Octavio M. Tinkham G.E. Blonder and T.M. Klapwijk Phys. Rev. B 27 6739 (1983).
- ²⁹ T.M. Klapwijk G.E. Blonder and M. Tinkham Physica B&C 109-110 1657 (1982).
- ³⁰ K. Flensberg J. Bindslev Hansen and M. Octavio Phys. Rev. B 38 8707 (1988).
- ³¹ A. Furusaki and M. Tsukada Physica B 165+166 967 (1990).
- ³² The dominant contribution of MAR in the linear conductance of S – N – S weak links has been recently pointed out by U. Günsenheimer and A.D. Zaikin Phys. Rev. B 50 6317 (1994).
- ³³ R. Rifkin and B.S. Deaver Jr. Phys. Rev. B 13 3894 (1976).

- ³⁴ A.B. Zorin et al. Fiz. Nizk. Temp. 5 1138 (1979) [Sov. J. Low Temp. Phys. 5 537 (1979)].
- ³⁵ A. Barone and G. Paterno Physics and Applications of the Josephson Effect (Wiley New York 1982).
- ³⁶ A. Martín-Rodero A. Levy Yeyati and F.J. García-Vidal Phys. Rev. B 53 R8891 (1996).
- ³⁷ R. de Bruyn Ouboter and A.N. Omelyanchouk Physica B 216 37 (1995).
- ³⁸ C.W. Beenakker in Proceedings of the 14th Taniguchi International Symposium on Transport Phenomena in Mesoscopic Systems edited by H. Fukuyama and T. Ando (Springer Berlin 1992).

From Quasiclassical Green's function approach to mesoscopic superconductivity by Belzig, Wilhelm, Bruder, Schön, Zaikin

References

- [1] G. Eilenberger, Z. Phys. **214**, 195 (1968).
- [2] A.I. Larkin and Yu.N. Ovchinnikov, Zh. Eksp. Teor. Fiz. **55** 2262 (1968) [Sov. Phys. JETP **26**, 1200 (1968)].
- [3] A.F. Andreev, Zh. Eksp. Teor. Fiz. **46**, 1823 (1964) [JETP **19**, 1228 (1964)].
- [4] A. Schmid in *Nonequilibrium Superconductivity*, edited by K.E. Gray (Plenum, N.Y., 1981).
- [5] A.I. Larkin and Yu.N. Ovchinnikov, in *Nonequilibrium Superconductivity*, edited by D.N. Langenberg and A.I. Larkin, (Elsevier, Amsterdam, 1984).
- [6] J. Rammer and H. Smith, Rev. Mod. Phys. **58**, 323 (1986).
- [7] A.A. Golubov and M.Yu. Kupriyanov, J. Low Temp. Phys. **70**, 83 (1988).
- [8] W. Belzig, C. Bruder, and G. Schön, Phys. Rev. B **54**, 9443 (1996).
- [9] S. Guéron *et al.*, Phys. Rev. Lett. **77**, 3025 (1996).
- [10] A.C. Mota, P. Visani, and A. Pollini, J. Low Temp. Phys **76**, 465 (1989).
- [11] F.B. Müller-Allinger, A.C. Mota, and W. Belzig, submitted to Phys. Rev. B.
- [12] W. Belzig, C. Bruder, and A.L. Fauchère, Phys. Rev. B **58**, 14531 (1998).
- [13] H. Courtois, Ph. Gandit, and B. Pannetier, Phys. Rev. B **52**, 1162 (1995).
- [14] H. Pothier *et al.*, Phys. Rev. Lett. **79**, 3490 (1997).
- [15] F.K. Wilhelm, G. Schön, and A.D. Zaikin, Phys. Rev. Lett. **81**, 1682 (1998).
- [16] A. Morpurgo, B.J. van Wees, and T.M. Klapwijk, Appl. Phys. Lett. **72**, 966 (1998).
- [17] J.J.A. Baselmans, B.J. van Wees, and T.M. Klapwijk, to appear in Nature.
- [18] B.J. van Wees *et al.*, this volume.
- [19] A.A. Golubov, F.K. Wilhelm, and A.D. Zaikin, Phys. Rev. B **55**, 1123 (1997).
- [20] Yu.V. Nazarov and T.H. Stoof, Phys. Rev. Lett. **76**, 823 (1996).

