Duality

Yury Holubeu *

October 7, 2024

This draft is not aimed for distribution.

Duality is discussed in details, especially duality in QFT, Strings by Álvarez-Gaumé Zamora, and Electromagnetic Duality for Children by J. M. Figueroa-O'Farrill. Problems and solutions are provided.

Duality was a topic that I studied in 2024 in the scope of the course "Advanced Field Theory" in KU Leuven. Later maybe I'll continue learning it, but it is too specific, big, and deep subject to study it without related research.

Before continuing this research I want to spend 1 week on differential geometry and 1 week on general preparation for special field theories.

Contents

1 Preface **7**

I Duality in a Nutshell **8**

2 Main dual constructions and formulas **8**
- 2.1 Main formulas and ideas of duality . . . 8
- 2.2 Special methods of duality . . . 13
- 2.3 Applications and examples in a nutshell . . . 17
  - 2.3.1 Main illustrative examples . . . 17
  - 2.3.2 Fermions in $N = 2$ rigid Yang-Mills theory . . . 19
  - 2.3.3 Example: duality transformations in $N = 1$ locally SUSY YM theories . . . 20
  - 2.3.4 Duality in electrodynamics in brief . . . 21
  - 2.3.5 On duality examples in Dirac, 't Hooft–Polyakov, BPS- monopoles . . . 21
  - 2.3.6 On duality examples in Poincare group . . . 22
  - 2.3.7 On Duality examples in black holes (??) . . . 22
  - 2.3.8 Overview of other applications . . . 22
- 2.4 Background for duality in a nutshell . . . 22
  - 2.4.1 On variation,.... . . . 22
  - 2.4.2 On supersymmetry and sypergravity basics . . . 22
- 2.5 Very special duality methods in a nutshell . . . 22
  - 2.5.1 On Effective Action for N=2 Supersymmetric Yang–Mills . . . 22
  - 2.5.2 On Monopoles for Arbitrary Gauge Groups by Figueroa-O'Farrill . . . 23

II Main Theory **24**

3 Duality in united formalism **24**

4 Most used theory of duality **24**
- 4.1 Duality in simple coupled theories by Freedman, Proeyen . . . 24
  - 4.1.1 Duality for one free electromagnetic field . . . 24
  - 4.1.2 Basic Properties of Dual Tensors . . . 25
  - 4.1.3 Duality for gauge field and complex scalar . . . 26
  - 4.1.4 Electromagnetic duality for coupled Maxwell fields . . . 28
- 4.2 Duality by Gaillard-Zumino . . . 34
  - 4.2.1 Introduction and formalism by Gaillard, Zumino . . . 34
  - 4.2.2 Derivations of properties of dual transformations by GZ . . . 37
  - 4.2.3 The conserved current . . . 41
  - 4.2.4 Invariance of the energy-momentum tensor . . . 43

*yuri.holubev@gmail.com

4.2.5 Construction of the lagrangian . . . 44
4.2.6 Compact duality rotations . . . 45
4.2.7 Non-linear realizations . . . 47
4.2.8 Non-compact duality transformations . . . 51
4.2.9 Invariance of a special derivative of a lagrangian . . . 53
4.2.10 Conclusion and examples . . . 54

III Questions and Problems in Duality **56**
4.3 Typical questions and problems . . . 56
4.3.1 Questions about understanding duality . . . 56
4.3.2 Problems about duality in electrodynamics . . . 56
4.3.3 Problems about duality in general field theories . . . 56

IV Other topics **57**

5 Duality Rotations in Nonlinear Electrodynamics and in Extended Supergravity by Aschieri, Ferrara, Zumino **57**
5.1 Introduction . . . 57
5.2 $U(1)$ gauge theory and duality symmetry . . . 60
5.2.1 Duality symmetry in nonlinear electromagnetism . . . 61
5.2.2 Legendre Transformations . . . 64
5.2.3 Hamiltonian theory . . . 65
5.2.4 Born-Infeld lagrangian . . . 66
5.2.5 Extended duality rotations . . . 67
5.3 General theory of duality rotations . . . 67
5.3.1 General nonlinear theory . . . 67
5.3.2 The main example and the scalar fields fractional transformations . . . 70
5.3.3 A basic example with fermi fields . . . 71
5.3.4 Compact and noncompact duality rotations . . . 72
5.3.5 Nonlinear sigma models on G/H . . . 75
5.3.6 Invariance of energy momentum tensor . . . 77
5.3.7 Generalized Born Infeld theory . . . 78
5.4 Dualities in $N > 2$ extended Supergravities . . . 84
5.4.1 Extended supergravities with target space G/H . . . 85
5.4.2 Specific cases . . . 92
5.5 Special Geometry and $N = 2$ Supergravity . . . 103
5.5.1 Special Geometry . . . 104
5.5.2 The $N = 2$ theory . . . 114
5.6 Duality rotations in Noncommutative Spacetime . . . 115
5.7 Appendix: Symplectic group and transformations . . . 119
5.7.1 Symplectic group (A, B, C, D and f, h and V matrices) . . . 119
5.7.2 The coset space $Sp(2n, \mathbb{R})/U(n)$ (M^* and $\mathcal{N}$ matrices) . . . 120
5.7.3 Lie algebra of $Sp(2n, \mathbb{R})$ and $U(n)$ (a, b, c, d matrices) . . . 122
5.8 Appendix: Unilateral Matrix Equations . . . 123

6 Classical Electromagnetic Duality for Children by JM Figueroa-O'Farrill **124**
6.1 1.1 The Dirac Monopole . . . 124
6.1.1 1.1.1 And in the beginning there was Maxwell... . . . 124
6.1.2 1.1.2 The Dirac quantisation condition . . . 126
6.1.3 1.1.3 Dyons and the Zwanziger-Schwinger quantisation condition . . . 127
6.2 1.2 The 't Hooft-Polyakov Monopole . . . 129
6.2.1 1.2.1 The bosonic part of the Georgi-Glashow model . . . 129

6.2.2 1.2.2 Finite-energy solutions: the 't Hooft-Polyakov Ansatz . . . 131
6.2.3 1.2.3 The topological origin of the magnetic charge . . . 134
6.3 1.3 BPS-monopoles . . . 136
6.3.1 1.3.1 Estimating the mass of a monopole: the Bogomol'nyi bound . . . 136
6.3.2 1.3.2 Saturating the bound: the BPS-monopole . . . 137
6.4 1.4 Duality conjectures . . . 138
6.4.1 1.4.1 The Montonen-Olive conjecture . . . 138
6.4.2 1.4.2 The Witten effect . . . 139
6.4.3 1.4.3 $SL(2,\mathbb{Z})$ duality . . . 141

7 Special geometry and symplectic transformations by de Wit, Van Proeyen 145
7.0.1 Introduction . . . 146
7.0.2 $N = 2$ actions . . . 147
7.0.3 Symplectic transformations . . . 150
7.0.4 Further results and conclusions . . . 157

8 Duality in supersymmetry for Children by JM Figueroa-O'Farrill 158
8.1 2.1 The super-Poincaré algebra in four dimensions . . . 158
8.1.1 2.1.1 Some notational remarks about spinors . . . 158
8.1.2 2.1.2 The Coleman-Mandula and Haag-LopuszańskiSohnius theorems . . . 160
8.2 2.2 Unitary representations of the supersymmetry algebra . . . 161
8.2.1 2.2.1 Wigner's method and the little group . . . 161
8.2.2 2.2.2 Massless representations . . . 162
8.2.3 2.2.3 Massive representations . . . 164
8.3 2.3 $N = 2$ Supersymmetric Yang-Mills . . . 167
8.3.1 2.3.1 $N = 1 d = 6$ supersymmetric Yang-Mills . . . 168
8.3.2 2.3.2 From $N = 1$ in $d = 6$ to $N = 2$ in $d = 4$ . . . 172
8.3.3 2.3.3 Higgsed $N = 2$ supersymmetric Yang-Mills . . . 173
8.3.4 2.3.4 $N = 2$ avatar of the BPS-monopole . . . 174
8.3.5 2.3.5 The supersymmetry bound is the Bogomol'nyi bound . . . 176
8.4 2.4 $N = 4$ Supersymmetric Yang-Mills . . . 180
8.4.1 2.4.1 $N = 1$ $d = 10$ supersymmetric Yang-Mills . . . 180
8.4.2 2.4.2 Reduction to $d = 4 : N = 4$ supersymmetric YangMills . . . 182
8.4.3 2.4.3 Monopoles and gauge bosons in $N = 4$ supersymmetric Yang-Mills . . . 186
8.4.4 2.4.4 The mass bound for $N = 4$ super Yang-Mills . . . 188

9 Collective Coordinates for Children by JM Figueroa-O'Farrill 190
9.1 3.1 The metric on the moduli space . . . 191
9.1.1 3.1.1 The physical configuration space . . . 191
9.1.2 3.1.2 The metric on the physical configuration space . . . 192
9.1.3 3.1.3 The metric on the moduli space . . . 193
9.1.4 3.1.4 The 1-monopole moduli space . . . 194
9.1.5 3.1.5 The quantisation of the effective action . . . 196
9.1.6 3.1.6 Some general properties of the monopole moduli space . . . 197
9.2 3.2 $\dim\mathcal{M}_k = 4k$ . . . 197
9.2.1 3.2.1 The dimension as an index . . . 197
9.2.2 3.2.2 Computing the index of $\mathcal{D}$ . . . 199
9.2.3 3.2.3 Computing the current $J_i(x,x)$ . . . 203
9.3 3.3 A quick motivation of hyperkähler geometry . . . 205
9.3.1 3.3.1 Riemannian geometry . . . 206
9.3.2 3.3.2 Kähler geometry . . . 208
9.3.3 3.3.3 Ricci flatness . . . 211
9.3.4 3.3.4 Hyperkähler geometry . . . 213
9.4 3.4 $\mathcal{M}_k$ is hyperkähler . . . 214

9.4.1 3.4.1 Symplectic quotients . . . 214
9.4.2 3.4.2 Kähler quotients . . . 215
9.4.3 3.4.3 Hyperkähler quotients . . . 217
9.4.4 3.4.4 $\mathcal{M}_k$ as a hyperkähler quotient . . . 218
9.4.5 3.4.5 Another proof that $\mathcal{M}_k$ is hyperkähler . . . 220

10 The Effective Action for $N = 2$ Supersymmetric Yang-Mills 222
10.1 4.1 Fermionic collective coordinates . . . 223
10.1.1 4.1.1 Computing the index . . . 223
10.1.2 4.1.2 Using supersymmetry . . . 224
10.1.3 4.2 The effective action . . . 226
10.2 4.3 $N = 4$ supersymmetry of the effective action . . . 228
10.2.1 4.3.1 $N = 4$ supersymmetry in $\mathbb{R}^4$: a toy model . . . 229
10.2.2 4.3.2 $N = 4$ supersymmetry in hyperkähler manifolds . . . 230
10.2.3 4.3.3 $N = 4$ supersymmetry of L_{eff} . . . 231
10.3 4.4 A brief review of harmonic theory . . . 233
10.3.1 4.4.1 Harmonic theory for riemannian manifolds . . . 233
10.3.2 4.4.2 Harmonic theory for Kähler manifolds . . . 235
10.3.3 4.4.3 Explicit formulas for $\bar{\partial}$ and $\bar{\partial}^*$ . . . 238
10.3.4 4.5 Quantisation of the effective action . . . 239
10.3.5 4.5.1 Canonical analysis . . . 239
10.3.6 4.5.2 The quantisation of the effective hamiltonian . . . 241

11 The Effective Action for $N = 4$ Supersymmetric Yang-Mills 243
11.0.1 5.1 Fermionic collective coordinates . . . 243

12 Monopoles for Arbitrary Gauge Groups for Children by JM Figueroa-O'Farrill 245
12.1 6.1 Topologically stable solutions . . . 245
12.1.1 6.1.1 Some elements of homotopy . . . 247
12.1.2 6.1.2 Homotopy classification of finite-energy configurations . . . 253
12.2 6.2 The Dirac quantisation condition . . . 260
12.3 6.3 Some facts about compact Lie groups and Lie algebras . . . 262
12.3.1 6.3.1 Compact Lie groups . . . 262
12.3.2 6.3.2 The Weyl group . . . 264
12.3.3 6.3.3 Root systems and simple Lie algebras . . . 266
12.3.4 6.3.4 Some simple examples . . . 273
12.3.5 6.4 The magnetic dual of a compact Lie group . . . 275

13 Duality Transformations in Supersymmetric Yang-Mills Theories coupled to Supergravity by Ceresole, Ferrara, Proeyen 282
13.1 1 Introduction . . . 283
13.2 2 Résumé of rigid special geometry . . . 284
13.2.1 2.1 Basics . . . 284
13.2.2 2.2 Symplectic transformations in the fermionic sector . . . 287
13.2.3 2.3 Fermions in $N = 2$ rigid Yang-Mills theory . . . 288
13.2.4 2.4 Positivity and monodromies . . . 289
13.3 3 Coupling to gravity . . . 290
13.3.1 3.1 Special geometry and symplectic transformations . . . 290
13.3.2 3.2 The fermionic sector . . . 294
13.3.3 3.3 The three-form cohomology . . . 295
13.3.4 3.4 Duality transformations in $N = 1$ locally supersymmetric Yang-Mills theories 296
13.4 4 Duality symmetries . . . 298
13.4.1 4.1 The facts . . . 298
13.4.2 4.2 Heterotic $N = 2$ superstring theories . . . 299

13.4.3 4.3 Duality symmetries in $N > 2$ supergravities . . . 302
13.5 5 On monodromies in string effective field theories . . . 302
13.5.1 5.1 Classical and quantum monodromies . . . 302
13.5.2 5.2 The BPS mass formula . . . 306
13.6 6 Conclusions . . . 308

14 Other methods and theories with duality 309

15 5d/4d U-dualities and $\mathcal{N}=8$ black holes Anna Ceresole, Ferrara, Gnecchi 309
15.0.1 Introduction . . . 309
15.0.2 4d/5d relations for the $\mathcal{N} = 8$ extremal black hole potential . . . 312
15.0.3 Attractors in the 5 dimensional theory . . . 316
15.0.4 Attractors in the 4 dimensional theory . . . 317
15.0.5 Maurer-Cartan equations of the four dimensional theory . . . 322
15.1 Self-Duality in Nonlinear Electromagnetism by Gaillard Zumino . . . 325
15.1.1 Duality rotations in four dimensions . . . 325
15.1.2 Born-Infeld theory . . . 327
15.1.3 Schrödinger's formulation of Born's theory . . . 328
15.1.4 Axion, dilaton and $SL(2, R)$ . . . 329
15.1.5 Connections to string theory . . . 331
15.2 Nonlinear Electromagnetic Self-Duality and Legendre Transformations by Gaillard Zumino . . . 331
15.2.1 Duality rotations in four dimensions . . . 331
15.2.2 Born-Infeld theory . . . 334
15.2.3 Schrödinger's formulation of Born's theory . . . 335
15.2.4 General solution of the self-duality equation . . . 336
15.2.5 Axion, dilaton and $SL(2, R)$ . . . 337
15.2.6 Duality as a Legendre transformation . . . 339
15.2.7 Concluding remarks . . . 342
15.3 Exact electromagnetic duality by David I Olive . . . 342
15.3.1 Introduction . . . 342
15.3.2 Zamolodchikov's Principle and Solitons . . . 344
15.3.3 Magnetic Charge and its Quantisation . . . 345
15.3.4 Magnetic Monopoles (and dyons) as Solitons . . . 346
15.3.5 Electromagnetic Duality Conjectures . . . 348
15.3.6 Catechism concerning the Duality Conjecture . . . 350
15.3.7 More on Supersymmetry and $N = 2$ versus $N = 4$ . . . 351
15.3.8 The Schwinger Quantisation Condition and the Charge Lattice . . . 353
15.3.9 Exact Electromagnetic Duality and the Modular Group . . . 355
15.3.10 Conclusion . . . 356
15.4 Duality in Quantum Field Theory and String Theory by Álvarez-Gaumé Zamora . . . 358
15.4.1 The duality symmetry. . . . 358
15.4.2 Dirac's charge quantization. . . . 359
15.4.3 A charge lattice and the $SL(2, \mathbf{Z})$ group. . . . 361
15.4.4 The Higgs Phase . . . 363
15.4.5 The Georgi-Glashow model and the Coulomb phase. . . . 365
15.4.6 The 't Hooft-Polyakov monopoles . . . 367
15.4.7 The Confining phase. . . . 371
15.4.8 The Higgs/confining phase. . . . 374
15.4.9 Supersymmetry . . . 374
15.4.10 The uses of supersymmetry. . . . 380
15.4.11 $N = 1$ SQCD. . . . 383
15.4.12 The vacuum structure of SQCD with $N_f < N_c$. . . . 386
15.4.13 The vacuum structure of SQCD with $N_f = N_c$. . . . 389

15.4.14 The vacuum structure of SQCD with $N_f = N_c + 1$. . . . 391
15.4.15 Seiberg's duality. . . . 393
15.4.16 $N = 2$ supersymmetry. . . . 394
15.4.17 $N = 2$ $SU(2)$ super Yang-Mills theory in perturbation theory. . . . 396
15.4.18 The low energy effective Lagrangian. . . . 398
15.4.19 BPS bound and duality. . . . 399
15.4.20 Singularities in the moduli space. . . . 401
15.4.21 The physical interpretation of the singularities. . . . 401
15.4.22 The Seiberg-Witten solution. . . . 403
15.4.23 Breaking $N = 2$ to $N = 1$. Monopole condensation and confinement. . . . 404
15.4.24 Breaking $N = 2$ to $N = 0$. . . . 406
15.4.25 String Theory in perturbation theory. . . . 408
15.4.26 D-branes. . . . 412
15.4.27 Some final comments on nonperturbative String Theory. . . . 414
15.5 Other dualities 415
15.5.1 Duality in Poincare group 415
15.5.2 Duality and coset manifolds (???) 415

V Appendix 416
.1 Additional general information 416
.1.1 Additional motivation 416
.1.2 Literature 416
.1.3 History of research in duality 417
.2 Мышление профессионала в дуальности 417
.2.1 Насколько же на самом деле дуальность полезна и работает? 417
.2.2 Об изучении дуальности 417
.3 Mathematical tricks for duality 417
.3.1 On variational methods in duality (!?) 417
.3.2 Hodge duality of forms 418
.3.3 Properties of the non-compact real symplectic group 420

A Bibliography 422

B Bibliography 422

C Bibliography 424

D Bibliography 426

E Bibliography 434

F Bibliography 435

G Bibliography 435

1 Preface

Description of the note

Overview of applications

We will specify the applications
(I'll reveal it later)

Main motivation

Обсудим самые типичные причины, зачем изучать дуальность?

Amazing facts (I'll reveal it later)

Puzzles for motivation

Обсудим теоретические задачи, которые мотивируют изучать дуальность.

Why Pauli-Lubanski vector works? (напишу, что мол до группы Пуанкаре можно то дойти, а вот что дальше делать - непонятно, а с помощью дуальности можно догадаться до Паули-Любанского вектора и так построить казимиры. Почему такое работает?)

Part I
Duality in a Nutshell

2 Main dual constructions and formulas

2.1 Main formulas and ideas of duality

Main idea of duality (!??!?!?)

(тут самые важные общие слова! потом напишу их)

$$\tilde{F}_{\mu\nu} = \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}F^{\rho\sigma},$$

Main formulas of Duality

Main basic properties of dual tensors

For now the convection is: $\tilde{H}^{\mu\nu} := -\frac{1}{2}\mathrm{i}\varepsilon^{\mu\nu\rho\sigma}H_{\rho\sigma}$

$$H^{\pm}_{\mu\nu} := \frac{1}{2}\left(H_{\mu\nu} \pm \tilde{H}_{\mu\nu}\right), \qquad H^{\pm}_{\mu\nu} := \left(H^{\mp}_{\mu\nu}\right)^{*}.$$

$$-\frac{1}{2}\mathrm{i}\varepsilon^{\mu\nu\rho\sigma}\tilde{H}_{\rho\sigma} = H^{\mu\nu}.$$

$$-\frac{1}{2}\mathrm{i}\varepsilon_{\mu\nu}{}^{\rho\sigma}H^{\pm}_{\rho\sigma} = \pm H^{\pm}_{\mu\nu}.$$

$$G^{+\mu v}H^{-}{}_{\mu\nu} = 0,$$

$$G^{\pm\rho(\mu}H^{\pm\nu)}{}_{\rho} = -\frac{1}{4}\eta^{\mu\nu}G^{\pm\rho\sigma}H^{\pm}{}_{\rho\sigma}, \qquad G^{+}{}_{\rho[\mu}H^{-}{}_{\nu]}{}^{\rho} = 0,$$

$$\tilde{G}^{\rho\mu}\tilde{H}^{v}{}_{\rho} = -\tfrac{1}{2}\eta^{\mu v}G^{\rho\sigma}H_{\rho\sigma} - G^{\rho v}H^{\mu}{}_{\rho}.$$

$$F\tilde{G} = \tilde{F}G.$$

This is because $F\tilde{G} \equiv \frac{1}{2}F^{\mu\nu}\varepsilon_{\mu\nu\rho\sigma}G^{\rho\sigma} = \frac{1}{2}G^{\rho\sigma}\varepsilon_{\rho\sigma\mu\nu}F^{\mu\nu} = \tilde{F}G.$

For any symmetric matrix B ($B = B^T$) and arbitrary matrix Gs

$$\frac{\partial(GB\tilde{G})}{\partial\chi} = 2\frac{\partial G}{\partial\chi}B\tilde{G}, \qquad \Leftrightarrow \qquad \frac{\partial G}{\partial\chi}B\tilde{G} = \frac{1}{2}\frac{\partial(GB\tilde{G})}{\partial\chi}$$

Proof:

$$GB\frac{\partial\tilde{G}}{\partial\chi} \equiv G^{\mu\nu}B\frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}\frac{\partial G^{\rho\sigma}}{\partial\chi} = \frac{\partial G^{\mu\nu}}{\partial\chi}B\frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}G^{\rho\sigma} = \frac{\partial G}{\partial\chi}B\tilde{G}$$

$$\frac{1}{2}\frac{\partial(GB\tilde{G})}{\partial\chi} = \frac{1}{2}\left(\frac{\partial G}{\partial\chi}B\tilde{G} + GB\frac{\partial\tilde{G}}{\partial\chi}\right) = \frac{\partial G}{\partial\chi}B\tilde{G}$$

5. For (?? which matrices???)

$$\tilde{F}BG = \tilde{G}B^T F.$$

Proof:

$$\tilde{F}BG = \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}F^{a\rho\sigma}B^{ab}G^{b\mu\nu} = \frac{1}{2}\varepsilon_{\rho\sigma\mu\nu}F^{a\mu\nu}B^{ab}G^{b\rho\sigma} = \frac{1}{2}\varepsilon_{\rho\sigma\mu\nu}G^{b\rho\sigma}B^{T\,ba}F^{a\mu\nu} = \tilde{G}B^T F.$$

$$F^T A^T\frac{\delta S}{\delta F} = (F^{+T}{+}F^{-T})_{\mu\nu}A^T\frac{(-\mathrm{i})}{2}\tilde{G}^{\mu\nu} = (F^{+T}{+}\,F^{-T})_{\mu\nu}A^T\frac{(-\mathrm{i})}{2}(G^{+\mu\nu}{-}\,G^{-\mu\nu}) = \frac{(-\mathrm{i})}{2}F^{+T}A^T G^{+} + \text{h.c.}$$

Dual tensors are not independent form original one,

$$\frac{\delta F^{+}_{\mu\nu}}{\delta F_{\rho\sigma}} = \frac{\delta}{\delta F_{\rho\sigma}}\frac{1}{2}\left(F_{\mu\nu} + \frac{-\mathrm{i}}{2}\varepsilon_{\mu\nu\tau\phi}F^{\tau\phi}\right) = \frac{1}{2}\left(\delta^{\rho\sigma}_{\mu\nu} + \frac{-\mathrm{i}}{2}\varepsilon_{\mu\nu}{}^{\rho\sigma}\right)$$

Duality for one free electromagnetic field We know that $\partial_\mu F^{\mu\nu} = 0, \partial_\mu\tilde{F}^{\mu\nu} = 0$. There is a dual symmetry - change of variables:

$$F^{\mu\nu} \to F'^{\mu\nu} = \mathrm{i}\tilde{F}^{\mu\nu}$$

(the “i” is included to make the transformation real).

Since $F'^{\mu\nu}$ also obeys both equations of (4.41) we have defined a symmetry of the free electromagnetic field.

The symmetry (4.42) exchanges the electric and magnetic fields: $E_i \to E'_i = -B_i$ and $B_i \to B'_i = E_i$.

$$\partial_\mu F^{\mu v} = 0, \quad \partial_\mu\tilde{F}^{\mu v} = 0.$$

We can now consider the change of variables (the i is included to make the transformation real):

$$F^{\mu v} \to F'^{\mu\nu} = \mathrm{i}\tilde{F}^{\mu v}.$$

Since $F'^{\mu\nu}$ also obeys both equations of (4.41) we have defined a symmetry of the free electromagnetic field.

The symmetry (4.42) exchanges the electric and magnetic fields: $E_i \to E'_i = -B_i$ and $B_i \to B'_i = E_i$.

The self-dual combinations $F^{\pm}_{\mu\nu}$contain only photons of one polarization in their plane wave expansions:

$$F^{\pm}_{\mu\nu} = 2\mathrm{i}\int\frac{\mathrm{d}^3k}{(2\pi)^3 2k^0}\left[\mathrm{e}^{\mathrm{i}k\cdot x}k_{[\mu}\epsilon_{v]}(\vec{k},\pm)a(\vec{k},\pm) - \mathrm{e}^{-\mathrm{i}k\cdot x}k_{[\mu}\epsilon^{*}_{\nu]}(\vec{k},\mp)a^{*}(\vec{k},\mp)\right].$$

To perform this exercise, check first that with the polarization vectors given in Sec. 4.1.2, one has

$$-\frac{1}{2}\mathrm{i}\varepsilon^{\mu\nu\rho\sigma}k_\rho\epsilon_\sigma(\vec{k},\pm) = \pm k^{[\mu}\epsilon^{\nu]}(\vec{k},\pm).$$

$$F_{\mu\nu}\tilde{F}^{\mu\nu} = -\mathrm{i}\partial_{\mu}\left(\varepsilon^{\mu\nu\rho\sigma}A_{v}F_{\rho\sigma}\right).$$

$$\left(F_{\mu\nu}\tilde{F}^{\mu\nu}\right)(x) \to \det\Lambda^{-1}\left(F_{\mu\nu}\tilde{F}^{\mu\nu}\right)(\Lambda x).$$

Thus $F_{\mu\nu}\tilde{F}^{\mu\nu}$ transforms as a scalar under proper Lorentz transformations but changes sign under space or time reflections.

$$F_{\mu\rho}\tilde{F}^{\rho}_{\nu} = \frac{1}{4}\eta_{\mu\nu}F_{\rho\sigma}\tilde{F}^{\rho\sigma}.$$

General duality rotations

$$L = L\left(F^{a}, \chi^{i}, \chi^{i}_{\mu}\right).$$

$$F^{a}_{\mu\nu} = \partial_{\mu}\mathscr{A}^{a}_{\nu} - \partial_{\nu}\mathscr{A}^{a}_{\mu},$$

$$\partial^{\mu}\tilde{F}^{a}_{\mu\nu} = 0.$$

$$\tilde{G}^{a}_{\mu\nu} = \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}G^{a\rho\sigma} \equiv 2\frac{\partial L}{\partial F^{a\mu\nu}},$$

$$\partial^{\mu}\tilde{G}^{a}_{\mu\nu} = 0.$$

$$\delta\begin{pmatrix}F\\G\end{pmatrix} = \begin{pmatrix}A & B\\C & D\end{pmatrix}\begin{pmatrix}F\\G\end{pmatrix},$$

$$\delta\chi^{i} = \xi^{i}(\chi),$$

Covariance of the equations of motion:

$$C = C^{\mathrm{T}}, \quad B = B^{\mathrm{T}},$$

$$D^{ab} + A^{ba} = \varepsilon\delta^{ab},$$

$$\frac{\partial}{\partial F^{a}}\delta L = \frac{\partial}{\partial F^{a}}\left(\frac{1}{4}FC\tilde{F} + \frac{1}{4}GB\tilde{G} + \varepsilon L\right).$$

$$E_{i} = \frac{\partial L}{\partial\chi^{i}} - \partial_{\mu}\frac{\partial L}{\partial\chi^{i}_{\mu}}$$

$$\delta L = \frac{1}{4}(FC\tilde{F} + GB\tilde{G})$$

Proof:
(?)

Conserved current under duality rotations

$$\hat{J}^\mu = \frac{1}{2}\left(\tilde{G}^{\mu\nu} A \mathscr{A}_\nu - \tilde{F}^{\mu\nu} C \mathscr{A}_\nu + \tilde{G}^{\mu\nu} B \mathscr{B}_\nu - \tilde{F}^{\mu\nu} D \mathscr{B}_\nu\right)$$

$$J^\mu = \xi^i \frac{\partial L}{\partial \chi^i_\mu} + \hat{J}^\mu$$
$$\partial_\mu J^\mu = 0.$$

Proof:
(?)
(???, see theory about it, ???)

Duality formalizm in vector+spinor+scalar field theory We considered there actions $S(F)$ that depend on field strengths $F^A_{\mu\nu}$, which are determined in terms of (abelian) vectors A^A_μ. We consider actions at most quadratic in spacetime derivatives, and thus also at most quadratic in $F^A_{\mu\nu}$.

$$\tilde{F}_{\mu\nu} = -\frac{1}{2}\mathrm{i}e\varepsilon_{\mu\nu\rho\sigma} F^{\rho\sigma}, \quad F^{\pm}_{\mu\nu} \equiv \frac{1}{2}\left(F_{\mu\nu} \pm \tilde{F}_{\mu\nu}\right),$$

$$\begin{aligned} \nabla^\mu \operatorname{Im} F^{A+}_{\mu\nu} &= 0 \quad \text{Bianchi identities,} \\ \nabla_\mu \operatorname{Im} G^{\mu\nu+}_A &= 0 \quad \text{Equations of motion of } A^A_\nu, \end{aligned}$$

$$G^{+\mu\nu}_A := 2\mathrm{i}e^{-1}\frac{\delta S\left(F^+, F^-, \phi\right)}{\delta F^{+A}_{\mu\nu}}, \quad \text{i.e.} \quad \tilde{G}^{\mu\nu}_A := 2\mathrm{i}e^{-1}\frac{\delta S(F,\phi)}{\delta F^A_{\mu\nu}}.$$

$S(F,\phi) = S\left(F^+, F^-, \phi\right)$
$\mathcal{N}_{AB} = -\mathrm{i}\bar{f}_{AB}$

$$G^+_{A\mu\nu} = \mathcal{N}_{AB}(\phi) F^{+B}_{\mu\nu} + H^+_{A\mu\nu}(\phi) = G^+_{\mathrm{b}A\mu\nu} + H^+_{A\mu\nu}(\phi).$$

$$\delta_d \begin{pmatrix} F^A_{\mu\nu} \\ G_{A\mu\nu} \end{pmatrix} = \begin{pmatrix} A^A_B & B^{AB} \\ C_{AB} & D^B_A \end{pmatrix} \begin{pmatrix} F^B_{\mu\nu} \\ G_{B\mu\nu} \end{pmatrix},$$
$$B^{AB} = B^{BA}, \quad C_{AB} = C_{BA}, \quad D^B_A = -A^A_B.$$

$$E_i \equiv \frac{\delta S}{\delta \phi^i},$$

$$S = \int \mathrm{d}^4 x e L\left(\phi, \partial_\mu \phi\right), \quad \frac{\delta S}{\delta \phi^i} = e\left[\frac{\delta L}{\delta \phi^i} - \nabla_\mu \frac{\delta L}{\delta \partial_\mu \phi^i}\right].$$

$$\delta_d \phi^i = \xi^i(\phi),$$

$$\delta_d = \xi^i \frac{\delta}{\delta \phi^i} + \left(A^A_B F^B_{\mu\nu} + B^{AB} G_{B\mu\nu}\right)\frac{\delta}{\delta F^A_{\mu\nu}},$$

$$\begin{aligned} \delta_d S &= \left(\xi^i \frac{\delta}{\delta \phi^i} + \left(F^T A^T + G^T B\right)\frac{\delta}{\delta F}\right) S \\ &= \xi^i E_i - \frac{1}{2}\left(\mathrm{i}\left(F^{+T} A^T + G^{+T} B\right) G^+ + \text{h.c. }\right) \end{aligned}$$

$$\frac{\delta}{\delta F^+} S = -\frac{1}{2}\mathrm{i}G^+, \quad \frac{\partial G_A^+}{\partial F^{+B}} = \mathcal{N}_{AB}.$$

$\partial_i = \frac{\delta}{\delta\phi^i}$,

$$\frac{\delta}{\delta F^+}\delta_d = \delta_d\frac{\delta}{\delta F^+} + \left(A^T + \mathcal{N}B\right)\frac{\delta}{\delta F^+}$$

$$\partial_i\delta_d = \delta_d\partial_i + \left(\partial_i\xi^j\right)\partial_j + \left(\partial_i G^{+T} B\frac{\delta}{\delta F^+} + \text{h.c.}\right).$$

$$\frac{\delta}{\delta F^+}\delta_d S = -\frac{1}{2}\mathrm{i}\left(\delta_d G^+ + \left(A^T + \mathcal{N}B\right)G^+\right) = -\frac{1}{2}\mathrm{i}\left[CF^+ + \mathcal{N}\mathcal{B}BG^+\right]$$

$$= -\frac{1}{4}\mathrm{i}\frac{\delta}{\delta F^+}\left[F^{+T}CF^+ + G^{+T}BG^+\right]$$

$$\partial_i\delta_d S = \delta_d E_i + \left(\partial_i\xi^j\right)E_j - \frac{1}{2}\left(\mathrm{i}\left(\partial_i G^{+T}B\right)G^+ + \text{h.c.}\right)$$

$$\delta_d S = -\frac{1}{4}\mathrm{i}\left[F^{+T}CF^+ + G^{+T}BG^+\right] + \text{h.c.}\ , \quad \delta_d E_i = -\left(\partial_i\xi^j\right)E_j.$$

$$\delta_d S = -\frac{1}{8}\int \mathrm{d}^4x\varepsilon^{\mu\nu\rho\sigma}\left(F^A_{\mu\nu}C_{AB}F^B_{\rho\sigma} + G_{A\mu\nu}D^{AB}G_{B\rho\sigma}\right),$$

$$S_{\text{non-inv}} = -\frac{1}{4}\mathrm{i}F^{+A}G^+_A + \text{h.c.}\ = S_2(F,\phi) - \frac{1}{4}\left(\mathrm{i}F^{+A}H^+_A + \text{h.c.}\right),$$

$$S_2(F,\phi) = -\frac{1}{4}\mathrm{i}F^{+A}G^+_{\mathrm{b}A} + \text{h.c.}\ = -\frac{1}{4}\mathrm{i}F^{+A}\mathcal{N}_{AB}F^{+B} + \text{ h.c.}$$

$$= \frac{1}{4}\int \mathrm{d}^4x\left[e\left(\mathrm{Im}\,\mathcal{N}_{AB}\right)F^A_{\mu\nu}F^{B\mu\nu} - \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}\left(\mathrm{Im}\,\mathcal{N}_{AB}\right)F^A_{\mu\nu}F^B_{\rho\sigma}\right],$$

$$S_1(F,\phi) = -\frac{1}{2}\mathrm{i}F^{+A}H^+_A + \text{h.c.}\ = -\frac{1}{4}\int \mathrm{d}^4x\varepsilon^{\mu\nu\rho\sigma}F^A_{\mu\nu}H_{A\rho\sigma}.$$

$$S_{\text{inv}} = \frac{1}{2}S_1(F,\phi) + S_0(\phi),$$

Symplectic vector notation

$$\delta_d\mathcal{N}_{AB}(\phi) = \xi^i\partial_i\mathcal{N}_{AB} = \left(C - \mathcal{N}A - A^T\mathcal{N} - \mathcal{N}B\mathcal{N}\right)_{AB}$$

$$\delta_d H^+(\phi) = \xi^i\partial_i H^+ = -\left(A^T + \mathcal{N}B\right)H^+.$$

$$\delta_d\frac{1}{2}S_1(F,\phi) = -\frac{1}{4}\mathrm{i}H^{+T}BH^+,$$

$\left(Q^{+A}, P^+_A\right)$
$P^+_A = \overline{\mathcal{N}}_{AB}Q^{+B}$.

$$H^+_{\mu\nu} = \left(\mathcal{N}Q^+ - P^+\right)_{\mu\nu} = 2\mathrm{i}(\mathrm{Im}\,\mathcal{N})Q^+_{\mu\nu}.$$

$$\frac{1}{2}S_1(F,\phi) = -\frac{1}{4}\mathrm{i}F^{+A}\left(\mathcal{N}Q^+ - P^+\right) = \frac{1}{4}\mathrm{i}\left(F^{+A}P^+_A - G^{+A}_{\mathrm{b}}Q^+_A\right) + \text{h.c.}\ .$$

If we replace G_{b} by the full G, this is a symplectic invariant.

$$S_0(\phi) = \frac{1}{4}\mathrm{i}H^+_{A\mu\nu}Q^{+A\mu\nu} + \text{ h.c. } + S_{0,\text{ inv}}\ .$$

Invariance of the energy-momentum tensor

$$\theta^{\mu}_{\lambda} = -\chi^{i}_{\lambda}\frac{\partial L}{\partial \chi^{i}_{\mu}} + \delta^{\mu}_{\lambda} L + \tilde{G}^{a\mu\nu} F^{a}_{\nu\lambda}$$

is conserved as a consequence of the equations of motion (2.3), (2.5) and (2.15):

$$\partial_{\mu}\theta^{\mu}_{\lambda} = 0.$$

(??? в чем идея в иоге его??? и идея вывода???)

On construction of the lagrangian

$$L = \frac{1}{4}F\tilde{G} + \frac{1}{4}(FI - GH) + L_{\text{inv}}(\chi),$$

$$jG - I = (F + jH)\frac{\partial jG}{\partial F}.$$

$$L = -\frac{1}{4}FKF + \frac{1}{2}F(I - jKH) + \frac{1}{4}jH(I - jKH) + L_{\text{inv}}(\chi).$$

(??? тоже впишу идею, что происходит???)

2.2 Special methods of duality

(потом сделаю несколько разделов, сгруппирую методы, пока тут методы, которые я плохо понимаю)

Basics

$N = 2$ supersymmetric gauge theory on a group G broken to $U(1)^r$, with $r = \operatorname{rank} G$, corresponds to a particular case of the most general $N = 1$ coupling of r chiral multiplets $\left(X^A, \chi^A\right)$ to $rN = 1$ abelian vector multiplets $\left(\mathcal{A}^A_{\mu}, \lambda^A\right)$ in which the Kähler potential K and the holomorphic kinetic term function $f_{AB}\left(X^A\right)$ are given by

$$\begin{aligned} K &= i\left(\bar{F}_A X^A - F_A \bar{X}^A\right), \quad (F_A = \partial_A F) \\ f_{AB} &= \partial_A \partial_B F \equiv F_{AB} \end{aligned} \tag{2.1}$$

$$R_{A\bar{B}C\bar{D}} = -\partial_A \partial_C \partial_P F \partial_{\bar{B}} \partial_{\bar{D}} \partial_{\bar{Q}} \bar{F} g^{P\bar{Q}} \tag{2.2}$$

$$g_{P\bar{Q}} = \partial_P \partial_{\bar{Q}} K = 2 \operatorname{Im} \partial_P \partial_Q F \tag{2.3}$$

$$\begin{aligned} \mathcal{L} = {} & g_{A\bar{B}} \partial_{\mu} X^A \partial_{\mu} \bar{X}^B + \left(g_{A\bar{B}} \lambda^{IA} \sigma^{\mu} \mathcal{D}_{\mu} \bar{\lambda}^{\bar{B}}_{I} + \text{ h.c. }\right) \\ & + \operatorname{Im}\left(F_{AB} \mathcal{F}^{-A}_{\mu\nu} \mathcal{F}^{-B}_{\mu\nu}\right) + \mathcal{L}_{\text{Pauli}} + \mathcal{L}_{4-\text{ fermi}} \end{aligned} \tag{2.4}$$

where $A, B, \ldots$ run on the adjoint representation of the gauge group $G, I = 1, 2$ and $\mathcal{F}^{+A}_{\mu\nu} = \mathcal{F}^{A}_{\mu\nu} - \frac{i}{2}\epsilon_{\mu\nu\rho\sigma}\mathcal{F}^{A\rho\sigma}$ (and $\mathcal{F}^{-A}_{\mu\nu} = \overline{\mathcal{F}}^{+A}_{\mu\nu}$). As we shall see, also $\mathcal{L}_{\text{Pauli}}$ and $\mathcal{L}_{4-\text{ Fermi}}$ contain the function F and its derivatives up to the fourth.