- [21] P. Charlat *et al.*, Czech. J. Phys. **46**, S6 3107 (1996).
- [22] H. Courtois *et al.*, this volume.
- [23] H. Courtois *et al.*, Phys. Rev. Lett. **76**, 130 (1996).
- [24] V.T. Petrashov *et al.*, Pis'ma Zh. Eksp. Teor. Fiz. **58**, 48 (1993) [JETP Lett. **58** 49 (1993)].
- [25] V.T. Petrashov *et al.*, Phys. Rev. Lett. **74**, 5268 (1995).
- [26] F.K. Wilhelm, A.D. Zaikin, and H. Courtois, Phys. Rev. Lett. **80**, 4289 (1998).
- [27] Articles in *Mesoscopic Superconductivity*, edited by F.W.J. Hekking, G. Schön, and D.V. Averin, Physica B **203**, 201-538 (1994).
- [28] Articles in *Mesoscopic Electron Transport*, edited by L.L. Sohn, L.P. Kouwenhoven, and G. Schön (Kluwer, Dordrecht, 1997).
- [29] G. Schön, in *Nonequilibrium Superconductivity*, edited by D.N. Langenberg and A.I. Larkin (Elsevier, Amsterdam, 1984).
- [30] C. Bruder, Superconductivity Review **1**, 261 (1996).
- [31] C.J. Lambert and R. Raimondi, J. Phys. Cond. Mat. **10**, 901 (1998).
- [32] *Quasiclassical Methods in Superconductivity and Superfluidity*, Proceedings of the Verditz Workshop 1996, edited by D. Rainer and J. Sauls (1998).
- [33] A.A. Abrikosov, L.P. Gorkov and I.E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Dover, 1963).
- [34] see e.g. G. Rickayzen, *Green's Functions and Condensed Matter* (Academic Press, London, 1980); G. Mahan, *Many-Particle Physics*, (Plenum, N.Y. and London, 1981).
- [35] L.P. Kadanoff and G. Baym, *Quantum Statistical Mechanics*, (Benjamin, N.Y., 1962).
- [36] Y. Nambu, Phys. Rev. **117**, 648 (1960).
- [37] I.O. Kulik, Zh. Eksp. Teor. Phys. **57**, 1745 (1969) [Sov. Phys. JETP **30**, 944 (1970)].
- [38] C. Ishii, Prog. Theor. Phys. **44**, 1525 (1970).
- [39] U. Eckern and A. Schmid, J. Low Temp. Phys. **45**, 137.
- [40] J. Eom, C.J. Chien, and V. Chandrasekhar, Phys. Rev. Lett. **81**, 437 (1998).
- [41] T. Matsubara, Prog. Theor. Phys. **14**, 351 (1955).
- [42] L.V. Keldysh, Zh. Eksp. Teor. Phys. **47**, 1515 (1964) [Sov. Phys. JETP **20**, 1018 (1965)].
- [43] H. Smith and H.H. Jensen, *Transport Phenomena* (Clarendon, Oxford, 1989).
- [44] U. Gunsenheimer and A.D. Zaikin, Phys. Rev. B **50**, 6317 (1994); Europhys. Lett. **41**, 195 (1998).
- [45] K.D. Usadel, Phys. Rev. Lett. **25**, 507 (1970).
- [46] A. Schmid and G. Schön, J. Low Temp. Phys. **20**, 207 (1975).