The previous formulation, derived from tensor calculus, is incomplete because it is not coordinate covariant. It is written in a particular coordinate system ("special coordinates") which is not uniquely selected. In fact, eq.(2.1) is left invariant under particular coordinate changes of the $X^A \to \widetilde{X}^A$ with some new function $\widetilde{F}(\widetilde{X})$ described by

$$\begin{aligned} \widetilde{X}^A(X) &= A^A_B X^B + B^{AB} F_B(X) + P^A \\ \widetilde{F}_A\left(\widetilde{X}^A(X)\right) &= C_{AB} X^B + D^B_A F_B(X) + Q_A \end{aligned} \tag{2.5}$$

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in Sp(2r, \mathbf{R})$$

$$A^T C - C^T A = 0 \quad , \quad B^T D - D^T B = 0 \quad , \quad A^T D - C^T B = \mathbb{1} \tag{2.6}$$

and P^A, Q_A can be complex constants which from now on will be set to zero.
It can be shown that a function $\widetilde{F}$ exists such that [3]

$$\widetilde{F}_A = \frac{\partial \widetilde{F}}{\partial \widetilde{X}^A} \tag{2.7}$$

provided the mapping $X^A \to \widetilde{X}^A$ is invertible.

$$\begin{aligned} \partial^\mu \operatorname{Im} \mathcal{F}^{-A}_{\mu\nu} &= 0 \quad \text{Bianchi identities} \\ \partial_\mu \operatorname{Im} G^{\mu\nu}_{-A} &= 0 \quad \text{Equations of motion} \end{aligned} \tag{2.8}$$

$\left(\mathcal{F}^{-A}_{\mu\nu}, G^{\mu\nu}_{-A}\right)$
$G^{\mu\nu}_{-A} \equiv i \frac{\delta \mathcal{L}}{\delta \mathcal{F}^{-A}_{\mu\nu}} = \overline{\mathcal{N}}_{AB} \mathcal{F}^{-B}_{\mu\nu} +$
$\mathcal{F}^{-A}_{\mu\nu} = \mathcal{F}^A$ and $G^{\mu\nu}_{-A} = G_A$

$$\begin{aligned} \operatorname{Im} \mathcal{F}^A \overline{\mathcal{N}}_{AB} \mathcal{F}^B \to \operatorname{Im} \widetilde{\mathcal{F}}^A \widetilde{G}_A &= \\ &= \operatorname{Im} \Big(\mathcal{F}^A G_A + 2 \mathcal{F}^A \left(C^T B\right)^B_A G_B + \\ &+ \mathcal{F}^A \left(C^T A\right)_{AB} \mathcal{F}^B + G_A \left(D^T B\right)^{AB} G_B \Big) \end{aligned} \tag{2.9}$$

If $C = B = 0$ the lagrangian is invariant. If $C \neq 0, B = 0$ it is invariant up to a four-divergence.

In presence of a topologically non-trivial $\mathcal{F}^{-A}_{\mu\nu}$ background, $\left(C^T A\right)_{AB} \int \operatorname{Im} \mathcal{F}^{-A}_{\mu\nu} \mathcal{F}^{-B}_{\mu\nu} \neq 0$, one sees that in the quantum theory duality transformations must be integral valued in $Sp(2r, \mathbb{Z})$ [1] and transformations with $B = 0$ will be called perturbative duality transformations.

If $B \neq 0$ the lagrangian is not invariant. As it is well known, then the duality transformation is only a symmetry of the equations of motion and not of the lagrangian.

$\widetilde{G}^{\mu\nu}_{-A} = \widetilde{\mathcal{N}}_{AB}\widetilde{\mathcal{F}}^{-B}_{\mu\nu}$,

$$\widetilde{\mathcal{N}} = (C + D\mathcal{N})(A + B\mathcal{N})^{-1} \tag{2.10}$$

$\widetilde{\mathcal{N}}(\widetilde{X}) = \mathcal{N}(\widetilde{X})$,
$\widetilde{F}(\widetilde{X}) = F(\widetilde{X})$.

$$\begin{pmatrix} A & 0 \\ C & \left(A^T\right)^{-1} \end{pmatrix}, A \subset GL(r), A^T C \text{ symmetric} \tag{2.11}$$

$\begin{pmatrix} A & 0 \\ 0 & \left(A^T\right)^{-1} \end{pmatrix}$

The general form of the central charge for BPS states in a generic $N = 2$ rigid theory is given by

$$|Z| = M = \left| n^A_{(m)} F_A - n^{(e)}_A X^A \right| \tag{2.12}$$

$A^T C = \langle \alpha_i \mid \alpha_j \rangle$ of $SU(r+1)$

$$U_i \equiv \partial_i V = \left(\partial_i X^A, \partial_i F_A\right) \qquad \text{with } i = 1, \ldots, r \tag{2.13}$$

In rigid special geometry the U_i satisfy the constraints

$$\begin{aligned} \mathcal{D}_i U_j &= i C_{ijk} g^{k\bar{l}} \bar{U}_{\bar{l}} \\ \partial_i \bar{U}_{\bar{j}} &= 0 \end{aligned} \tag{2.14}$$

$$\begin{aligned} g_{i\bar{j}} = \partial_i \partial_{\bar{j}} K &= i \left(\partial_{\bar{j}} \bar{F}_A \partial_i X^A - \partial_{\bar{j}} \bar{X}^A \partial_i F_A\right) \\ &= i \partial_i X^A \partial_{\bar{j}} \bar{X}^B \left(\mathcal{N}_{AB} - \overline{\mathcal{N}}_{AB}\right) \end{aligned} \tag{2.15}$$

$$\partial_{\bar{\imath}} \bar{F}_A = \mathcal{N}_{AB} \partial_{\bar{\imath}} \bar{X}^B \tag{2.16}$$

$$\begin{aligned} C_{ikp} &= \partial_i X^A \mathcal{D}_k \partial_p F_A - \partial_i F_A \mathcal{D}_k \partial_p X^A \\ &= \partial_i X^B \left(\partial_k \partial_p F_B - \partial_k \partial_p X^A \overline{\mathcal{N}}_{AB}\right) \end{aligned} \tag{2.17}$$

$$R_{i\bar{j}k\bar{l}} = -C_{ikp} \bar{C}_{\bar{j}\bar{l}\bar{p}} g^{p\bar{p}} \tag{2.18}$$

$$C_{ijk} = \partial_i X^A \partial_j X^B \partial_k X^C \partial_A \partial_B \partial_C F \tag{2.19}$$

$$C_{ABC} = \partial_A \partial_B \partial_C F \tag{2.20}$$

Symplectic transformations in the fermionic sector

$$\mathcal{L} = -\frac{i}{2}\overline{\mathcal{N}}_{AB}\mathcal{F}^{-A\mu\nu}\mathcal{F}^{-B}_{\mu\nu} - i\mathcal{F}^{-A\mu\nu}H^{-}_{A\mu\nu} + \text{c.c.} \ + \mathcal{L}_{4f} \tag{2.21}$$

$H^{-}_{A\mu\nu}$ are quadratic, and $\mathcal{L}_{4f}$ are the quartic terms in fermions.

$$G^{-}_{A\mu\nu} \equiv i\frac{\delta\mathcal{L}}{\delta\mathcal{F}^{-A\mu\nu}} = \overline{\mathcal{N}}_{AB}\mathcal{F}^{-B}_{\mu\nu} + H^{-}_{A\mu\nu} = G^{-}_{bA\mu\nu} + H^{-}_{A\mu\nu} \tag{2.22}$$

$$\begin{aligned}
\mathcal{L}_V &\equiv -\frac{i}{2}\overline{\mathcal{N}}\mathcal{F}^{-}_{\mu\nu}\mathcal{F}^{-\mu\nu} + \text{ c.c.} \\
&= -\frac{i}{2}G^{-}_{b\mu\nu}\mathcal{F}^{\mu\nu} + \text{ c.c.} \\
&= i\partial^{\mu}G^{-}_{b\mu\nu}A^{\nu} + \text{ c.c.} \\
&= -i\partial^{\mu}H^{-}_{\mu\nu}A^{\nu} + \text{ c.c.} \ - 2\partial^{\mu}\operatorname{Im}G^{-}_{\mu\nu}A^{\nu} \\
&= \frac{i}{2}H^{-}_{\mu\nu}\mathcal{F}^{-\mu\nu} + \text{ c.c.} \ - 2\partial^{\mu}\operatorname{Im}G^{-}_{\mu\nu}A^{\nu}
\end{aligned} \tag{2.23}$$

$$\mathcal{L}|_{\frac{\delta\mathcal{L}}{\delta\mathcal{A}}=0} = -\frac{i}{2}H^{-}_{\mu\nu}\mathcal{F}^{-\mu\nu} + \text{ c.c.} \ + \mathcal{L}_{4f} \equiv \mathcal{L}_{inv} \tag{2.24}$$

$$\mathcal{L} = -\frac{i}{2}\mathcal{F}^{-A\mu\nu}G^{-}_{A\mu\nu} + \text{c.c.} \ + \mathcal{L}_{inv} \tag{2.25}$$

$$H^{-}_{A\mu\nu} = \left(P_{Aa} - \overline{\mathcal{N}}_{AB}Q^{B}_{a}\right)\mathcal{T}^{-a}_{\mu\nu} \tag{2.26}$$

a denotes a new index, whose meaning depends on the model. $\mathcal{T}^{-a}_{\mu\nu}$ is a tensor not transforming under the symplectic group.

$$\begin{aligned}
\mathcal{L}_{inv} &= -\frac{i}{2}\mathcal{F}^{-A\mu\nu}\left(P_{Aa} - \overline{\mathcal{N}}_{AB}Q^{B}_{a}\right)\mathcal{T}^{-a}_{\mu\nu} + \text{ c.c.} \ + \mathcal{L}_{4f} \\
&= -\frac{i}{2}\left(\mathcal{F}^{-A\mu\nu}P_{Aa} - G^{-\mu\nu}_{bA}Q^{A}_{a}\right)\mathcal{T}^{-a}_{\mu\nu} + \text{ c.c.} \ + \mathcal{L}_{4f}
\end{aligned} \tag{2.27}$$

Invariance of $\mathcal{L}_{\text{inv}}$ is then guaranteed if $\left(Q^{A}, P_{A}\right)$ is a symplectic vector, and $\mathcal{L}_{4f}$ is constructed as the completion of G_b to G in the above formula (plus possible completely invariant terms). These completions are thus

$$\mathcal{L}_{4f} = \frac{i}{2}H^{-\mu\nu}_{A}Q^{A}_{a}\mathcal{T}^{-a}_{\mu\nu} + \text{ c.c.} \ + \text{ invariant terms} \tag{2.28}$$

Compact duality rotations

$$\delta\begin{pmatrix} F + iG \\ F - iG \end{pmatrix} = \begin{pmatrix} T & 0 \\ 0 & T^{*} \end{pmatrix}\begin{pmatrix} F + iG \\ F - iG \end{pmatrix},$$

$$FI - GH = \frac{1}{2}\left(F^{2} + G^{2}\right) = \frac{1}{2}(F - iG)(F + iG),$$

Special methods by Gaillard, Zumino

Non-linear realizations

$$\begin{aligned}
&D_\mu g = \partial_\mu g - gQ_\mu,\\
&D_\mu g \to (D_\mu g)\,[k(x)]^{-1},\\
&g^{-1}D_\mu g \to kg^{-1}D_\mu gk^{-1}.
\end{aligned}$$

$$\delta L = \operatorname{Tr}\delta g g^{-1}\partial_\mu\left(D_\mu g g^{-1}\right) = 0.$$

Now, $\delta g g^{-1}$ is an arbitrary element of the Lie algebra of $\mathcal{G}$; therefore, one has the equations of motion

$$\partial_\mu\left(D_\mu g g^{-1}\right) = 0,$$

or

$$\partial_\mu\left(gP_\mu g^{-1}\right) = 0.$$

(?? ничего тут пока не знаю.)

Non-compact duality transformations (?? ничего тут пока не знаю. но тут кстати много всего)

2.3 Applications and examples in a nutshell

2.3.1 Main illustrative examples

(напишу, подумать чуть нужно про них, пока не понял их как следует ещё)

Changing little coupling constant to a big one (???)

(??? have plans to find such examples)

Duality for gauge field and complex scalar

$$\mathcal{L} = -\frac{1}{4}(\operatorname{Im} Z)F_{\mu\nu}F^{\mu\nu} - \frac{1}{8}(\operatorname{Re} Z)\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}.$$

$$\partial_\mu\tilde{F}^{\mu v} = 0, \quad \partial_\mu\left[(\operatorname{Im} Z)F^{\mu v} + \mathrm{i}(\operatorname{Re} Z)\tilde{F}^{\mu v}\right] = 0.$$

$$G^{\mu v} \equiv \varepsilon^{\mu\nu\rho\sigma}\frac{\delta S}{\delta F^{\rho\sigma}} = -\mathrm{i}(\operatorname{Im} Z)\tilde{F}^{\mu v} + (\operatorname{Re} Z)F^{\mu\nu},$$

$$G^{\mu\nu-} = ZF^{\mu\nu-}, \quad G^{\mu\nu+} = \bar{Z}F^{\mu v+}.$$

$$\partial_\mu \operatorname{Im} F^{\mu v-} = 0, \quad \partial_\mu \operatorname{Im} G^{\mu\nu-} = 0.$$

$$\mathrm{SL}(2,\mathbb{R})\mathcal{S} \equiv \begin{pmatrix} d & c\\ b & a\end{pmatrix}, \quad ad - bc = 1$$

$$\begin{pmatrix} F'^{-}\\ G'^{-}\end{pmatrix} = \mathcal{S}\begin{pmatrix} F^{-}\\ G^{-}\end{pmatrix}.$$

$G'^{\mu\nu-} = Z'F'^{\mu\nu-}$,

$$Z' = \frac{aZ+b}{cZ+d}.$$

$$\mathcal{L}(F,Z) = -\frac{1}{2}\,\mathrm{Im}\left(ZF^{-}_{\mu\nu}F^{\mu\nu-}\right).$$

$\mathrm{SL}(2,\mathbb{R})$, $a=d=1$, $b=0\Rightarrow$

$$\mathcal{L}\left(F',Z'\right) = -\frac{1}{2}\,\mathrm{Im}\left(Z(1+cZ)F^{-}_{\mu\nu}F^{\mu v-}\right) \neq \mathcal{L}(F,Z).$$

$$\Theta^{\mu v} = (\mathrm{Im}\,Z)\left(F^{\mu\rho}F^{\nu}_{\rho} - \frac{1}{4}\eta^{\mu v}F_{\rho\sigma}F^{\rho\sigma}\right). \qquad \text{TEM}$$

$$\mathrm{Im}\,Z' = \frac{\mathrm{Im}\,Z}{(cZ+d)(c\bar{Z}+d)}.$$

$$\tilde{F}_{\mu\rho}\tilde{F}_{\nu}{}^{\rho} = -F_{\mu\rho}F_{\nu}{}^{\rho} + \frac{1}{2}\eta_{\mu\nu}F_{\rho\sigma}F^{\rho\sigma}.$$

$$F'_{\mu\rho}F'^{\rho}_{\nu} - \frac{1}{4}\eta_{\mu\nu}F'_{\rho\sigma}F'^{\rho\sigma} = |cZ+d|^2\left[F_{\mu\rho}F_{\nu}{}^{\rho} - \frac{1}{4}\eta_{\mu\nu}F_{\rho\sigma}F^{\rho\sigma}\right].$$

Electromagnetic duality for coupled Maxwell fields (!?!?!)

$$\begin{aligned}
\mathcal{L} &= -\frac{1}{4}\left(\mathrm{Re}\,f_{AB}\right)F^{A}_{\mu\nu}F^{\mu\nu B} + \frac{1}{4}\mathrm{i}\left(\mathrm{Im}\,f_{AB}\right)F^{A}_{\mu\nu}\tilde{F}^{\mu\nu B},\\
&= -\frac{1}{2}\,\mathrm{Re}\left(f_{AB}F^{-A}_{\mu\nu}F^{\mu\nu-B}\right)\\
&= -\frac{1}{4}\left(f_{AB}F^{-A}_{\mu\nu}F^{\mu\nu-B} + f^{*}_{AB}F^{+A}_{\mu\nu}F^{\mu\nu+B}\right),
\end{aligned}$$

$$\begin{aligned}
G^{\mu\nu}_{A} &= \varepsilon^{\mu\nu\rho\sigma}\frac{\delta S}{\delta F^{\rho\sigma A}} = -\left(\mathrm{Im}\,f_{AB}\right)F^{\mu\nu B} - \mathrm{i}\left(\mathrm{Re}\,f_{AB}\right)\tilde{F}^{\mu\nu B} \equiv G^{\mu\nu+}_{A} + G^{\mu\nu-}_{A},\\
G^{\mu\nu-}_{A} &= -2\mathrm{i}\frac{\delta S\left(F^{+},F^{-}\right)}{\delta F^{-A}_{\mu\nu}} = \mathrm{i}f_{AB}F^{\mu\nu-B},\\
G^{\mu\nu+}_{A} &= 2\mathrm{i}\frac{\delta S\left(F^{+},F^{-}\right)}{\delta F^{+A}_{\mu\nu}} = -\mathrm{i}f^{*}_{AB}F^{\mu\nu+B}.
\end{aligned}$$

Since the field equation for the action containing (4.67) is

$$0 = \frac{\delta S}{\delta A^{A}_{\nu}} = -2\partial_{\mu}\frac{\delta S}{\delta F^{A}_{\mu\nu}},$$

the Bianchi identity and the equation of motion can be expressed in the concise form

$$\begin{aligned}
\partial^{\mu}\,\mathrm{Im}\,F^{A-}_{\mu\nu} &= 0 \quad \text{Bianchi identities,}\\
\partial_{\mu}\,\mathrm{Im}\,G^{\mu\nu-}_{A} &= 0 \quad \text{equations of motion.}
\end{aligned}$$

(The same equations hold for $\mathrm{Im}\,F^{A+}$ and $\mathrm{Im}\,G_{A}{}^{+}$.)

$$\begin{pmatrix} F'^{-} \\ G'^{-} \end{pmatrix} = \mathcal{S}\begin{pmatrix} F^{-} \\ G^{-} \end{pmatrix} \equiv \begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} F^{-} \\ G^{-} \end{pmatrix},$$

with real $m \times m$ submatrices A, B, C, D. Owing to the reality of these matrices, the same relations hold for the self-dual tensors $F^+$and G^+. In Sec. 4.2.3, these matrices were just numbers:

$$A = d, \quad B = c, \quad C = b, \quad D = a.$$

We require that the transformed field tensors F'^A and G'_A are also related by the definitions (4.68), with appropriately transformed f_{AB}. We work out this requirement in the following steps:

$$G'^- = (C + \mathrm{i}Df)F^- = (C + \mathrm{i}Df)(A + \mathrm{i}Bf)^{-1}F'^-,$$

such that we conclude that

$$\mathrm{i}f' = (C + \mathrm{i}Df)(A + \mathrm{i}Bf)^{-1}.$$

$$A^TC - C^TA = 0, \quad B^TD - D^TB = 0, \quad A^TD - C^TB = \mathbb{1}.$$

These relations among A, B, C, D are the defining conditions of a matrix of the symplectic group in dimension $2m$ so we reach the conclusion that

$$\mathcal{S} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Sp}(2m, \mathbb{R}).$$

The conditions (4.75) may be summarized as

$$\mathcal{S}^T\Omega\mathcal{S} = \Omega \quad \text{where} \quad \Omega = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}.$$

(!!???!??!?!?!?!??!??!??!?!?)

The duality transformations in four dimensions are transformations in the symplectic group Sp(2m, R).

(!?!!?!??!?!!??!?!?!?!??)

$$\mathcal{L} = -\frac{1}{2}\,\mathrm{Re}\left(f_{AB}F^{-A}_{\mu\nu}F^{\mu\nu-B}\right) = -\frac{1}{2}\,\mathrm{Im}\left(F^{-A}_{\mu\nu}G^{\mu\nu-}_A\right),$$

we obtain

$$\mathrm{Im}\,F'^-G^- = \mathrm{Im}\left(F^-G^-\right) + \mathrm{Im}\left[2F^-\left(C^TB\right)G^- + F^-\left(C^TA\right)F^- + G^-\left(D^TB\right)G^-\right].$$

If $C \neq 0, B = 0$ the Lagrangian is invariant up to a 4-divergence, since $\mathrm{Im}\,F^-F^- = -\frac{1}{4}\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}$ and the matrices A and C are real constants. For $B \neq 0$ neither the Lagrangian nor the action is invariant.

2.3.2 Fermions in $N = 2$ rigid Yang-Mills theory

The coordinate independent description of fermions is given by $SU(2)$ doublets $\left(\lambda^{iI}, \lambda^{\bar{\imath}}_I\right)$ where upper and lower $SU(2)$ indices I mean positive and negative chiralities respectively. As such the spinors are symplectic invariant and contravariant world vector fields.

$$\left(X^A, \partial_i X^A \lambda^{iI}_\alpha, \mathcal{F}^{-A}_{\alpha\beta}\right) \tag{2.29}$$

$\alpha, \beta \in SL(2, \mathbb{C})$.

$$\star\mathcal{F}^{-A}_{\alpha\beta} \text{ is } \sigma^{\mu\nu}_{\alpha\beta}\mathcal{F}^{-A}_{\mu\nu}.$$

$$Q^A_{\bar\imath} = \partial_{\bar\imath}\bar X^A; \quad P_{A\bar\imath} = \partial_{\bar\imath}\bar F_A$$
$$\mathcal{T}^{\bar\imath}_{\alpha\beta} = k g^{\bar\imath j} C_{jkp}\lambda^{kI}_\alpha \lambda^{pJ}_\beta \epsilon_{IJ} \tag{2.30}$$

$$H^{\alpha\beta}_{-A} = k\partial_{\bar\imath}\bar X^B \left(\mathcal{N}_{BA} - \overline{\mathcal{N}}_{BA}\right) g^{\bar\imath j} C_{jkp}\lambda^{\alpha kI}\lambda^{\beta pJ}\epsilon_{IJ} \tag{2.31}$$

$$\mathcal{L}_{\text{Pauli}} = -i(\mathcal{N} - \overline{\mathcal{N}})_{AB}\partial_{\bar\imath}\bar X^A \mathcal{T}^{\bar\imath}_{\alpha\beta}\mathcal{F}^{B\alpha\beta} + \text{ c.c.}$$
$$\mathcal{L}_{4\text{f}} = \frac{i}{2}\partial_{\bar\imath}\bar X^A \partial_{\bar\jmath}\bar X^B \left(\overline{\mathcal{N}}_{AB} - \mathcal{N}_{AB}\right)\mathcal{T}^{\bar\imath}_{\alpha\beta}\mathcal{T}^{\bar\jmath\alpha\beta} + c.c. + \text{ invariant terms} \tag{2.32}$$

$$\mathcal{L}_{\text{Pauli}} = -k\partial_A\partial_B\partial_C F \left(\chi^A_\alpha\lambda^B_\beta - \lambda^A_\alpha\chi^B_\beta\right)\mathcal{F}^{-C\alpha\beta} + \text{ c.c.} \tag{2.33}$$

$$\lambda^{iI}_\alpha\lambda^{kJ}_\beta\epsilon^{\alpha\beta}\bar\lambda^{\bar\jmath}_{\dot\alpha I}\bar\lambda^{\bar l}_{\dot\beta J}\epsilon^{\dot\alpha\dot\beta}R_{i\bar\jmath k\bar l} \tag{2.34}$$

$$\mathcal{D}_i C_{jlm}\lambda^{iI}_\alpha\lambda^{jK}_\beta\epsilon^{\alpha\beta}\lambda^{lJ}_\gamma\lambda^{mL}_\delta\epsilon^{\gamma\delta}\epsilon_{IJ}\epsilon_{KL} \tag{2.35}$$

2.3.3 Example: duality transformations in $N = 1$ locally SUSY YM theories

$$\mathcal{V} = \left(\mathcal{F}^{-A}_{\mu\nu}, G^{-\mu\nu}_A = i\frac{\partial\mathcal{L}}{\partial\mathcal{F}^{-A}_{\mu\nu}}\right)$$
$$\mathcal{U}_\alpha = \left(\lambda^A_\alpha, f_{AB}(z)\lambda^B_\alpha\right) \tag{3.50}$$

$$\mathcal{V}\to\mathcal{S}\mathcal{V}, \quad \mathcal{U}\to\mathcal{S}\mathcal{U} \quad , \quad f\to(C+Df)(A+Bf)^{-1} \quad , \quad \mathcal{S}\in Sp(2,r;\mathbf{R}) \tag{3.51}$$

$$W^A_\alpha = T\left(\mathcal{D}_\alpha V^A\right) \tag{3.52}$$

$$4\,\mathrm{Im}\, W^A_\alpha\mathcal{D}_\beta U_A\epsilon^{\alpha\beta}\big|_D + i f_{AB}(z)W^A_\alpha W^B_\beta\epsilon^{\alpha\beta}\big|_F \tag{3.53}$$

$$\mathcal{D}^\alpha W^A_\alpha = \overline{\mathcal{D}}_{\dot\alpha}\bar W^{\dot\alpha A} \tag{3.54}$$

$$W^A_\alpha = T\left(\mathcal{D}_\alpha V^A\right) \tag{3.55}$$

$$W^A_\alpha = \left(f^{-1}\right)^{AB} W^{(D)}_{\alpha B} \tag{3.56}$$

$$\mathcal{L}^D = -\,i\left(f^{-1}\right)^{AB} W^{(D)}_{\alpha A}W^{(D)}_{\beta B}\epsilon^{\alpha\beta}\Big|_F \tag{3.57}$$

$$\widetilde{f}_{AB}(\widetilde{z}) = f_{AB}(\widetilde{z}) \tag{3.58}$$

$$\|\widetilde{W}(\widetilde{z})\|^2 = \|W(\widetilde{z})\|^2 \tag{3.59}$$

$\|W(z)\|^2 = |W(z)|^2 e^K \equiv e^G.$

$$
\begin{aligned}
e^{-1}\mathcal{L}_{\text{kin}}(\lambda,\bar{\lambda}) &= i\boldsymbol{\mathcal{U}}_\alpha \Omega\,(\sigma^\mu)^{\alpha\dot{\alpha}}\,\mathcal{D}_\mu \overline{\boldsymbol{\mathcal{U}}}_{\dot{\alpha}} \\
e^{-1}\mathcal{L}_{\text{Pauli}}(\psi,\lambda) &= \operatorname{Im}\left(\overline{\boldsymbol{\mathcal{U}}}_{\dot{\alpha}}\Omega\,(\sigma^\mu)^{\dot{\alpha}\beta}\,\boldsymbol{\mathcal{V}}_{b\beta\gamma}\psi^\gamma_\mu\right) \\
e^{-1}\mathcal{L}_{\text{Pauli}}(\chi,\lambda) &= \operatorname{Im}\left(\partial_i f_{AB}\lambda^A_\alpha \bar{\chi}^i_\beta \boldsymbol{\mathcal{F}}^{-B\alpha\beta}\right)
\end{aligned}
\tag{3.60}
$$

$$
\mathcal{T}^{\dot{\alpha}}_{\beta\gamma} = -\frac{1}{2}\,(\sigma^\mu)^{\dot{\alpha}}{}_\beta \psi_{\mu\gamma}.
$$

$$
H_{A\alpha\beta} = \frac{1}{2}\partial_i f_{AB}\lambda^B_\alpha \bar{\chi}^i_\beta
$$

$$
\begin{aligned}
Q^A_{i\alpha} &\equiv (\operatorname{Im} f)^{-1AB}\partial_i f_{BC}\lambda^C_\alpha; \quad P_{Ai\alpha} \equiv \bar{f}_{AB} Q^B_{i\alpha} \\
\mathcal{T}^{i\alpha}_{\beta\gamma} &= \tfrac{i}{4}\delta^\alpha_{(\beta}\chi^i_{\gamma)}
\end{aligned}
$$

$$
\begin{aligned}
\tilde{f} &= (C+Df)(A+Bf)^{-1} = \left(A^T + fB^T\right)^{-1}\left(C^T + fD^T\right) \\
\partial_i \tilde{f} &= D\partial_i f(A+Bf)^{-1} - (C+Df)(A+Bf)^{-1}B\partial_i f(A+Bf)^{-1} \\
&= \left(A^T + fB^T\right)^{-1}\partial_i f(A+Bf)^{-1} \\
\operatorname{Im}\tilde{f} &= \left(A^T + fB^T\right)^{-1}(\operatorname{Im} f)(A+B\bar{f})^{-1} \\
\widetilde{\lambda} &= (A+Bf)\lambda
\end{aligned}
\tag{3.61}
$$

2.3.4 Duality in electrodynamics in brief

(all about it here.)

2.3.5 On duality examples in Dirac, 't Hooft–Polyakov, BPS- monopoles

The Dirac Monopole

111 And in the beginning there was Maxwell12
112 The Dirac quantisation condition 14

Dyons and the Zwanziger–Schwinger quantisation condition (!?) (читал, но нужно разобарться ещё, отдельная тоже тема)

The 't Hooft–Polyakov Monopole

121 The bosonic part of the Georgi–Glashow model 18
122 Finite-energy solutions: the 't Hooft–Polyakov Ansatz 20
123 The topological origin of the magnetic charge 24

BPS-monopoles

131 Estimating the mass of a monopole: the Bogomol'nyi bound 27
132 Saturating the bound: the BPS-monopole 28

14 Duality conjectures 30

141 The Montonen–Olive conjecture 30
142 The Witten effect 31
143 SL(2, Z) duality

2.3.6 On duality examples in Poincare group

(впишу потом основные формулы)

2.3.7 On Duality examples in black holes (??)

(пока не до этого, смотрел немного, но это отдельное направление вообще)

2.3.8 Overview of other applications

2.4 Background for duality in a nutshell

2.4.1 On variation,....

Properties of variation

2.4.2 On supersymmetry and sypergravity basics

The super-Poincaré algebra in four dimensions

2.1.1 Some notational remarks about spinors 40

2.1.2 The Coleman–Mandula and Haag–ÃLopusza´nski–Sohnius theorems 42

Reminder of idea and main formulas of supergravity

(без этого мало кто не поймет то, что выше, при этом на 1/2-2/3 страницы всего формул и слов планирую.)

2.2 Unitary representations of the supersymmetry algebra 44

2.2.1 Wigner's method and the little group 44

2.2.2 Massless representations 45

2.2.3 Massive representations 47

No central charges

Adding central charges

2.5 Very special duality methods in a nutshell

(пока это не пишу, в теории буду писать чуть что. тут то, что крайне редко нужно для редких методов - часть выше есть.)

2.5.1 On Effective Action for N=2 Supersymmetric Yang–Mills

(look at Figueroa-O'Farrill, his last part, I don't need it now)

2.5.2 On Monopoles for Arbitrary Gauge Groups by Figueroa-O'Farrill

(look at Figueroa-O'Farrill, his last part, I don't need it now)

Part II
Main Theory

(I'll add especially useful theory form part below to here later)

3 Duality in united formalism

(My theory about all the theory below in one consistent formalism. No need to write it now and I need several months of practice to do it)

4 Most used theory of duality

4.1 Duality in simple coupled theories by Freedman, Proeyen

(!!! тут то же, что Гиллард Зумино, но лучшим как сказал Проен методом, так что переделаю!)

4.1.1 Duality for one free electromagnetic field

Duality operates as an interesting symmetry of field theories containing one or more abelian gauge fields which may interact with other fields, principally scalars. In this section we discuss the simplest case, namely a single free gauge field. First note that, after contraction with the ε-tensor, the Bianchi identity (4.11) can be expressed as $\partial_\mu \tilde{F}^{\mu\nu} = 0$.

6 The definition (4.35) is valid in Minkowski space, but must be modified in curved spacetimes as we will discuss in Ch. 7.

So we can temporarily ignore the vector potential and regard $F_{\mu\nu}$ as the basic field variable which must satisfy both the Maxwell and Bianchi equations:

$$\partial_\mu F^{\mu\nu} = 0, \quad \partial_\mu \tilde{F}^{\mu\nu} = 0.$$

We can now consider the change of variables (the i is included to make the transformation real):

$$F^{\mu\nu} \to F'^{\mu\nu} = \mathrm{i}\tilde{F}^{\mu\nu}$$

Since $F'^{\mu\nu}$ also obeys both equations of (4.41) we have defined a symmetry of the free electromagnetic field.

Exercise 4.8 Show that the symmetry (4.42) exchanges the electric and magnetic fields: $E_i \to E'_i = -B_i$ and $B_i \to B'_i = E_i$.

It is not possible to extend the symmetry to the vector potentials $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and $F'_{\mu\nu} = \partial_\mu A'_\nu - \partial_\nu A'_\mu$ because A_μ and A'_μ are not related by any local transformation.

Here are some basic exercises involving the duality transform of the field strength tensor $F_{\mu\nu}$.

Exercise 4.9 Show that the self-dual combinations $F^\pm_{\mu\nu}$contain only photons of one polarization in their plane wave expansions:

$$F^\pm_{\mu\nu} = 2\mathrm{i} \int \frac{\mathrm{d}^3 k}{(2\pi)^3 2k^0} \left[\mathrm{e}^{\mathrm{i}k\cdot x} k_{[\mu}\epsilon_{\nu]}(\vec{k}, \pm) a(\vec{k}, \pm) - \mathrm{e}^{-\mathrm{i}k\cdot x} k_{[\mu}\epsilon^*_{\nu]}(\vec{k}, \mp) a^*(\vec{k}, \mp) \right]$$

To perform this exercise, check first that with the polarization vectors given in Sec. 4.1.2, one has

$$-\frac{1}{2}\mathrm{i}\varepsilon^{\mu\nu\rho\sigma}k_\rho\epsilon_\sigma(\vec{k},\pm) = \pm k^{[\mu}\epsilon^{\nu]}(\vec{k},\pm).$$

Exercise 4.10 Show that the quantity $F_{\mu\nu}\tilde{F}^{\mu\nu}$ is a total derivative, i.e.

$$F_{\mu\nu}\tilde{F}^{\mu\nu} = -\mathrm{i}\partial_\mu\left(\varepsilon^{\mu\nu\rho\sigma}A_\nu F_{\rho\sigma}\right)$$

Show, using (1.45), that under a Lorentz transformation

$$\left(F_{\mu\nu}\tilde{F}^{\mu\nu}\right)(x) \to \det\Lambda^{-1}\left(F_{\mu\nu}\tilde{F}^{\mu\nu}\right)(\Lambda x)$$

Thus $F_{\mu\nu}\tilde{F}^{\mu\nu}$ transforms as a scalar under proper Lorentz transformations but changes sign under space or time reflections. Use the Schouten identity (3.11) to prove that

$$F_{\mu\rho}\tilde{F}^{\rho}_{\nu} = \frac{1}{4}\eta_{\mu\nu}F_{\rho\sigma}\tilde{F}^{\rho\sigma}$$

4.1.2 Basic Properties of Dual Tensors

For a second rank antisymmetric tensors $H_{\mu\nu}$ in four-dimensional Minkowski spacetime we define a dual tensor as

$$\tilde{H}^{\mu\nu} \equiv -\frac{1}{2}\mathrm{i}\varepsilon^{\mu\nu\rho\sigma}H_{\rho\sigma}$$

In this conventions the dual tensor is imaginary. The indices of $\tilde{H}$ can be raised and lowered with the Minkowsk metric $\eta_{\mu\nu}$. We also define the self-dual and anti-self-dual tensors as linear combinations

$$H^{\pm}_{\mu\nu} = \frac{1}{2}\left(H_{\mu\nu} \pm \tilde{H}_{\mu\nu}\right), \quad H^{\pm}_{\mu\nu} = \left(H^{\mp}_{\mu\nu}\right)^*$$

They obey

$$-\frac{1}{2}\mathrm{i}\varepsilon^{\mu\nu\rho\sigma}\tilde{H}_{\rho\sigma} = H^{\mu\nu}$$

(???? proof!!???) The validity of this property is the reason for the i in the definition (4.35).

Also

$$-\frac{1}{2}\mathrm{i}\varepsilon_{\mu\nu}{}^{\rho\sigma}H^{\pm}_{\rho\sigma} = \pm H^{\pm}_{\mu\nu}$$

(???? proof!!???)

Let $G_{\mu\nu}$ be another antisymmetric tensor with $G^{\pm}_{\mu\nu}$defined as in (4.36). Prove the following relations (where $(\mu\nu)$ means symmetrization between the indices):

$$G^{+\mu\nu}H^{-}_{\mu\nu} = 0, \quad G^{\pm\rho(\mu}H^{\pm\nu)}_{\rho} = -\frac{1}{4}\eta^{\mu\nu}G^{\pm\rho\sigma}H^{\pm}_{\rho\sigma}, \quad G^{+}_{\rho[\mu}H^{-}_{\nu]}\rho$$

Hint: you could first prove

$$\tilde{G}^{\rho\mu}\tilde{H}^{\nu}_{\rho} = -\frac{1}{2}\eta^{\mu\nu}G^{\rho\sigma}H_{\rho\sigma} - G^{\rho v}H^{\mu}_{\rho}.$$

The duality operation can also be applied to matrices of the Clifford algebra. Define the quantity $L_{\mu\nu} = \gamma_{\mu\nu}P_L$. Show that this is anti-self-dual. Hint: check first that $\gamma_{\mu\nu}\gamma_* = \frac{1}{2}\mathrm{i}\varepsilon_{\mu\nu\rho\sigma}\gamma^{\rho\sigma}$.

4.1.3 Duality for gauge field and complex scalar

The simplest case of electromagnetic duality in an interacting field theory occurs with one abelian gauge field $A_\mu(x)$ and a complex scalar field $Z(x)$. The electromagnetic part of the Lagrangian is

$$\mathcal{L} = -\frac{1}{4}(\operatorname{Im} Z)F_{\mu\nu}F^{\mu\nu} - \frac{1}{8}(\operatorname{Re} Z)\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}.$$

Actions in which the gauge field kinetic term is multiplied by a function of complex scalar fields are quite common in supersymmetry and supergravity. We now define an extension of the duality transformation (4.42) which gives a non-abelian global SL(2, $\mathbb{R}$) symmetry of the gauge field equations of this theory. In Sec. 7.12.2 we will discuss a generalized scalar kinetic term that is invariant under SL$(2,\mathbb{R})$. The field $Z(x)$ carries dynamics, and the equations of motion of the combined vector and scalar theory are also invariant.

The gauge Bianchi identity and equation of motion of our theory are

$$\partial_\mu \tilde{F}^{\mu\nu} = 0, \quad \partial_\mu \left[(\operatorname{Im} Z)F^{\mu\nu} + \mathrm{i}(\operatorname{Re} Z)\tilde{F}^{\mu\nu}\right] = 0.$$

It is convenient to define the real tensor

$$G^{\mu\nu} \equiv \varepsilon^{\mu\nu\rho\sigma}\frac{\delta S}{\delta F^{\rho\sigma}} = -\mathrm{i}(\operatorname{Im} Z)\tilde{F}^{\mu\nu} + (\operatorname{Re} Z)F^{\mu\nu}$$

and to consider the self-dual combinations $F^{\mu\nu\pm}$ and $G^{\mu\nu\pm}$. Note that these are related by

$$G^{\mu\nu-} = ZF^{\mu\nu-}, \quad G^{\mu\nu+} = \bar{Z}F^{\mu\nu+}.$$

The information in (4.49) can then be reexpressed as

$$\partial_\mu \operatorname{Im} F^{\mu\nu-} = 0, \quad \partial_\mu \operatorname{Im} G^{\mu\nu-} = 0$$

We define a matrix of the group SL$(2,\mathbb{R})$ by

$$\mathcal{S} \equiv \begin{pmatrix} d & c \\ b & a \end{pmatrix}, \quad ad - bc = 1$$

The group SL$(2,\mathbb{R})$ acts on the tensors $F^-$and G^-as follows:

$$\begin{pmatrix} F'^- \\ G'^- \end{pmatrix} = \mathcal{S}\begin{pmatrix} F^- \\ G^- \end{pmatrix}$$

Since $\mathcal{S}$ is real, the conjugate tensors $F^+$and G^+also transform in the same way.

Exercise 4.11 Assume that $\operatorname{Im} F^-$and $\operatorname{Im} G^-$satisfy (4.52), and show that $\operatorname{Im} F'^-$ and $\operatorname{Im} G'^-$ also obey the same equations. Show that G'^- and a transformed scalar Z' satisfy $G'^{\mu\nu-} = Z'F'^{\mu\nu-}$, if Z' is defined as the following nonlinear transform of Z:

$$Z' = \frac{aZ+b}{cZ+d}$$

The two equations (4.54) and (4.55) specify the SL$(2,\mathbb{R})$ duality transformation on the field strength and complex scalar of our system. The exercise shows that the Bianchi identity and generalized Maxwell equations are duality invariant. In general the duality transform is not a symmetry of the Lagrangian or the action integral. The following exercise illustrates this.