- [47] M. Tinkham, *Introduction to Superconductivity*, 2nd ed., (McGraw Hill, N.Y., 1996).
- [48] A.V. Galaktionov, private communication.
- [49] A.V. Zaitsev, Zh. Eksp. Teor. Fiz. **86**, 1742 (1984) [Sov. Phys. JETP **59**, 1015 (1984)].
- [50] M.Yu. Kuprianov and V.F. Lukichev, Zh. Eksp. Teor. Fiz. **94**, 139 (1988) [Sov. Phys. JETP **67**, 1163 (1988)].
- [51] C.J. Lambert *et al.*, Phys. Rev. B **55**, 6015 (1997).
- [52] N. Schopohl and K. Maki, Phys. Rev. B **52**, 490 (1995).
- [53] N. Schopohl, in [32].
- [54] M. Eschrig, PhD thesis, Universität Bayreuth 1997; published in [32].
- [55] A.D. Zaikin and G.F. Zharkov, Fiz. Nizk. Temp. **7**, 375 (1981) [Sov. J. Low Temp. Phys. (1981)].
- [56] A.V. Volkov *et al.*, this volume
- [57] E. Scheer, private communication.
- [58] E. Scheer *et al.*, Phys. Rev. Lett. **78**, 3535 (1997); Nature **394**, 154-157 (1998).
- [59] W.L. McMillan, Phys. Rev. **175**, 537 (1968).
- [60] A.A. Golubov *et al.*, Phys. Rev. B **51**, 1073 (1995).
- [61] A.F. Volkov in [27].
- [62] A.F. Volkov *et al.*, Physica C **242**, 261 (1995).
- [63] K. M. Frahm *et al.*, Phys. Rev. Lett. **76**, 2981 (1996).
- [64] D. Saint-James, J. Phys. **25**, 899 (1964).
- [65] Orsay Group on Superconductivity, in *Quantum Fluids*, edited by D. Brewer (North-Holland, Amsterdam, 1966).
- [66] Y. Oda and H. Nagano, Solid State Commun. **35**, 631 (1980).
- [67] A.D. Zaikin, Solid State Commun. **41**, 533 (1982).
- [68] A.C. Mota, D. Marek, and J.C. Weber, Helv. Phys. Acta **55**, 647 (1982); J.C. Weber, A.C. Mota, and D. Marek, J. Low Temp. Phys **66**, 41 (1987).
- [69] Th. Bergmann *et al.*, J. Low. Temp. Phys. **66**, 209 (1987).
- [70] P. Visani, A.C. Mota, and A. Pollini, Phys. Rev. Lett. **65**, 1514 (1990).
- [71] A.C. Mota *et al.*, Physica B **197**, 95 (1994).
- [72] S. Higashitani and K. Nagai, J. Phys. Soc. Jpn **64**, 549 (1995).
- [73] W. Belzig, C. Bruder, and G. Schön, Phys. Rev. B **53**, 5727 (1996).
- [74] A.L. Fauchère and G. Blatter, Phys. Rev. B **56**, 14102 (1997).

- [75] A. B. Pippard, Proc. Roy. Soc. **216**, 547 (1953).
- [76] F. and H. London, Proc. Phys. Soc., **A149**, 71 (1935).
- [77] P.G. de Gennes, Rev. Mod. Phys. **36**, 225 (1964).
- [78] M.Yu. Kupriyanov, Sov. J. Superconductivity **2**, 5 (1989).
- [79] F.K. Wilhelm, A.D. Zaikin, and G. Schön, J. Low Temp. Phys. **106**, 305 (1997).
- [80] A.D. Zaikin, in *Nonequilibrium Superconductivity*, edited by V.L. Ginzburg (Nova Science, N.Y., 1988).
- [81] P. Dubos *et al.*, unpublished.
- [82] A.A. Golubov and M.Yu. Kupriyanov, JETP **69**, 805 (1989).
- [83] F. Zhou *et al.*, J. Low Temp. Phys. **110**, 841 (1998).
- [84] L.N. Bulaevskii, V.V. Kuzii, and A.A. Sobyenin, Solid St. Comm. **25**, 1053 (1977).
- [85] S.N. Artemenko, A.F. Volkov, and A.V. Zaitsev, Sol. State Comm. **30**, 771 (1979).
- [86] T.H. Stoof and Yu.V. Nazarov, Phys. Rev. B **53**, 14496 (1996).
- [87] K. Maki, Prog. Theor. Phys. **39**, 897 (1968); *ibid* **40**, 193 (1969); R.S. Thompson, Phys. Rev. B **1**, 327 (1970).
- [88] S.G. den Hartog *et al.*, Phys. Rev. B **56**, 13738 (1997).
- [89] P. Charlat *et al.*, Phys. Rev. Lett. **77**, 4950 (1996).
- [90] T.H. Stoof and Yu.V. Nazarov, Phys. Rev. B **54**, R772 (1996).
- [91] A.F. Volkov, N. Allsopp, and C.J. Lambert, J. Phys. Cond. Mat. **8** L45 (1996).
- [92] A.V. Zaitsev, Physica B **203**, 274 (1994).
- [93] Yu.V. Nazarov, Phys. Rev. Lett. **73**, 134 (1994).
- [94] Yu.V. Nazarov, this volume.
- [95] V.N. Antonov *et al.*, submitted to Phys. Rev. B.
- [96] A.F. Volkov, A.V. Zaitsev, and T.M. Klapwijk, Physica C **59**, 21 (1993).
- [97] A.F. Volkov and A.V. Zaitsev, Phys. Rev. B **53**, 9267 (1996).
- [98] R. Vaglio *et al.*, Phys. Rev. B **47**, 15302 (1993).