Exercise 4.12 Show that the Lagrangian (4.48) can be rewritten as

$$\mathcal{L}(F,Z) = -\frac{1}{2}\operatorname{Im}\left(ZF^-_{\mu\nu}F^{\mu\nu-}\right).$$

Consider the SL$(2,\mathbb{R})$ transformation with parameters $a = d = 1$ and $b = 0$. Show that

$$\mathcal{L}(F', Z') = -\frac{1}{2}\operatorname{Im}\left(Z(1+cZ)F^-_{\mu\nu}F^{\mu\nu-}\right) \neq \mathcal{L}(F, Z)$$

The symmetric gauge invariant stress tensor of this theory is

$$\Theta^{\mu\nu} = (\operatorname{Im} Z)\left(F^{\mu\rho}F^{\nu}_{\rho} - \frac{1}{4}\eta^{\mu\nu}F_{\rho\sigma}F^{\rho\sigma}\right)$$

As we will see in Ch. 8, when the theory is coupled to gravity, it is this stress tensor that is the source of the gravitational field; see (8.4). It is then important that $\operatorname{Im} Z$ is positive, which restricts the domain of Z to the upper half-plane. It is also important that the stress tensor is invariant under the duality transformations (4.54) and (4.55). This is the reason for the duality symmetry of many black hole solutions of supergravity,

Exercise 4.13 Prove that the energy-momentum tensor (4.58) is invariant under duality. Here are some helpful relations which you will need:

$$\operatorname{Im} Z' = \frac{\operatorname{Im} Z}{(cZ+d)(c\bar{Z}+d)}$$

Further you need again (4.47) and a similar identity (proven by contracting ε-tensors)

$$\tilde{F}_{\mu\rho}\tilde{F}^{\rho}_{\nu} = -F_{\mu\rho}F_{\nu}{}^{\rho} + \frac{1}{2}\eta_{\mu\nu}F_{\rho\sigma}F^{\rho\sigma}.$$

This leads to

$$F'_{\mu\rho}F'^{\rho}_{\nu} - \frac{1}{4}\eta_{\mu\nu}F'_{\rho\sigma}F'^{\rho\sigma} = |cZ+d|^2\left[F_{\mu\rho}F_{\nu}{}^{\rho} - \frac{1}{4}\eta_{\mu\nu}F_{\rho\sigma}F^{\rho\sigma}\right].$$

When the SL$(2,\mathbb{R})$ duality transformation appears in supergravity, there is also a scalar kinetic term in the Lagrangian which is invariant under the symmetry, specifically under the transformation (4.55). The prototype Lagrangian with this symmetry is the nonlinear σ-model whose target space is the Poincaré plane. This model and its SL$(2,\mathbb{R})$ symmetry group will be discussed in Sec. 7.12; see (7.151) and (7.152). The Poincaré plane is the upper half-plane $\operatorname{Im} Z > 0$. The relation (4.59) shows that duality transformations map the upper half-plane into itself. The positive sign is preserved by SL$(2,\mathbb{R})$ transformations and the energy density obtained from the stress tensor Θ^{00} above will be positive!

Exercise 4.14 The free Maxwell theory is the special case of (4.48) with fixed $Z = \mathrm{i}$. Suppose that the gauge field is coupled to a conserved current as in (4.14). Check that the electric charge can be expressed in terms of F or G by

$$q \equiv \int \mathrm{d}^3\vec{x}J^0 = \int \mathrm{d}^3\vec{x}\partial_i F^{0i} = -\frac{1}{2}\int \mathrm{d}^3\vec{x}\varepsilon^{ijk}\partial_i G_{jk}$$

A magnetic charge can be introduced in Maxwell theory as the divergence of $\vec{B}$ (recall $E^i = F^{0i}$ and $B^i = \frac{1}{2}\varepsilon^{ijk}F_{jk}$). This leads to a definition [7]

$$p \equiv -\frac{1}{2}\int \mathrm{d}^3\vec{x}\varepsilon^{ijk}\partial_i F_{jk}$$

Show that $\begin{pmatrix} p \\ q \end{pmatrix}$ is a vector that transforms under SL$(2,\mathbb{R})$ in the same way as the tensors $F^-$and G^-in (4.54).

In many applications of electromagnetic duality, magnetic and electric charges appear as sources for the Bianchi 'identity' and generalized Maxwell equations of (4.49). As exemplified in

Ex. 4.14 this leads to an $\mathrm{SL}(2,\mathbb{R})$ vector of charges. Particles that carry both electric and magnetic charge are called dyons. In quantum mechanics, dyon charges must obey the Schwinger-Zwanziger quantization condition. If a theory contains two dyons with charges (p_1, q_1) and (p_2, q_2), these charges must satisfy $p_1 q_2 - p_2 q_1 = 2\pi n$, where n is an integer. [8] This condition is invariant under $\mathrm{SL}(2,\mathbb{R})$ transformations of the charges. However, one can show [25] that there is a lowest non-zero value of the electric charge and that all allowed charges are restricted to an infinite discrete set of points called the charge lattice. The allowed SL(2) transformations must take one lattice point to another, and this restricts the group parameters in (4.53) to be integers. This restriction defines the subgroup $\mathrm{SL}(2,\mathbb{Z})$, often called the modular group. [9] One can show that this subgroup is generated by the following choices of $\mathcal{S}$:

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$
$$Z' = Z + 1, \quad Z' = -\frac{1}{Z}.$$

This means that one can express any element of $\mathrm{SL}(2,\mathbb{Z})$ as the product of (finitely many) factors of the two generators above and their inverses.

Exercise 4.15 In (4.48), the kinetic terms of the electromagnetic fields are determined by a variable Z that was treated as a scalar field. Z can also be replaced by a coupling constant, and typically one takes Z to be the imaginary number [10]i/g^2, where g is a coupling constant. Observe that the first transformation of (4.64) does not preserve the restriction that Z is imaginary. However, the second one does. Prove that this transformation is of the type (4.42), interchanging the electric and magnetic fields. It transforms g to its inverse, and thus relates the strong and weak coupling descriptions of the theory. In

7 In order to obtain a symplectic vector (p, q) and not $(-p, q)$, we changed the sign of the magnetic charge with respect to some classical works. This implies that we have $\vec{\nabla} \cdot \vec{B} = -j^0_{\mathrm{m}}$, where j^0_{m} is the magnetic charge density.

8 For the case $(p_1, q_1) = (p, 0)$ and $(p_2, q_2) = (0, q)$, this reduces to condition $pq = 2\pi n$ found by Dirac in 1933.

9 The modular group generated by the matrices (4.64) is in fact PSL(2, $\mathbb{Z}$). In PSL(2, $\mathbb{Z}$), the elements M and $-M$ of $\mathrm{SL}(2,\mathbb{Z})$ are identified. Both these elements give in fact the same transformation $Z'(Z)$.

10 One often adds an extra term that is a real so-called θ-parameter, but we will omit this here.

Secs. 4.1 and 4.2.2 we considered $Z = \mathrm{i}g = \mathrm{i}$. Check that general duality transformations in this case are of the form

$$F'^{-}_{\mu\nu} = (d + \mathrm{i}c) F^{-}_{\mu\nu}, \qquad \text{i.e.} \qquad F'_{\mu\nu} = dF_{\mu\nu} - \mathrm{i}c\tilde{F}_{\mu\nu}.$$

4.1.4 Electromagnetic duality for coupled Maxwell fields

In this section we explore how the duality symmetry is extended to systems containing a set of abelian gauge fields $A^A_\mu(x)$, indexed by $A = 1, 2, \ldots, m$ together with scalar fields ϕ^i. Scalars enter the theory through complex functions $f_{AB}(\phi) = f_{BA}(\phi)$. We consider the action

$$S = \int \mathrm{d}^4 x \mathcal{L}, \quad \mathcal{L} = -\frac{1}{4} \left(\mathrm{Re}\, f_{AB}\right) F^A_{\mu\nu} F^{\mu\nu B} + \frac{1}{4}\mathrm{i} \left(\mathrm{Im}\, f_{AB}\right) F^A_{\mu\nu} \tilde{F}^{\mu\nu B}$$

which is real since $\tilde{F}^{\mu\nu}$ is pure imaginary, as defined in (4.35). The first term is a generalized kinetic Lagrangian for the gauge fields, so we usually require that $\mathrm{Re}\, f_{AB}$ is a positive definite matrix. This ensures that gauge field kinetic energies are positive. Although $F_{\mu\nu}\tilde{F}^{\mu\nu}$ is a total derivative, the second term does contribute to the equations of motion when $\mathrm{Im}\, f_{AB}$ is a function of the scalars ϕ^i. Our discussion will not involve the scalars directly. However, as in Sec. 4.2.3, additional terms to specify the scalar dynamics will appear when theories of this type are encountered in extended $D = 4$ supergravity. The treatment that follows is modeled on Sec. 4.2.3 (where f_{AB} was taken to be $-\mathrm{i}Z$).

Using the self-dual tensors of (4.36), we then rewrite the Lagrangian (4.66) as

$$\begin{aligned} \mathcal{L}\left(F^{+}, F^{-}\right) &= -\frac{1}{2}\,\mathrm{Re}\left(f_{AB} F_{\mu\nu}^{-A} F^{\mu\nu-B}\right) \\ &= -\frac{1}{4}\left(f_{AB} F_{\mu\nu}^{-A} F^{\mu\nu-B} + f_{AB}^{*} F_{\mu\nu}^{+A} F^{\mu\nu+B}\right), \end{aligned}$$

and define the new tensors

$$\begin{aligned} G_A^{\mu\nu} &= \varepsilon^{\mu\nu\rho\sigma} \frac{\delta S}{\delta F^{\rho\sigma A}} = -\left(\mathrm{Im}\, f_{AB}\right) F^{\mu\nu B} - \mathrm{i}\left(\mathrm{Re}\, f_{AB}\right) \tilde{F}^{\mu\nu B} = G_A^{\mu\nu+} + G_A^{\mu\nu-}, \\ G_A^{\mu\nu-} &= -2\mathrm{i} \frac{\delta S\left(F^{+}, F^{-}\right)}{\delta F_{\mu\nu}^{-A}} = \mathrm{i} f_{AB} F^{\mu\nu-B}, \\ G_A^{\mu\nu+} &= 2\mathrm{i} \frac{\delta S\left(F^{+}, F^{-}\right)}{\delta F_{\mu\nu}^{+A}} = -\mathrm{i} f_{AB}^{*} F^{\mu\nu+B}. \end{aligned}$$

Since the field equation for the action containing (4.67) is

$$0 = \frac{\delta S}{\delta A_v^A} = -2\partial_\mu \frac{\delta S}{\delta F_{\mu\nu}^A}$$

the Bianchi identity and the equation of motion can be expressed in the concise form

$$\begin{aligned} \partial^\mu\, \mathrm{Im}\, F_{\mu\nu}^{A-} &= 0 \qquad\qquad \text{Bianchi identities} \\ \partial_\mu\, \mathrm{Im}\, G_A^{\mu\nu-} &= 0 \qquad \text{equations of motion.} \end{aligned}$$

(The same equations hold for $\mathrm{Im}\, F^{A+}$ and $\mathrm{Im}\, G_A{}^{+}$.)

Duality transformations are linear transformations of the $2m$ tensors $F^{A\mu\nu}$ and $G_A^{\mu\nu}$ (accompanied by transformations of the f_{AB}) which mix Bianchi identities and equations of motion, but preserve the structure that led to (4.70). Since the equations (4.70) are real, we can mix them by a real $2m \times 2m$ matrix. We extend these transformations to the (anti-)self-dual tensors, and consider

$$\begin{pmatrix} F'^{-} \\ G'^{-} \end{pmatrix} = \mathcal{S} \begin{pmatrix} F^{-} \\ G^{-} \end{pmatrix} \equiv \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} F^{-} \\ G^{-} \end{pmatrix}$$

with real $m \times m$ submatrices A, B, C, D. Owing to the reality of these matrices, the same relations hold for the self-dual tensors $F^{+}$and G^{+}. In Sec. 4.2.3, these matrices were just numbers:

$$A = d, \quad B = c, \quad C = b, \quad D = a.$$

We require that the transformed field tensors F'^{A} and G'_A are also related by the definitions (4.68), with appropriately transformed f_{AB}. We work out this requirement in the following steps:

$$G'^{-} = (C + \mathrm{i}Df)F^{-} = (C + \mathrm{i}Df)(A + \mathrm{i}Bf)^{-1}F^{-},$$

such that we conclude that

$$\mathrm{i}f' = (C + \mathrm{i}Df)(A + \mathrm{i}Bf)^{-1}$$

The last equation gives the symmetry transformation relating f'_{AB} to f_{AB}. If $G'^{-}_{\mu\nu}$ is to be the variational derivative of a transformed action, as (4.68) requires, then the matrix f' must be symmetric. For a generic [11] symmetric f, this requires that the matrices A, B, C, D satisfy

$$A^T C - C^T A = 0, \quad B^T D - D^T B = 0, \quad A^T D - C^T B = \mathbb{1}.$$

These relations among A, B, C, D are the defining conditions of a matrix of the symplectic group in dimension $2m$ so we reach the conclusion that

$$\mathcal{S} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Sp}(2m, \mathbb{R})$$

The conditions (4.75) may be summarized as

$$\mathcal{S}^T \Omega \mathcal{S} = \Omega \quad \text{where} \quad \Omega = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}$$

11 If the initial f_{AB} is non-generic, then the matrix $\mathbb{1}$ in the last equation can be replaced by any matrix which commutes with f_{AB}. For generic f_{AB}, this must be a constant multiple of the unit matrix. The constant, which should be positive to preserve the sign of the kinetic energy of the vectors, can be absorbed by rescaling the matrices A, B, C, D.

The duality transformations in four dimensions are transformations in the symplectic group $\mathrm{Sp}(2m, \mathbb{R})$.

The matrix Ω is often called the symplectic metric, and the transformations (4.71) are then called symplectic transformations. This is the main result originally derived in [26]. Duality transformations in four spacetime dimensions are transformations of the group $\mathrm{Sp}(2m, \mathbb{R})$, which is a non-compact group.

Exercise 4.16 The dimension of the group $\mathrm{Sp}(2m, \mathbb{R})$ is the number of elements of the matrix $\mathcal{S}$, namely $4m^2$ minus the number of independent conditions contained in (4.77). Show that the dimension is $m(2m+1)$.

Duality transformations have two types of applications: they can describe symmetries of one theory and they can describe transformations from one theory to another. In the first case, the symmetries concerned form a subgroup of the 'maximal' duality group $\mathrm{Sp}(2m, \mathbb{R})$ discussed above. The subgroup consists of transformations (4.74) of $f_{AB}\,(\phi^i)$ induced by the symmetry transformations of the elementary scalars ϕ^i. These scalar transformations must be symmetries of the scalar kinetic term and other parts of the Lagrangian. The model of Sec. 4.2.3 is one example. The transformation of Z defined in (4.55) is the standard $\mathrm{SL}(2, \mathbb{R})$ symmetry of the Poincaré plane. This could be part of the full symmetry group of all the scalar fields of the theory. In extended supergravities it turns out that all the symmetry transformations that act on the scalars appear also as transformations of the vector kinetic matrix. Hence, the symmetry group is then a subgroup of the 'maximal' group $\mathrm{Sp}(2m, \mathbb{R})$ discussed above.

However, another application is of the type that we encountered in Ex. 4.15. In that case constants that specify the theory under consideration change under the duality transformations. The constants that transform are sometimes called 'spurionic quantities'. The transformations thus relate two different theories. Solutions of one theory are mapped into solutions of the other one. This is the basic idea of dualities in M-theory.

Symplectic transformations always transform solutions of (4.70) into other solutions. However, they are not always invariances of the action. Indeed, writing

$$\mathcal{L} = -\frac{1}{2} \operatorname{Re} \left(f_{AB} F^{-A}_{\mu\nu} F^{\mu\nu - B} \right) = -\frac{1}{2} \operatorname{Im} \left(F^{-A}_{\mu\nu} G^{\mu\nu -}_{A} \right)$$

we obtain

$$\operatorname{Im} F'^{-}G'^{-} = \operatorname{Im}\left(F^{-}G^{-}\right) + \operatorname{Im}\left[2F^{-}\left(C^{T}B\right)G^{-} + F^{-}\left(C^{T}A\right)F^{-} + G^{-}\left(D^{T}B\right)G^{-}\right].$$

If $C \neq 0, B = 0$ the Lagrangian is invariant up to a 4-divergence, since $\operatorname{Im} F^{-}F^{-} = -\frac{1}{4}\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}$ and the matrices A and C are real constants. For $B \neq 0$ neither the Lagrangian nor the action is invariant.

Electromagnetic duality has important applications to black hole solutions of extended supergravity theories. Supergravity is also very relevant to the analysis of black hole solutions of string theory. Many black holes are dyons; they carry both magnetic and electric charges for the gauge fields of the system. The general situation is a generalization of what was discussed at the end of Sec. 4.2.3. The charges form a symplectic vector $\begin{pmatrix} q^{A}_{\rm m} \\ q_{{\rm e}A} \end{pmatrix}$ which must transform as in (4.71). The Dirac-Schwinger-Zwanziger quantization condition restricts these charges to a lattice. Invariance of this lattice restricts the symplectic transformations of (4.71) to a discrete subgroup $\mathrm{Sp}(2m,\mathbb{Z})$, which is analogous to the $\mathrm{SL}(2,\mathbb{Z})$ group discussed previously.

Finally, we comment that symplectic transformations with $B \neq 0$ should be considered as non-perturbative for the following reasons. A system with no magnetic charges as in classical electromagnetism is transformed to a system with magnetic charges. The elements of f_{AB} may be regarded as coupling constants (see Ex. 4.15), and a system with weak coupling is transformed to one with strong coupling. A duality transformation which mixes electric and magnetic fields cannot be realized by transformation of the vector potential A_μ. One would need a 'magnetic' partner of A_μ to reexpress the $F'_{\mu\nu}$ and $G'_{\mu\nu}$ in terms of potentials.

The important properties of the matrix f_{AB} are that it is symmetric and that $\operatorname{Re} f_{AB}$ define a positive definite quadratic form in order to have positive gauge field energy. These properties are preserved under symplectic transformations defined by (4.74).

On general formulas

The historical standard reference on dualities in field theories is [2], though they appeared before in [13, 14, 15] and some extensions are in [16]. Duality transformations were explained in [1, Sec. 4.2] for coupled Maxwell fields in $D = 4$. We want to extend these transformations to the other parts of the action, including all other fields ϕ (bosons and fermions). Here is first a summary of the main concepts that were in [1, Sec. 4.2]. We considered there actions $S(F)$ that depend on field strengths $F^{A}_{\mu\nu}$, which are determined in terms of (abelian) vectors A^{A}_{μ}. We consider actions at most quadratic in spacetime derivatives, and thus also at most quadratic in $F^{A}_{\mu\nu}$. Having introduced the dual and self-dual combinations in [1, (4.35 – 36)][1]

$$\tilde{F}_{\mu\nu} = -\frac{1}{2}\mathrm{i}e\varepsilon_{\mu\nu\rho\sigma}F^{\rho\sigma}, \quad F^{\pm}_{\mu\nu} \equiv \frac{1}{2}\left(F_{\mu\nu} \pm \tilde{F}_{\mu\nu}\right),$$

the Bianchi identities and equation of motions for the vectors can be written as

$$\begin{aligned} \nabla^{\mu} \operatorname{Im} F^{A+}_{\mu\nu} &= 0 \quad \text{Bianchi identities,} \\ \nabla_{\mu} \operatorname{Im} G^{\mu\nu+}_{A} &= 0 \quad \text{Equations of motion of } A^{A}_{\nu}, \end{aligned}$$

where

$$G^{+\mu\nu}_{A} \equiv 2\mathrm{i}e^{-1}\frac{\delta S\left(F^{+},F^{-},\phi\right)}{\delta F^{+A}_{\mu\nu}}, \quad \text{i.e.} \quad \tilde{G}^{\mu\nu}_{A} = 2\mathrm{i}e^{-1}\frac{\delta S(F,\phi)}{F^{A}_{\mu\nu}}.$$

In the first part above, the action $S(F,\phi) = S\left(F^{+},F^{-},\phi\right)$ is considered as a function of the self-dual and anti-self-dual parts of F. In [1, Sec. 4.2] only the part of the action quadratic in F was considered, and thus G was linear in F. Now we consider that there can also be parts independent of F (terms in the action linear in F) and thus (here $\mathcal{N}_{AB} = -\mathrm{i}\bar{f}_{AB}$ in terms of f_{AB} in chapter 4 (as also done in ch. 21))

$$G^+_{A\mu\nu} = \mathcal{N}_{AB}(\phi)F^{+B}_{\mu\nu} + H^+_{A\mu\nu}(\phi) = G^+_{\mathrm{b}A\mu\nu} + H^+_{A\mu\nu}(\phi).$$

Since $\mathcal{N}$ is by the last two equations the second derivative of the action w.r.t. F, it is a symmetric tensor. Since the indices μ,ν should go somewhere and we do not consider higher derivative actions, the H will in practice depend on fermion bilinears, but we just write that they are functions of all the fields ϕ.

The dynamical equations (A.2) are then invariant under real symplectic transformations [2]

$$\delta_d \begin{pmatrix} F^A_{\mu\nu} \\ G_{A\mu\nu} \end{pmatrix} = \begin{pmatrix} A^A_B & B^{AB} \\ C_{AB} & D_A^B \end{pmatrix} \begin{pmatrix} F^B_{\mu\nu} \\ G_{B\mu\nu} \end{pmatrix},$$
$$B^{AB} = B^{BA}, \quad C_{AB} = C_{BA}, \quad D_A^B = -A^A_B.$$

[1] We insert here factors e as in [1, (7.59)] to take into account that we can be in curved space-time, though this is not important for the duality transformations. After the first definitions, we will omit the factors e.

[2] We write here the infinitesimal transformations, while in [1, (4.71)], A, B, C, D were used for global transformations

Now we can consider also the field equations for the other fields, which we denote as ϕ^i. We can omit the frame fields e^a_μ, which are inert under the duality transformations. We write

$$E_i \equiv \frac{\delta S}{\delta \phi^i},$$

using DeWitt notation, which means for an action that is function of fields and derivatives of fields (and adding a total derivative)

$$S = \int \mathrm{d}^4 x e L\left(\phi, \partial_\mu \phi\right), \quad \frac{\delta S}{\delta \phi^i} = e\left[\frac{\delta L}{\delta \phi^i} - \nabla_\mu \frac{\delta L}{\delta \partial_\mu \phi^i}\right].$$

We will use the notation that a derivative w.r.t. a field is a left derivative. For bosons left or right derivative makes no difference.

We consider transformations of these fields under the duality transformations:

$$\delta_d \phi^i = \xi^i(\phi),$$

where ξ^i is not dependent on $F_{\mu\nu}$. Full duality transformation can then be written as

$$\delta_d = \xi^i \frac{\delta}{\delta \phi^i} + \left(A^A_B F^B_{\mu\nu} + B^{AB} G_{B\mu\nu}\right) \frac{\delta}{\delta F^A_{\mu\nu}},$$

E.g. the total transformation of the action $S(F,\phi)$ is

$$\begin{aligned} \delta_d S &= \left(\xi^i \frac{\delta}{\delta \phi^i} + \left(F^T A^T + G^T B\right) \frac{\delta}{\delta F}\right) S \\ &= \xi^i E_i - \frac{1}{2}\left(\mathrm{i}\left(F^{+T} A^T + G^{+T} B\right) G^+ + \text{h.c. }\right) \end{aligned}$$

Check this equation, and also (A.3), using properties of dual tensors explained in [1, Sec. 4.2.1 and Ex. 4.6]. From here onwards we use simplifications in the equations. We consider matrix multiplication to write the expression in the brackets in (A.9). We omit the indices $[\mu\nu]$ on F^A and G_A and they are summed over in (A.10). We use DeWitt notation, [3] i.e.

$\xi^i E_i$ contains an integral over spacetime, and derivatives w.r.t. spacetime can be treated as in (A.7). Furthermore we will omit dependence on the frame. fields. They can be reinserted in an obvious way such that they reinstall general coordinate transformations. It is easier to work with self-dual combinations as in the last expression of (A.10). Some of the above relations are then simpler written as

$$\frac{\delta}{\delta F^+}S = -\frac{1}{2}\mathrm{i}G^+, \quad \frac{\partial G_A^+}{\partial F^{+B}} = \mathcal{N}_{AB}.$$

[3] This simplifies many expressions in [2], where the first term would be written as $\int \mathrm{d}^4x \left[\xi^i \frac{\delta L}{\delta\phi^i} + \partial_\mu \xi^i \frac{\delta L}{\delta\partial_\mu\phi^i}\right]$

It is useful to write the commutator of field derivatives with δ_d: (with $\partial_i = \frac{\delta}{\delta\phi^i}$ and using that ξ^i does not depend on F^A)

$$\begin{aligned}\frac{\delta}{\delta F^+}\delta_d &= \delta_d \frac{\delta}{\delta F^+} + \left(A^T + \mathcal{N}B\right)\frac{\delta}{\delta F^+}\\ \partial_i\delta_d &= \delta_d\partial_i + \left(\partial_i\xi^j\right)\partial_j + \left(\partial_i G^{+T}B\frac{\delta}{\delta F^+} + \text{h.c.}\right).\end{aligned}$$

On the action we thus obtain

$$\begin{aligned}\frac{\delta}{\delta F^+}\delta_d S &= -\frac{1}{2}\mathrm{i}\left(\delta_d G^+ + \left(A^T + \mathcal{N}B\right)G^+\right) = -\frac{1}{2}\mathrm{i}\left[CF^+ + \mathcal{N}\mathcal{B}BG^+\right]\\ &= -\frac{1}{4}\mathrm{i}\frac{\delta}{\delta F^+}\left[F^{+T}CF^+ + G^{+T}BG^+\right]\\ \partial_i\delta_d S &= \delta_d E_i + \left(\partial_i\xi^j\right)E_j - \frac{1}{2}\left(\mathrm{i}\left(\partial_i G^{+T}B\right)G^+ + \text{h.c.}\right)\end{aligned}$$

Thus we find

$$\delta_d S = -\frac{1}{4}\mathrm{i}\left[F^{+T}CF^+ + G^{+T}BG^+\right] + \text{h.c.}\,, \quad \delta_d E_i = -\left(\partial_i\xi^j\right)E_j.$$

For the first equation, note that this is, reinserting all notations,

$$\delta_d S = -\frac{1}{8}\int \mathrm{d}^4x\varepsilon^{\mu\nu\rho\sigma}\left(F_{\mu\nu}^A C_{AB}F_{\rho\sigma}^B + G_{A\mu\nu}D^{AB}G_{B\rho\sigma}\right),$$

and the first term is a total derivative. The second equation in (A.14) then implies that field equations transform to field equations, hence preserving the dynamics.

Note that the transformation $\delta_d S$ is the transformation of

$$S_{non-inv} = -\frac{1}{4}\mathrm{i}F^{+A}G_A^+ + \text{h.c.} \;= S_2(F,\phi) - \frac{1}{4}\left(\mathrm{i}F^{+A}H_A^+ + \text{h.c.}\right),$$

where $S_2(F,\phi)$ is the part of the action that is quadratic in F and can be written in various ways:

$$\begin{aligned}S_2(F,\phi) &= -\frac{1}{4}\mathrm{i}F^{+A}G_{\mathrm{b}A}^+ + \text{h.c.} \;= -\frac{1}{4}\mathrm{i}F^{+A}\mathcal{N}_{AB}F^{+B} + \text{h.c.}\\ &= \frac{1}{4}\int \mathrm{d}^4x\left[e\left(\operatorname{Im}\mathcal{N}_{AB}\right)F_{\mu\nu}^A F^{B\mu\nu} - \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}\left(\operatorname{Im}\mathcal{N}_{AB}\right)F_{\mu\nu}^A F_{\rho\sigma}^B\right],\end{aligned}$$

where we gave also the full expression in curved space. The linear part in F is

$$\begin{aligned}S_1(F,\phi) &= -\frac{1}{2}\mathrm{i}F^{+A}H_A^+ + \text{h.c.}\,,\\ &= -\frac{1}{4}\int \mathrm{d}^4x\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}^A H_{A\rho\sigma}.\end{aligned}$$

This implies that one half of this $S_1(F, \phi)$ sits in $S_{non-inv}$ and the other half is in the invariant part

$$S_{\text{inv}} = \frac{1}{2} S_1(F, \phi) + S_0(\phi),$$

where $S_0(\phi)$ is the part of the action without gauge fields: Note that this is also the remaining part when $G_{\mu\nu} = 0$, i.e. when the equations of motion of the vector fields are satisfied.

We can obtain aspects of the duality transformations more in detail. First we consider the consistency of (A.4) with (A.5). This implies that

$$\begin{aligned}\delta_d \boldsymbol{\mathcal{N}}_{AB}(\phi) &= \xi^i \partial_i \boldsymbol{\mathcal{N}}_{AB} = \left(C - \boldsymbol{\mathcal{N}} A - A^T \boldsymbol{\mathcal{N}} - \boldsymbol{\mathcal{N}} B \boldsymbol{\mathcal{N}}\right)_{AB} \\ \delta_d H^+(\phi) &= \xi^i \partial_i H^+ = -\left(A^T + \boldsymbol{\mathcal{N}} B\right) H^+.\end{aligned}$$

The transformation of the first term of the invariant part (A.19) under the duality is

$$\delta_d \frac{1}{2} S_1(F, \phi) = -\frac{1}{4} \mathrm{i} H^{+T} B H^+,$$

which is only function of the ϕ^i and should thus be compensated by a transformation of $S_0(\phi)$.

The way in which a H^+in agreement with the transformation in (A.20) appears, is from a symplectic vector $\left(Q^{+A}, P_A^+\right)$with $P_A^+ = \overline{\boldsymbol{\mathcal{N}}}_{AB} Q^{+B}$. Check that this is then a 'symplectic vector', which means that the vector transforms as $\left(F^A, G_A\right)$ in (A.5. The $H^+$from this symplectic vector is

$$H_{\mu\nu}^+ = \left(\boldsymbol{\mathcal{N}} Q^+ - P^+\right)_{\mu\nu} = 2\mathrm{i}(\operatorname{Im} \boldsymbol{\mathcal{N}}) Q_{\mu\nu}^+.$$

Then

$$\frac{1}{2} S_1(F, \phi) = -\frac{1}{4} \mathrm{i} F^{+A} \left(\boldsymbol{\mathcal{N}} Q^+ - P^+\right) = \frac{1}{4} \mathrm{i} \left(F^{+A} P_A^+ - G_{\mathrm{b}}^{+A} Q_A^+\right) + \text{h.c.} \ .$$

If we replace G_{b} by the full G, this is a symplectic invariant. Hence this determines which parts of $S_0(\phi)$ are separately invariant:

$$S_0(\phi) = \frac{1}{4} \mathrm{i} H_{A\mu\nu}^+ Q^{+A\mu\nu} + \text{ h.c. } + S_{0, \text{ inv}} \ .$$

Note that by the remarks after (A.4) the first term is in practice a 4 -fermion term. The scalar action is thus in the invariant part, which means that the scalar transformations should be isometries. This will determine a subalgebra of the symplectic algebra that is the symmetry of the theory. Hopefully this will be clarified with the example.

4.2 Duality by Gaillard-Zumino

(напишу потом)

4.2.1 Introduction and formalism by Gaillard, Zumino

Introduction by Gaillard, Zumino

It has long been known that the free Maxwell's equations are invariant under a rotation of the electric field and the magnetic field into each other. In relativistic notation this means that the electromagnetic field strength $F_{\mu\nu}$ and its dual,

$$\tilde{F}_{\mu\nu} = \frac{1}{2} \varepsilon_{\mu\nu\rho\sigma} F^{\rho\sigma},$$

rotate into each other. For this reason the transformation has been called a "duality rotation". It is easy to see [1, 2] that duality rotation invariance can be extended to the case when the electromagnetic field interacts with the gravitational field, which does not transform under duality. On the other hand, it is obvious that duality invariance is violated by electromagnetic couplings of the minimal type. Related to this is the fact that there is no non-abelian generalization of duality under which the pure Yang-Mills equations could be invariant [2].

Non-minimal couplings of the magnetic moment type can, however, be made duality invariant, and this invariance can in fact be generalized to a non-abelian group. This is the situation [3] in extended supergravity theories without gauging. The assumption that the theory is invariant under duality rotations has been used [4-6] to simplify the search for the correct supersymmetric lagrangian. For $N = 4$ supergravity it was discovered [6] that the U(4) duality could be extended to a larger $\mathrm{SU}(4) \times \mathrm{SU}(1,1)$ non-compact duality invariance. Similar situations occur for $N > 4$ supergravities; in the particular case of $N = 8$ supergravity, Cremmer and Julia [7] have shown that the theory is invariant under a non-compact E_7 duality. A non-compact duality invariance is possible only when there are scalar fields in the theory, and is related to non-linear transformations of the scalars.

Our purpose in this paper is to clarify the structure of theories admitting both compact and non-compact duality. Our analysis will not make explicit reference to supersymmetry, although of course we have in mind the application to supergravity theories. In fact, we have been strongly influenced by the work of Cremmer and Julia [7] on $N = 8$ supergravity, and some of our results can be found in their paper although in a less explicit and systematic form.

The first point we wish to emphasize is that the requirement that the equations of motion be duality invariant is not identical to the invariance of the lagrangian. This is already apparent in the case of the free electromagnetic field, where the lagrangian is

$$L = \frac{1}{2}\left(E^2 - H^2\right)$$

and is obviously not invariant under rotations * of E into H. Nevertheless, when the equations of motion are duality invariant, the lagrangian has some special properties; in particular it changes under a duality transformation in a specific way which will be described below.

A second point we make is that the invariance of the equations of motion under duality rotations implies the existence of conserved currents, both for the compact and the non-compact cases. These currents are constructed in terms of the basic fields of the theory and in terms of a set of dual vector potentials, and are not invariant under the abelian gauge transformations up to which these potentials are determined. However, the currents change under these gauge transformations by the divergence of an antisymmetric tensor, and therefore the integrated charges are gauge invariant. These charges are in fact the generators of the duality rotations.

In sect. 2 we discuss the general properties of duality transformations. We find that the most general group which can be realized with n field strengths is the non-compact real symplectic group $\mathrm{Sp}(2n, \mathrm{R})$, which has $\mathrm{U}(n)$ as its maximal compact subgroup. In the absence of scalar fields, $\mathrm{U}(n)$ is the largest group of duality transformations **. In specific examples the actual group of duality transformations can be smaller. For example, in $N = 8$ supergravity there are 28 field strengths. The non-compact invariance is the $\mathrm{E}_{7(+7)}$ subgroup of $\mathrm{Sp}(56, \mathrm{R})$, and its maximal compact subgroup is the SU(8) subgroup of U(28). We derive the transformation property of the lagrangian which is required for the equations of motion to be duality invariant, and show that this property implies the existence of conserved currents and the invariance of the energy-momentum tensor. We further exploit this property for the explicit construction of the lagrangian, which we illustrate first by specializing to the compact case in sect. 3.

By the way, for the free Maxwell equations with n field strengths, the largest duality group is $\mathrm{GL}(n, \mathrm{C})$, the non-compact general linear group of complex $n \times n$ matrices, which operate on the complex field strengths $F + i\tilde{F}$. In the presence of interactions without scalars this is reduced to its maximal compact subgroup U(n) which can be enlarged to Sp($2n$, R) in the presence of interacting scalars. Observe however that, even for the free Maxwell theory, the energy-momentum tensor is only invariant under U(n).

In order to generalize to the non-compact case it is necessary to consider non-linear realizations of the symmetry group and the corresponding scalar lagrangians and currents which we discuss in sect. 4. In sect. 5 we describe in more detail non-compact duality invariant theories. In sect. 6 we conclude with some comments relevant to the $N = 8$ supergravity theory.

Lagrangian formalism in electrodynamics

We consider a lagrangian which is a function of n real field strengths $F^a_{\mu\nu}$ and of a certain number of additional fields $\chi^i \equiv \chi^i(x)$ (scalars or fermions) and their derivatives $\chi^i_\mu = \partial_\mu \chi^i \equiv \frac{\partial \chi^i}{\partial x^\mu}$:

$$L = L\big) F^a, \chi^i, \chi^i_\mu \big) .$$

Note that x are coordinates, but they will appear rear, so there should not be confusions. So from definitions we have properties:

$$\partial_\mu \xi^j = \chi^k_\mu \frac{\partial \xi^j}{\partial \chi^k}$$
$$\frac{\partial \chi_k}{\partial x} \frac{\partial^2 L}{\partial \chi^j \partial \chi^i} = \partial_\mu \frac{\partial L}{\partial \chi^j}.$$

The field strengths are the curls of vector potentials

$$F^a_{\mu\nu} \equiv \partial_\mu \mathcal{A}^a_\nu - \partial_\nu \mathcal{A}^a_\mu,$$

and therefore they satisfy

$$\partial^\mu \tilde{F}^a_{\mu\nu} = 0, \qquad \text{where} \qquad \tilde{F}_{\mu\nu} := \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma} F^{\rho\sigma},$$

because of symmetry-antisymmetry of the terms. (??? no, I don't know. черт, я забыл спросить!!!)

If we define the antisymmetric tensors $G^a_{\mu\nu}$ by

$$\tilde{G}^a_{\mu\nu} \equiv \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma} G^{a\rho\sigma} := 2\frac{\partial L}{\partial F^{a\mu\nu}}.$$

So:

$$\frac{\partial L}{\partial F^{a\mu\nu}} \equiv \frac{1}{2}\tilde{G}^a_{\mu\nu}.$$

The equations of motion obtained by varying $\mathcal{A}^a_\mu$ are

$$\partial^\mu \tilde{G}^a_{\mu\nu} = 0.$$

because

(тут эта гениальная идея про вариацию, напишу потом!!!!!!!!)

if current is non-zero, $\partial^\mu \tilde{G}^a_{\mu\nu} = \frac{4\pi}{c} j^\nu$. (??? yes??? why don't we write this for χ fields???)

Duality definitions

The definitions were:

$$\tilde{F}_{\mu\nu} := \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}F^{\rho\sigma},$$
$$\tilde{G}^{a}_{\mu\nu} \equiv \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}G^{a\rho\sigma} := 2\frac{\partial L}{\partial F^{a\mu\nu}}$$

The system of equations $\partial^\mu \tilde{F}^a_{\mu\nu} = 0$ and $\partial^\mu \tilde{G}^a_{\mu\nu} = 0$ is invariant under linear transformations among the F's and G's. We therefore consider an infinitesimal transformation of the form

$$\delta \begin{pmatrix} F \\ G \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} F \\ G \end{pmatrix},$$
$$\delta\chi^i = \xi^i(\chi(x)),$$

where A, B, C, D are arbitrary real $n \times n$ matrices and $\xi^i(\chi)$ non-derivative functions of the additional fields $\chi^i(x)$.

In "Duality Rotations in Nonlinear Electrodynamics and in Extended Supergravity" it is written that "a practical convention is to define $\frac{\partial F_{\rho\sigma}}{\partial F_{\mu\nu}} = \delta^\mu_\rho \delta^\nu_\sigma$ rather than $\frac{\partial F_{\rho\sigma}}{\partial F_{\mu\nu}} = \delta^\mu_\rho \delta^\nu_\sigma - \delta^\nu_\rho \delta^\mu_\sigma$. This explains the factor 2 in $\tilde{G}^{\mu\nu} = 2\frac{\partial \mathcal{L}(F)}{\partial F_{\mu\nu}}$. (????? what?????)

(!!! абзац про то, что от чего зависит! вообще, у нас F и G зависят друг от друга, а также от фермионных полей!)

In general F and G are independent (an exception is the free field case, $\tilde{G} = -F$). Because we have other field that consist in G (Gailard, Zumino wrote that).

In indices we write:

$$\delta F^a = A^{ab}F^b + B^{ab}G^b, \qquad \delta \tilde{F}^a = A^{ab}\tilde{F}^b + B^{ab}\tilde{G}^b,$$
$$\delta G^a = C^{ab}F^b + D^{ab}G^b. \qquad \delta \tilde{G}^a = C^{ab}\tilde{F}^b + D^{ab}\tilde{G}^b,$$

We omit $\mu\nu$ indexes, because it is obvious, where they should be.

(допишу, что вообще-то про матрицы все известно.)

We shall study the conditions under which G by the definition is invariant (????) and the equations of motion for the χ^i are invariant. Notice that we do not impose the invariance of the lagrangian itself: we shall see that the system of the equations of motion can be invariant only if δL does not vanish. Instead the variation of L is required to have a specific form which can be used to demonstrate the existence of a conserved current and leads to an essentially unique construction of the vector field couplings.

(???? I'll think more about the idea later, until I really understand what exactly we want and why we are doing it specifically.)

List of meaning of symbols, parameters and letters for GZ methods

(later I'll add a list of them not to get confused in a lot of letters!!!)

4.2.2 Derivations of properties of dual transformations by GZ

Variation of the lagrangian and its derivative

Итак, идея в том, что...(?? додумаю!!!)