From An introduction to the quasiclassical theory of superconductivity for diffusive proximity-coupled system by Chandrasekhar

References

- [1] G. Eilenberger, Z. Phys **214**, 195 (1968).
- [2] A. Schmid and G. Schön, J. Low. Temp. Phys. **20**, 207 (1972).
- [3] G.M. Eliashberg, Zh. Eksp. Teor. Fiz. **61**, 1254 (1971) [Sov. Phys. JETP **34**, 668 (1972)].
- [4] A.I. Larkin and Yu. N. Ovchinnikov, Zh. Eksp. Teor. Fiz. **68**, 1915 (1976) [Sov. Phys. JETP **41**, 960 (1976)].
- [5] A.I. Larkin and Yu. N. Ovchinnikov, Zh. Eksp. Teor. Fiz. **73**, 299 (1977) [Sov. Phys. JETP **46**, 155 (1978)].
- [6] C-R Hu, Phys. Rev. B **21**, 2775 (1980).
- [7] J.W. Serene and D. Rainer, Physics Reports **101**, 221 (1983).
- [8] G. Schön, in **Nonequilibrium Superconductivity**, edited by D.N. Langenberg and A.I. Larkin [Elsevier, 1986]
- [9] A.G. Aronov, Yu. M. Gal'perin, V.L. Gurevich and V.I. Kozub, *Nonequilibrium Properties of Superconductors (Transport Equation Approach)*, in **Nonequilibrium Superconductivity**, edited by D.N. Langenberg and A.I. Larkin [Elsevier, 1986]; Advances in Physics **30**, 539 (1981).
- [10] J. Beyer, C.J. Pethick and J. Rammer, *Pair Breaking in Nonequilibrium Superconductors*, in **Nonequilibrium Superconductivity**, edited by D.N. Langenberg and A.I. Larkin [Elsevier, 1986].
- [11] J. Rammer and H. Smith, Rev. Mod. Phys. **58**, 323 (1986).
- [12] H. Smith and H.H. Jensen, *Transport Phenomena* [Clarendon, Oxford, 1989].
- [13] A.F. Volkov, A.V. Zaitsev and T.M Klapwijk, Physica C **210**, 21 (1993).
- [14] C.J. Lambert and R. Raimondi, J. Phys.: Condens. Matter **10**, 901 (1998).
- [15] A.F. Volkov and V.V. Pavlovskii, in *Proceedings of the AIP conference: lectures on superconductivity in networks and mesoscopic systems*, New York, 1998.
- [16] A.A. Golubov, F.K. Wilhelm, and A.D. Zaikin, Phys. Rev. B **55**, 1123 (1997).
- [17] F.K. Wilhelm, A.D. Zaikin and G. Schön, J. Low Temp. Phys. **106**, 305 (1997).
- [18] W. Belzig, F.K. Wilhelm, C. Bruder, G. Schön and A.D. Zaikin, Superlattices and Microstructures **25**, 1251 (1999).
- [19] Yu. V. Nazarov, Superlattices and Microstructures **25**, 1221 (1999).
- [20] K.E. Nagaev, Phys. Lett. A **169**, 103 (1992); Phys. Rev. B **52**, 4740 (1995).

- [21] See, for example, F. Pierre, H. Pothier, D. Esteve, M.H. Devoret, A.B. Gougam, and N.O. Birge, in *Kondo effect and dephasing in low-dimensional metallic systems*, edited by V. Chandrasekhar, C. Van Haesendonck and A. Zawadowski, NATO Science Series, vol. II/50 [Kluwer Academic Publishers, Dordrecht, 2001].
- [22] J.M. Ziman, *Principles of the theory of solids*, 2nd edition [Cambridge University Press, Cambridge, 1972].
- [23] E.M. Lifshitz and L.P. Pitaevskii, *Physical Kinetics*, vol. 10 of the Landau and Lifshitz *Course of Theoretical Physics* [Pergamon Press, Oxford, 1981].
- [24] L.P. Kadanoff and G. Baym, *Quantum Statistical Mechanics* [Addison-Wesley, Redwood City, 1962].
- [25] J.B. Ketterson and S.N. Song, *Superconductivity* [Cambridge University Press, Cambridge 1999].
- [26] A.A. Abrikosov, L.P. Gorkov and I.E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* [Dover, New York, 1975].
- [27] M. Tinkham, Phys. Rev. B **6**, 1747 (1972).
- [28] A.L. Shelankov, J. Low Temp. Phys. **60**, 29 (1985).
- [29] K.D. Usadel, Phys. Rev. Lett. **25**, 507 (1970).
- [30] A.V. Zaitsev, Zh. Eksp. Teor. Fiz. **86**, 1742 (1984) [Sov. Phys. JETP **59**, 1015 (1985)].
- [31] M.Yu. Kuprianov and V.F. Lukichev, Zh. Eksp. Teor. Fiz. **94**, 139 (1988) [Sov. Phys. JETP **67**, 1163 (1988)].
- [32] B.L. Altshuler, P.A. Lee and R.A. Webb, *Mesoscopic Phenomena in Solids*, [North-Holland, Amsterdam, 1991].
- [33] See, for example, P. Charlat, H. Courtois, Ph. Gandit, D. Mailly, A.F. Volkov and B. Pannetier, Phys. Rev. Lett. **77**, 4950 (1996).
- [34] P.G. de Gennes, *Superconductivity of Metals and Alloys*, [Addison-Wesley, Redwood City, 1966].
- [35] S. Guéron, H. Pothier, N.O. Birge, D. Esteve, and M.H. Devoret, Phys. Rev. Lett. **77**, 3025 (1996).
- [36] A.F. Andreev, Zh. Eksp. Teor. Fiz. **46**, 1823 (1964) [Sov. Phys. JETP **19**, 1228 (1964)].
- [37] J.M. Martinis, G.C. Hilton, K.D. Irwin and D.A. Wollman, Nuclear Instruments and Methods in Physics Research A **444**, 23 (2000).
- [38] S. Guéron, PhD thesis, Université Paris 6, (1997).
- [39] S.-K. Yip, Phys. Rev. B **58**, 5803 (1998).
- [40] F.K. Wilhelm, G. Schön and A.D. Zaikin, Phys. Rev. Lett. **81**, 1682 (1998).
- [41] J.J.A. Baselmans, A.F. Morpugo, B.J. van Wees and T.M Klapwijk, Nature **397**, 43 (1999).
- [42] H. Pothier, S. Guéron, D. Esteve, and M.H. Devoret, Phys. Rev. Lett. **73**, 2488 (1994).

-
- [43] V.T. Petrashov, V.N. Antonov, P. Delsing, and T. Claeson, Phys. Rev. Lett. **74**, 5268 (1995).
 - [44] S.G. den Hartog, B.J. van Wees, T.M. Klapwijk, Yu.V. Nazarov and G. Borghs, Phys. Rev. B **56**, 13738 (1997).
 - [45] J. Eom, C-J Chien and V. Chandrasekhar, Phys. Rev. Lett. **81**, 437 (1998).
 - [46] F. Wilhelm, PhD thesis, Universität Karlsruhe, 2000.