The variation of the lagrangian $L = L\left(F^a, \chi^i, \chi^i_\mu\right)$ under the linear transformations is

$$\begin{aligned}
\delta L &= \left[\delta\chi^i \frac{\partial}{\partial\chi^i} + \delta\chi^i_\mu \frac{\partial}{\partial\chi^i_\mu} + \delta F^b \frac{\partial}{\partial F^b}\right] L = \\
&= \left[\xi^i \frac{\partial}{\partial\chi^i} + \frac{\partial\delta\chi^i}{\partial x^\mu}\frac{\partial}{\partial\chi^i_\mu} + \left(F^c A^{bc} + G^c B^{bc}\right)\frac{\partial}{\partial F^b}\right] L = \\
&= \left[\xi^i \frac{\partial}{\partial\chi^i} + \frac{\partial\xi^i}{\partial\chi^j}\frac{\partial\chi^j}{\partial x^\mu}\frac{\partial}{\partial\chi^i_\mu} + \left(F^c A^{bc} + G^c B^{bc}\right)\frac{\partial}{\partial F^b}\right] L = \\
&= \left[\xi^i \frac{\partial}{\partial\chi^i} + \chi^j_\mu \frac{\partial\xi^i}{\partial\chi^j}\frac{\partial}{\partial\chi^i_\mu} + \left(F^c A^{bc} + G^c B^{bc}\right)\frac{\partial}{\partial F^b}\right] L
\end{aligned}$$

We differentiate δL with respect to F^a, remembering that here $\delta \equiv \delta(F^a)$, so $\frac{\partial(\delta L)}{\partial F^a} = \frac{\partial\delta(F)}{\partial F^a} L + \delta\left(\frac{\partial L}{\partial F^a}\right)$:

$$\begin{aligned}
\frac{\partial}{\partial F^a}\delta L &= A^{ba}\frac{\partial L}{\partial F^b} + \frac{\partial G^c}{\partial F^a} B^{bc}\frac{\partial L}{\partial F^b} + \delta\frac{\partial L}{\partial F^a} = \\
&= \frac{1}{2}A^{ba}\tilde{G}^b + \frac{1}{2}\frac{\partial G^c}{\partial F^a}B^{bc}\tilde{G}^b + \frac{1}{2}\delta\tilde{G}^a.
\end{aligned}$$

If we now require our linear transformation $\delta\tilde{G}^a = C^{ab}\tilde{F}^b + D^{ab}\tilde{G}^b$, we obtain

$$\begin{aligned}
2\frac{\partial\delta L}{\partial F^a} =& C^{ab}\tilde{F}^b + \left(D^{ab} + A^{ba}\right)\tilde{G}^b + \frac{\partial G^c}{\partial F^a}B^{bc}\tilde{G}^b = \\
=& \frac{1}{2}C^{ab}\tilde{F}^b + \frac{1}{2}\frac{\partial G^c}{\partial F^a}B^{bc}\tilde{G}^b + \left(D^{ab} + A^{ba}\right)\cdot 2\frac{\partial L}{\partial F^{a\mu\nu}} + \frac{1}{2}C^{ab}\tilde{F}^b + \frac{1}{2}\frac{\partial G^c}{\partial F^a}B^{bc}\tilde{G}^b = \\
=& \frac{1}{2}C^{ab}\tilde{F}^b + \frac{1}{2}C^{ba}\tilde{F}^b + \frac{1}{2}\frac{\partial G^c}{\partial F^a}B^{bc}\tilde{G}^b + \frac{1}{2}\frac{\partial G^c}{\partial F^a}B^{bc}\tilde{G}^b + 2\left(D^{ab} + A^{ba}\right)\frac{\partial L}{\partial F^{a\mu\nu}} + \\
& + \left[\frac{1}{2}C^{ab}\tilde{F}^b - \frac{1}{2}C^{ba}\tilde{F}^b + \frac{1}{2}\frac{\partial G^c}{\partial F^a}B^{bc}\tilde{G}^b - \frac{1}{2}\frac{\partial G^c}{\partial F^a}B^{bc}\tilde{G}^b\right] = \\
=& \frac{1}{2}\frac{\partial(FC\tilde{F})}{\partial F^a} + \frac{1}{2}\frac{\partial(GB^{\mathrm{T}}\tilde{G})}{\partial F^a} + 2\left(D^{ab} + A^{ba}\right)\frac{\partial L}{\partial F^{a\mu\nu}} + \frac{1}{2}\left[\left(C^{ab} - C^{ba}\right)\tilde{F}^b + \frac{\partial G^c}{\partial F^a}\left(B^{bc} - B^{cb}\right)\tilde{G}^b\right] = \\
=& \frac{1}{2}\frac{\partial}{\partial F^a}\left(FC\tilde{F} + GB^{\mathrm{T}}\tilde{G}\right) + 2\left(D^{ab} + A^{ba}\right)\frac{\partial L}{\partial F^b} + \frac{1}{2}\left[\left(C^{ab} - C^{ba}\right)\tilde{F}^b + \frac{\partial G^c}{\partial F^a}\left(B^{bc} - B^{cb}\right)\tilde{G}^b\right].
\end{aligned}$$

Here, as discussed before, (???? напишу это!!!) "grey" terms vanish because this expression should be a derivative with respect to F^a, and we have used the property

$$\begin{aligned}
\frac{\partial\tilde{F}^a_{\mu\nu}}{\partial F^{b\rho\sigma}} &\equiv \frac{1}{2}\frac{\partial\varepsilon_{\mu\nu\lambda\theta}F^{a\lambda\theta}}{\partial F^{b\rho\sigma}} = \frac{1}{2}\varepsilon_{\mu\nu\lambda\theta}\delta_{\lambda\rho}\delta_{\theta\sigma}\delta^b_a = \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}\delta^b_a = \frac{1}{2}\varepsilon_{\rho\sigma\mu\nu}\delta^b_a, \\
\frac{\partial(FC\tilde{F})}{\partial F^a} &\equiv \frac{\partial(F^iC^{ib}\tilde{F}^b)}{\partial F^a} = \delta^a_i C^{ib}\tilde{F}^b + F^iC^{ib}\frac{\partial\tilde{F}^b}{\partial F^a} = C^{ab}\tilde{F}^b + F^{i\mu\nu}C^{ib}\frac{\varepsilon_{\rho\sigma\mu\nu}}{2}\delta^b_a \equiv \\
&\equiv C^{ab}\tilde{F}^b + \tilde{F}^iC^{ia} = C^{ab}\tilde{F}^b + C^{ba}\tilde{F}^b, \\
\frac{\partial(GB^{\mathrm{T}}\tilde{G})}{\partial F^a} &\equiv \frac{\partial(G^cB^{bc}\tilde{G}^b)}{\partial F^a} = \frac{\partial G^c}{\partial F^a}B^{bc}\tilde{G}^b + G^cB^{bc}\frac{\partial\tilde{G}^b}{\partial F^a} = \frac{\partial G^c}{\partial F^a}B^{bc}\tilde{G}^b + G^bB^{cb}\frac{\partial\tilde{G}^c}{\partial F^a} \equiv \\
&\equiv \frac{\partial G^c}{\partial F^a}B^{bc}\tilde{G}^b + \frac{\partial\tilde{G}^c}{\partial F^a}B^{cb}G^b.
\end{aligned}$$

In general $\tilde{F}$ and $\tilde{G}$ are independent (an exception is the free field case, $\tilde{G} = -F$), so to vanish the "grey" terms we must impose the conditions below on the infinitesimal transformation matrices:

$$C = C^{\mathrm{T}},$$
$$B = B^{\mathrm{T}},$$
$$D^{ab} + A^{ba} = \varepsilon\delta^{ab}.$$

Now we have

$$\frac{\partial}{\partial F^a}\delta L = \frac{\partial}{\partial F^a}\left(\frac{1}{4}FC\tilde{F} + \frac{1}{4}GB\tilde{G} + \varepsilon L\right).$$

Equation of motion

Next we study the equations of motion for χ^i using the properties of matrices in the linear transformation. We define

$$E_i := \frac{\partial L}{\partial \chi^i} - \partial_\mu \frac{\partial L}{\partial \chi^i_\mu},$$

so that the equations of motion are

$$E_i = 0.$$

The covariance of these equations under the linear dual transformation requires

$$\delta E_i = -\frac{\partial \xi^j}{\partial \chi^i} E_j$$

since $\delta_F + \delta_\chi + \delta_{\chi_k} = 0$.

Euler-Lagrange operator applies to δL as showed below. But first let us note that

$$(\partial_\mu F^c)\frac{\partial}{\partial F^b}\frac{\partial L}{\partial \chi^i_\mu} = (\partial_\mu F^c)\frac{\partial \tilde{G}}{\partial \chi^i_\mu} = (\partial_\mu \tilde{F}^c)\frac{\partial G}{\partial \chi^i_\mu} \equiv 0 \qquad (\partial_\mu \tilde{F}^c = 0)$$
$$(\partial_\mu G^b)\frac{\partial^2 L}{\partial F^b \partial \chi^i_\mu} = (\partial_\mu \tilde{G}^b)\frac{\partial G}{\partial \chi^i_\mu} = 0 \qquad (\partial_\mu \tilde{G}^c = 0)$$
$$\frac{\partial^2 \xi^\lambda}{\partial \chi^i \chi^j} \equiv 0,$$
$$\partial_\mu \chi^i_\rho \equiv 0$$
$$\partial_\mu \chi^j_\rho \equiv 0,$$

since we are working with partial derivatives. Also note that below i is a free index.

(ниже все еще индексы нужно расставить, ну и еще раз пересмотрю это.)

We have:

$$\left(\frac{\partial}{\partial \chi^i} - \partial_\mu \frac{\partial}{\partial \chi^i_\mu}\right)\delta L \equiv \left(\frac{\partial}{\partial \chi^i} - \partial_\mu \frac{\partial}{\partial \chi^i_\mu}\right)\left[\xi^j \frac{\partial}{\partial \chi^j} + \chi^j_\mu \frac{\partial \xi^\lambda}{\partial \chi^j}\frac{\partial}{\partial \chi^\lambda_\mu} + \left(F^c A^{bc} + G^c B^{bc}\right)\frac{\partial}{\partial F^b}\right] L =$$

$$\left(\text{act with only } \frac{\partial}{\partial \chi^i} \text{ and } \frac{\partial}{\partial \chi^i_\mu}\right)$$

$$= \frac{\partial \xi^j}{\partial \chi^i}\frac{\partial L}{\partial \chi^j} + \xi^i \frac{\partial^2 L}{\partial \chi^i \partial \chi^j} + \chi^j_\mu \frac{\partial^2 \xi^\lambda}{\partial \chi^j \partial \chi^i}\frac{\partial L}{\partial \chi^\lambda_\mu} + \chi^j_\mu \frac{\partial \xi^\lambda}{\partial \chi^j}\frac{\partial^2 L}{\partial \chi^i \partial \chi^\lambda_\mu} + \frac{\partial G^c}{\partial \chi^i} B^{bc}\frac{\partial L}{\partial F^b} + \left(F^c A^{bc} + G^c B^{bc}\right)\frac{\partial^2 L}{\partial F^b \partial \chi^i} -$$
$$- \partial_\mu \left[\xi^j \frac{\partial^2 L}{\partial \chi^j \partial \chi^i_\mu} + \frac{\partial \xi^\lambda}{\partial \chi^i}\frac{\partial L}{\partial \chi^\lambda_\mu} + \chi^j_\rho \frac{\partial \xi^\lambda}{\partial \chi^j}\frac{\partial^2 L}{\partial \chi^\lambda_\rho \partial \chi^i_\mu} + F^c A^{bc}\frac{\partial}{\partial F^b}\frac{\partial L}{\partial \chi^i_\mu} + \frac{\partial G^b}{\partial \chi^i_\mu} B^{bc}\frac{\partial L}{\partial F^b} + G^b B^{bc}\frac{\partial^2 L}{\partial F^b \partial \chi^i_\mu}\right] =$$

(colorise some terms and act with ∂_μ on terms in the bracket)

$$=\frac{\partial\xi^j}{\partial\chi^i}\frac{\partial L}{\partial\chi^j}+\xi^i\frac{\partial^2 L}{\partial\chi^i\partial\chi^j}+\chi^j_\mu\frac{\partial^2\xi^\lambda}{\partial\chi^j\partial\chi^i}\frac{\partial L}{\partial\chi^\lambda_\mu}+\chi^j_\mu\frac{\partial\xi^\lambda}{\partial\chi^j}\frac{\partial^2 L}{\partial\chi^i\partial\chi^\lambda_\mu}+\frac{\partial G^c}{\partial\chi^i}B^{bc}\frac{\partial L}{\partial F^b}+\left(F^cA^{bc}+G^cB^{bc}\right)\frac{\partial^2 L}{\partial F^b\partial\chi^i}-$$

$$-\chi^k_\mu\frac{\partial\xi^j}{\partial\chi^k}\frac{\partial^2 L}{\partial\chi^j\partial\chi_\mu}-\xi^j\partial_\mu\frac{\partial^2 L}{\partial\chi^j\partial\chi^i_\mu}-\chi^j_\rho\frac{\partial\xi^\lambda}{\partial\chi^j}\partial_\mu\frac{\partial^2 L}{\partial\chi^\lambda_\rho\partial\chi^i_\mu}-(\partial_\mu\chi^i_\rho)\frac{\partial\xi^\lambda}{\partial\chi}\frac{\partial^2 L}{\partial\chi^i\chi}-\chi^i_\rho\left(\partial_\mu\frac{\partial\xi^\lambda}{\partial\chi}\right)\frac{\partial^2 L}{\partial\chi^i\chi}$$

$$-(\partial_\mu F^c)A^{bc}\frac{\partial}{\partial F^b}\frac{\partial L}{\partial\chi^i_\mu}-F^cA^{bc}\frac{\partial}{\partial F^b}\partial_\mu\frac{\partial L}{\partial\chi^i_\mu}-\partial_\mu\left(\frac{\partial G^b}{\partial\chi^i_\mu}B^{bc}\frac{\partial L}{\partial F^b}\right)-\partial_\mu\left(G^bB^{bc}\frac{\partial^2 L}{\partial F^b\partial\chi^i_\mu}\right)=$$

(collect colorized terms)

$$=\frac{\partial\xi^j}{\partial\chi^i}\frac{\partial L}{\partial\chi^j}-\chi^k_\mu\frac{\partial\xi^j}{\partial\chi^k}\frac{\partial^2 L}{\partial\chi^j\partial\chi^i_\mu}+\xi^i\frac{\partial^2 L}{\partial\chi^i\partial\chi^j}-\xi^j\partial_\mu\frac{\partial^2 L}{\partial\chi^j\partial\chi^i_\mu}+\chi^j_\mu\frac{\partial\xi^\lambda}{\partial\chi^j}\frac{\partial^2 L}{\partial\chi^i\partial\chi^\lambda_\mu}-\chi^j_\rho\frac{\partial\xi^\lambda}{\partial\chi^j}\partial_\mu\frac{\partial^2 L}{\partial\chi^\lambda_\rho\partial\chi^i_\mu}+$$

$$+\left(F^cA^{bc}+G^cB^{bc}\right)\frac{\partial^2 L}{\partial F^b\partial\chi^i}-F^cA^{bc}\frac{\partial}{\partial F^b}\partial_\mu\frac{\partial L}{\partial\chi^i_\mu}-\partial_\mu\left(G^bB^{bc}\frac{\partial^2 L}{\partial F^b\partial\chi^i_\mu}\right)+$$

$$+\frac{\partial G^c}{\partial\chi^i}B^{bc}\frac{\partial L}{\partial F^b}-\partial_\mu\left(\frac{\partial G^b}{\partial\chi^i_\mu}B^{bc}\frac{\partial L}{\partial F^b}\right)=$$

$$\left(\text{simplify terms by } \partial_\mu\xi^j=\chi^k_\mu\frac{\partial\xi^j}{\partial\chi^k},\qquad \frac{\partial\chi_k}{\partial x}\frac{\partial^2 L}{\partial\chi^j\partial\chi^i}=\partial_\mu\frac{\partial L}{\partial\chi^j}\right)$$

$$=\frac{\partial\xi^j}{\partial\chi^i}\frac{\partial L}{\partial\chi^j}-\frac{\partial\xi^i}{\partial\chi^j}\partial_\mu\frac{\partial L}{\partial\chi_\mu}+\xi^i\frac{\partial}{\partial\chi^i}\left(\frac{\partial L}{\partial\chi^j}-\partial_\mu\frac{\partial L}{\partial\chi^j_\mu}\right)+\chi^j_\mu\frac{\partial^2\xi^\lambda}{\partial\chi^j\partial\chi^i}\frac{\partial L}{\partial\chi^\lambda_\mu}-\chi^j_\rho\frac{\partial\xi^\lambda}{\partial\chi^j}\partial_\mu\frac{\partial^2 L}{\partial\chi^\lambda_\rho\partial\chi^i_\mu}+$$

$$+\left(F^cA^{bc}+G^cB^{bc}\right)\frac{\partial^2 L}{\partial F^b\partial\chi^i}-F^cA^{bc}\frac{\partial}{\partial F^b}\partial_\mu\frac{\partial L}{\partial\chi^i_\mu}-G^bB^{bc}\frac{\partial}{\partial F^b}\partial_\mu\frac{\partial^2 L}{\partial\chi^i_\mu}-$$

$$-(\partial_\mu G^b)B^{bc}\frac{\partial^2 L}{\partial F^b\partial\chi^i_\mu}+\frac{\partial G^c}{\partial\chi^i}B^{bc}\frac{\partial L}{\partial F^b}-\partial_\mu\left(\frac{\partial G^b}{\partial\chi^i_\mu}B^{bc}\frac{\partial L}{\partial F^b}\right)=$$

(rewrite terms in form of action to equation of motion)

$$=\frac{\partial\xi^j}{\partial\chi^i}\left(\frac{\partial L}{\partial\chi^j}-\partial_\mu\frac{\partial L}{\partial\chi_\mu}\right)+\xi^i\frac{\partial}{\partial\chi^i}\left(\frac{\partial L}{\partial\chi^j}-\partial_\mu\frac{\partial L}{\partial\chi^j_\mu}\right)+\xi^j_\mu\frac{\partial}{\partial\chi^j_\mu}\left(\frac{\partial L}{\partial\chi^j}-\partial_\mu\frac{\partial L}{\partial\chi_\mu}\right)+$$

$$+\left(F^cA^{bc}+G^cB^{bc}\right)\frac{\partial}{\partial F^b}\left(\frac{\partial L}{\partial\chi^i}-\partial_\mu\frac{\partial L}{\partial\chi^i_\mu}\right)+\frac{\partial G^c}{\partial\chi^i}B^{bc}\frac{\partial L}{\partial F^b}-\partial_\mu\left(\frac{\partial G^b}{\partial\chi^i_\mu}B^{bc}\frac{\partial L}{\partial F^b}\right)$$

(collect purple, pink and navy terms into a big navy term, insert E_i)

$$=\frac{\partial\xi^j}{\partial\chi^i}E_j+\left[\xi^i\frac{\partial}{\partial\chi^i}+\frac{\partial\delta\chi^i}{\partial x}\frac{\partial}{\partial\chi^i_\mu}+\left(F^cA^{bc}+G^cB^{bc}\right)\frac{\partial}{\partial F^b}\right]E_i+\frac{\partial G^c}{\partial\chi^i}B^{bc}\frac{\partial L}{\partial F^b}-\partial_\mu\left(\frac{\partial G^c}{\partial\chi^i_\mu}B^{bc}\frac{\partial L}{\partial F^b}\right)=$$

$$=\frac{\partial\xi^j}{\partial\chi^i}E_j+\delta E_i+\frac{\partial G^c}{\partial\chi^i}B^{bc}\frac{\partial L}{\partial F^b}-\partial_\mu\left(\frac{\partial G^c}{\partial\chi^i_\mu}B^{bc}\frac{\partial L}{\partial F^b}\right).$$

We have proven that

$$\left(\frac{\partial}{\partial\chi^i}-\partial_\mu\frac{\partial}{\partial\chi^i_\mu}\right)\delta L=\frac{\partial\xi^j}{\partial\chi^i}E_j+\delta E_i+\frac{\partial G^c}{\partial\chi^i}B^{bc}\frac{\partial L}{\partial F^b}-\partial_\mu\left(\frac{\partial G^c}{\partial\chi^i_\mu}B^{bc}\frac{\partial L}{\partial F^b}\right).$$

Also we know that $\delta E_i=-\frac{\partial\xi^j}{\partial\chi^i}E_j$ and $B=B^{\mathrm{T}}$, $\frac{\partial L}{\partial F^{a\mu\nu}}\equiv\frac{\tilde{G}^a_{\mu\nu}}{2}$, so we have

$$\left(\frac{\partial}{\partial\chi^i}-\partial_\mu\frac{\partial}{\partial\chi^i_\mu}\right)\delta L=\frac{1}{2}\frac{\partial G^c}{\partial\chi^i}B^{bc}\tilde{G}^b-\frac{1}{2}\partial_\mu\left(\frac{\partial G^c}{\partial\chi^i_\mu}B^{bc}\tilde{G}^b\right)=$$

$$=\frac{1}{4}\frac{\partial(GB\tilde{G})}{\partial\chi^i}-\frac{1}{4}\partial_\mu\frac{\partial(GB\tilde{G})}{\partial\chi^i}=$$

$$=\frac{1}{4}\left(\frac{\partial}{\partial\chi^i}-\partial_\mu\frac{\partial}{\partial\chi^i_\mu}\right)(GB\tilde{G})$$

so
$$\left(\frac{\partial}{\partial\chi^i}-\partial_\mu\frac{\partial}{\partial\chi^i_\mu}\right)\left(\delta L-\frac{1}{4}GB\tilde{G}\right)=0$$
We see that this equation is consistent with $\frac{\partial}{\partial F^a}\delta L=\frac{\partial}{\partial F^a}\left(\frac{1}{4}FC\tilde{F}+\frac{1}{4}GB\tilde{G}+\varepsilon L\right)$ only if
$$\varepsilon=0\qquad\Longleftrightarrow\qquad D=-A^{\mathrm{T}}$$
(???? no, I am not sure why it is so). We have obtained the transformation property:
$$\delta L=\frac{1}{4}(FC\tilde{F}+GB\tilde{G}).$$

We learned that dual transformations
$$\delta\begin{pmatrix}F\\G\end{pmatrix}=\begin{pmatrix}A&B\\C&D\end{pmatrix}\begin{pmatrix}F\\G\end{pmatrix},$$
$$C=C^{\mathrm{T}},\quad B=B^{\mathrm{T}},\qquad D=-A^{\mathrm{T}}.$$
is a transformation of $\mathrm{Sp}(2n,\mathrm{R})$, the real non-compact form of $\mathrm{Sp}(2n)$ in a real basis for the $2n$-dimensional representation.

Given two vectors F,G and H,I which transform by the same duality matrix, one can construct an invariant (??? don't fully see, why, maybe it is in the appendix)
$$(H\ \ I)\begin{pmatrix}0&-1\\1&0\end{pmatrix}\begin{pmatrix}F\\G\end{pmatrix}=IF-HG.$$
This property that will prove useful in the construction of the lagrangian. In specific examples the group of duality transformations leaving invariant the equations of motion may be a subgroup of $\mathrm{Sp}(2n,\mathrm{R})$.

(??? why they are not always a subgroup???)

Same derivations on example (!!????)

(давно хотел это же проделать, но на конкретном лагранжиане!!! вот и сделаю - и будет понятно, правильная логика моя была или нет.)

Идея для доказательства свойств дуальности (мб как Проен сказал, идея в том, мы просто обобщая темы, знаем, что доказать хотим, это и получаем. додумаю и напишу)

4.2.3 The conserved current

Derivation of the current

(! I'll rewrite it and start with the idea, how we will get it!!!)

In order to construct a conserved current we extract the variation of L due to the variation of the χ's alone:
$$\delta_\chi L:=\delta\chi^i\frac{\partial L}{\partial\chi^i}+\delta\chi^i_\mu\frac{\partial L}{\partial\chi^i_\mu}\equiv\delta L-\delta_F L$$
where
$$\delta_F L:=\delta F^a\frac{\partial L}{\partial F^a}=(A^{ab}F^b+B^{ab}G^b)\frac{1}{2}\tilde{G}^a=\frac{1}{2}\left(F^b(A^{\mathrm{T}})^{ba}\tilde{G}^a+G^b(B^{\mathrm{T}})^{ba}\tilde{G}^a\right)=\frac{1}{2}\left(FA^{\mathrm{T}}\tilde{G}+GB\tilde{G}\right),$$
$$\delta L=\frac{1}{4}(FC\tilde{F}+GB\tilde{G}),$$

which was obtained before. We have

$$\delta_\chi L = \frac{1}{4}FC\tilde{F} + \frac{1}{4}GB\tilde{G} + \frac{1}{2}FD\tilde{G} - \frac{1}{2}GB\tilde{G} = \frac{1}{4}(\tilde{F}CF - \tilde{G}BG) + \frac{1}{2}\tilde{F}DG =$$
$$= \frac{1}{4}\left(\tilde{F}CF - \tilde{G}BG + \tilde{F}DG + \tilde{F}DG\right) = \frac{1}{4}\left(\tilde{F}CF - \tilde{G}BG + \tilde{F}DG - \tilde{G}AF\right),$$

because of general property $\tilde{F}BG = GB^T\tilde{F}$.

From basic electrodynamics we know that equations of motion $\partial^\mu \tilde{G}^a_{\mu\nu} = 0$ imply that we can introduce a vector potential $\mathcal{B}_\mu$ of which $G_{\mu\nu}$ is the curl: $G^a_{\mu\nu} = \partial_\mu \mathcal{B}^a_\nu - \partial_\nu \mathcal{B}^a_\mu$.

If we consider transformations $\delta\phi^i(x) \equiv \epsilon^A \Delta_A \phi^i(x)$, we can obtain the current as

$$J^\mu{}_A := -\frac{\delta L}{\delta \partial_\mu \phi^i}\Delta_A \phi^i + K^\mu_A.$$

We need to find K^μ_A, such that by definition $\delta L = \epsilon^A \partial_\mu K^\mu_A$. We will consider only transformations of the scalar fields χ^i, so $\epsilon^A = 1, \delta\phi^i(x) \equiv \delta\chi = \xi^i$. We can check that if we define

$$\hat{K}^\mu := \frac{1}{2}\left(\tilde{G}^{\mu\nu}A\mathcal{A}_\nu - \tilde{F}^{\mu\nu}C\mathcal{A}_\nu + \tilde{G}^{\mu\nu}B\mathcal{B}_\nu - \tilde{F}^{\mu\nu}D\mathcal{B}_\nu\right),$$

than it has divergence

$$\partial_\mu \hat{K}^\mu = -\delta_\chi L,$$

so we need exactly $K^\mu = -\hat{K}^\mu$ for the term in the formula for the current. Indeed, the divergence has four parts:

$$2\partial_\mu K^\mu_A := \partial_\mu(\tilde{G}^{\mu\nu}A\mathcal{A}_\nu) = \partial_\mu(\tilde{G}^{\mu\nu})A\mathcal{A}_\nu + \tilde{G}^{\mu\nu}A\partial_\mu(\mathcal{A}_\nu) = 0 + \frac{1}{2}\tilde{G}^{\mu\nu}AF_{\mu\nu}$$

(!!! допишу нужные слова про этот переход, что у нас G тут есть антисимм, из-за нее так заменяем!!!)

Basically we showed that if we have some arbitraty electronagnetic tensors $P_{\mu\nu}$ with $\partial_\mu \tilde{P}^{\mu\nu} = 1$, arbitrary matrix M and $R_{\mu\nu} := \partial_\mu \mathcal{R}_\nu - \partial_\nu \mathcal{R}_\mu$, then

$$\partial_\mu(\tilde{P}^{\mu\nu}M\mathcal{R}_\nu) = \frac{1}{2}\tilde{P}_{\mu\nu}MR^{\mu\nu} \equiv \frac{1}{2}\tilde{P}MR.$$

So we see how the divergence of different parts of $\hat{K}$ looks like. We see that

$$\partial_\mu K^\mu_A = \frac{1}{2}\left(\frac{1}{2}\tilde{G}AF - \frac{1}{2}\tilde{F}CF + \frac{1}{2}\tilde{G}BG - \frac{1}{2}\tilde{F}DG\right) \equiv -\delta_\chi L.$$

The current off-shell is not a total derivative, because $G_{\mu\nu}$ is a curl only in virtue of the equations of motion $\partial^\mu \tilde{G}^a_{\mu\nu} = 0$.

Now, by the usual Noether argument, the equations of $E_i \equiv \frac{\partial L}{\partial \chi^i} - \partial_\mu \frac{\partial L}{\partial \chi^i_\mu} = 0$ for χ^i imply that

$$\delta_\chi L \equiv \delta\chi \frac{\partial L}{\partial \chi} + \delta\chi_\mu \frac{\partial L}{\partial \chi^i_\mu} = \delta\chi \partial_\mu \frac{\partial L}{\partial \chi^i_\mu} + \partial_\mu(\delta\chi)\frac{\partial L}{\partial \chi_\mu} = \partial_\mu\left(\xi^i \frac{\partial L}{\partial \chi^i_\mu}\right),$$

so backwards

$$\partial_\mu\left(\xi^i \frac{\partial L}{\partial \chi^i_\mu}\right) = \delta_\chi L.$$

Therefore, the current

$$-J^\mu := +\xi^i \frac{\partial L}{\partial \chi^i_\mu} - \hat{K}^\mu = +\xi^i \frac{\partial L}{\partial \chi^i_\mu} + K^\mu$$

is conserved:

$$\partial_\mu J^\mu = -(\delta_\chi L - \delta_\chi L) = 0.$$

So, indeed we've obtained the current for the duality transformations:

$$J^\mu = -\xi^i \frac{\partial L}{\partial \chi^i_\mu} + \frac{1}{2}\left(\tilde{G}^{\mu\nu}A\mathcal{A}_\nu - \tilde{F}^{\mu\nu}C\mathcal{A}_\nu + \tilde{G}^{\mu\nu}B\mathcal{B}_\nu - \tilde{F}^{\mu\nu}D\mathcal{B}_\nu\right).$$

Current is not gauge invariant

The current is not invariant under the gauge transformations

$$\mathcal{A}^a_\mu \to \mathcal{A}^a_\mu + \partial_\mu \alpha^a, \qquad \mathcal{B}^a_\mu \to \mathcal{B}^a_\mu + \partial_\mu \beta^a,$$

which leave invariant $F^a_{\mu\nu} := \partial_\mu \mathcal{A}^a_\nu - \partial_\nu \mathcal{A}^a_\mu$, and $G^a_{\mu\nu} := \partial_\mu \mathcal{B}^a_\nu - \partial_\nu \mathcal{B}^a_\mu$. It changes by the divergence of an antisymmetric tensor

$$J^\mu \to J^\mu + \frac{1}{2}\partial_\nu \left(\tilde{G}^{\mu\nu} A\alpha - \tilde{F}^{\mu\nu} C\alpha + \tilde{G}^{\mu\nu} B\beta - \tilde{F}^{\mu\nu} D\beta\right).$$

Therefore the corresponding charge $\int J^0 \; \mathrm{d}^3 x$ is gauge invariant and is actually the generator of the duality rotations, as one can see by constructing the potentials $\mathcal{A}_\mu$ and $\mathcal{B}_\mu$ in a special gauge.

A conserved current of this kind does not preclude [8, 9] the existence of massless spin-one states which couple to the charge operator, as we shall discuss in more detail in sect. 6.

(??? here is discussion of a charge from this current. Don't know an answer yet !?!?!?)

4.2.4 Invariance of the energy-momentum tensor

Theory

The energy-momentum tensor is obtained by

$$T^\mu_\nu := \frac{\partial L}{\partial\left(\partial_\mu \phi_\alpha\right)} \partial_\nu \phi_\alpha - \delta^\mu_\nu L,$$

(???? no, forgot the idea)

we will modify it a little (??? why do we need it???) and consider another energy-momentum tensor:

$$\theta^\mu{}_\lambda := -T^\mu{}_\lambda + \psi^\mu{}_\lambda = \qquad\qquad \psi^\mu{}_\lambda : \; \partial_\mu \psi^\mu{}_\lambda = 0; \; \psi^\mu{}_\lambda := \tilde{G}^{a\mu\nu} F^a_{\nu\lambda}.$$

$$= -\chi^i_\lambda \frac{\partial L}{\partial \chi^i_\mu} + \delta^\mu_\lambda L + \tilde{G}^{a\mu\nu} F^a_{\nu\lambda};$$

only because of this property we can add $\psi^\mu{}_\lambda$.

Indeed

$$\partial_\mu \psi^\mu{}_\lambda = \partial_\mu (\tilde{G}^{a\mu\nu} F^a_{\nu\lambda}) = \partial_\mu (\tilde{G}^{a\mu\nu}) F^a_{\nu\lambda} + \tilde{G}^{a\mu\nu} \partial_\mu (F^a_{\nu\lambda}) = 0 + G^{a\mu\nu} \partial_\mu (\tilde{F}^a_{\nu\lambda}) = 0,$$

because of the equations of motion $\partial^\mu \tilde{F}^a_{\mu\nu} = 0$ and $\partial^\mu \tilde{G}^a_{\mu\nu} = 0$ (???? no, I don't understand this!!!!).

The energy-momentum tensor is conserved as a consequence of the equations of motion (2.3), (2.5) and (2.15):

$$\partial_\mu \theta^\mu{}_\lambda = 0$$

We will show that $\delta L = \frac{1}{4}(FC\tilde{F} + GB\tilde{G})$ implies also that $\theta^\mu{}_\lambda$ is invariant under duality rotations. We must evaluate

$$\delta\theta^\mu_\lambda = -\delta\left(\frac{\partial L}{\partial \chi^i_\mu}\right)\chi^i_\lambda - \frac{\partial L}{\partial \chi^i_\mu}\delta\chi^i_\lambda + \delta^\mu_\lambda \delta L + \delta\tilde{G}^{\mu\nu} F_{\nu\lambda} + \tilde{G}^{\mu\nu}\delta F_{\nu\lambda}$$

Now, from (2.7),

$$\delta\frac{\partial L}{\partial \chi^i_\mu} = \frac{\partial}{\partial \chi^i_\mu}\delta L - \frac{\partial \xi^j}{\partial \chi^i}\frac{\partial L}{\partial \chi^j_\mu} - \frac{1}{2}\frac{\partial G^{\rho\nu}}{\partial \chi^i_\mu} B\tilde{G}_{\rho\nu}.$$

Therefore

$$\delta \frac{\partial L}{\partial \chi^i_\mu} \chi^i_\lambda + \frac{\partial L}{\partial \chi^i_\mu} \delta \chi^i_\lambda = \chi^i_\lambda \frac{\partial}{\partial \chi^i_\mu} \left(\delta L - \frac{1}{4} G B \tilde{G} \right)$$

In order to evaluate the other terms of (2.35) we take δF from (2.6a) and $\delta \tilde{G}$ from (2.6a) or (2.9). Using also (2.19), we obtain

$$\delta \tilde{G}^{\mu\nu} F_{\nu\lambda} + \tilde{G}^{\mu\nu} \delta F_{\nu\lambda} = \tilde{F}^{\mu\nu} C F_{\nu\lambda} + \tilde{G}^{\mu\nu} B G_{\nu\lambda}.$$

This expression can be simplified if one remembers that, for any two antisymmetric tensors $F_{1\mu\nu}$ and $F_{2\mu\nu}$,

$$\tilde{F}^{\mu\nu}_{l'} F_{2\nu\lambda} + \tilde{F}^{\mu\nu}_2 F_{1\nu\lambda} = -\frac{1}{2} \delta^\mu_\lambda \tilde{F}^{\rho\sigma}_1 F_{2\rho\sigma}.$$

Since the matrices C and B are symmetric we can use (2.39) in (2.38). Combining the result with (2.37) we obtain finally

$$\delta \theta^\mu_\lambda = \left(-\chi^i_\lambda \frac{\partial}{\partial \chi^i_\mu} + \delta^\mu_\lambda \right) \left(\delta L - \frac{1}{4} \tilde{F} C F - \frac{1}{4} \tilde{G} B G \right)$$

since the field F does not depend on χ_μ. This equation shows that our basic relation (2.20) implies

$$\delta \theta^\mu_\lambda = 0.$$

The invariance of the symmetric energy-momentum tensor will be shown in appendix B. Since the hamiltonian and the equations of motion are invariant under duality rotations, it follows that the S-matrix is also invariant.

Идея для вывода тока и ТЭИ (тут суть того, что дальше будем делать. это же во многих предложениях перед формулами.)

4.2.5 Construction of the lagrangian

(это не актуально, пока не смотрю)

Theory

We start by noting the identity

$$\delta L = \frac{1}{4} \delta (F \tilde{G}),$$

which follows from the comparison of the transformation (2.6a) subject to the constraints (2.11) and (2.19), with the property (2.20). As a consequence of (2.42) the lagrangian can be written in the form

$$L = \frac{1}{4} F \tilde{G} + L_{\text{inv}} .$$

Under the most general group of duality transformations, namely $\mathrm{Sp}(2n, R)$, the only * invariant which can be constructed from vectors in the fundamental $2n$-dimensional representation is an antisymmetric bilinear. If we introduce two antisymmetric Lorentz-tensor functions of the fields χ^i, $(H_{\mu\nu}(\chi), I_{\mu\nu}(\chi))$, which transform under (2.6) like $(F_{\mu\nu}, G_{\mu\nu})$, then the lagrangian

$$L = \frac{1}{4} F\tilde{G} + \frac{1}{4}(FI - GH) + L_{\text{inv}}(\chi)$$

where $L_{\text{inv}}(\chi)$ is an invariant function of the χ (and their derivatives), has the property that the equations of motion are invariant. The definition (2.4) of $\tilde{G}$,

$$\frac{\partial L}{\partial F} = \frac{1}{2}\tilde{G} = \frac{1}{4}\tilde{G} + \frac{1}{4}F\frac{\partial \tilde{G}}{\partial F} + \frac{1}{4}I + \frac{1}{4}\frac{\partial \tilde{G}}{\partial F}\tilde{H}$$

where we have used $F_1F_2 = -\tilde{F}_1\tilde{F}_2$, gives a differential equation for $\tilde{G}$:

$$\tilde{G} - I = (F + \tilde{H})\frac{\partial \tilde{G}}{\partial F}$$

At this point it is convenient to introduce the operator j which changes an antisymmetric tensor into its dual

$$jT_{\mu\nu} = \tilde{T}_{\mu\nu}, \quad (j)^2 = -1$$

and to rewrite (2.46) in the form

$$jG - I = (F + jH)\frac{\partial jG}{\partial F}$$

If the true invariance group is a subgroup of $\text{Sp}(2n, \text{R})$, other invariants may exist.

The similarity of the operator j with the usual imaginary unit i has been noted and used before [1, 7]. The general solution of (2.48) is

$$jG - I = -K(\chi)(F + jH)$$

where $K(\chi)$ is a priori an arbitrary $n \times n$ symmetric matrix function of the χ and can contain j. To verify that (2.49) is a solution of (2.48), observe that the definition (2.4) implies that

$$\frac{\partial \tilde{G}^a_{\mu\nu}}{\partial F^b_{\rho\sigma}} = \frac{\partial \tilde{G}^b_{\rho\sigma}}{\partial F^a_{\mu\nu}}$$

The transformation law of K under (2.6) is determined by those of (F, G) and (H, I). Varying both sides of (2.49) one finds

$$\delta K(x) = -jC - jKBK + DK - KA$$

where the matrices A, B, C, D satisfy (2.11) and (2.19). In addition, it is clear that the form of the kinetic energy term for the vector fields requires

$$K(\chi) = 1 + f(\chi)$$

with $f(0) = 0$. As we shall see in sect. 5, these properties permit the determination of the matrix K as a function of the fields χ. Substituting (2.49) in (2.44) we find the lagrangian

$$L = -\frac{1}{4}FKF + \frac{1}{2}F(I - jKH) + \frac{1}{4}jH(I - jKH) + L_{\text{inv}}(\chi).$$

4.2.6 Compact duality rotations

(это не актуально, пока не смотрю)

Theory

As an example, let us consider the case where $f(\chi) = 0$ in (2.52), i.e. $K(\chi) = 1$ in (2.49). Then, since $\delta K = 0$, the transformation law (2.51) restricts the matrices further:

$$B = -C = B^{\mathrm{T}}, \quad A = D = -A^{\mathrm{T}}.$$

The subgroup of $\mathrm{Sp}(2n, \mathbf{R})$ determined by (3.1) is just its maximal compact subgroup $\mathrm{U}(n)$. This becomes obvious if one writes (2.6a) in the complex basis given by the combinations $F \pm iG$. One finds

$$\delta \begin{pmatrix} F + iG \\ F - iG \end{pmatrix} = \begin{pmatrix} T & 0 \\ 0 & T^* \end{pmatrix} \begin{pmatrix} F + iG \\ F - iG \end{pmatrix}$$

where

$$T = A - iB = -T^\dagger$$

is an antihermitian $n \times n$ matrix. Whether or not the invariance group is actually $\mathrm{U}(n)$ or some (compact) subgroup thereof, one sees that the tensors $F + iG$ and $F - iG$ transform according to an n-dimensional representation (which may be reducible) and to its conjugate.

The complex basis defined in (3.2) allows a simple physical interpretation: spin-one states of opposite helicity transform according to conjugate representations of the duality group, just as massless fermions do under chiral transformations. However, the equations become more concise and are more easily generalized to the non-compact case if we choose a similar basis with $i \to j$, where j is the operator defined in eqs. (2.47), i.e.

$$F \pm iG \to F \pm jG$$
$$\mathrm{Re}\, T \pm i \,\mathrm{Im}\, T \to \mathrm{Re}\, T \pm j \,\mathrm{Im}\, T$$

in eq. (3.2). Then if we introduce antisymmetric tensor functions of χ,

$$jH_\pm \equiv (H \pm jI) = \pm j(I \mp jH)$$

transforming like $F \pm jG$, the lagrangian (2.53) with $K = 1$ becomes

$$L = -\frac{1}{4}F^2 + \frac{1}{2}FH_+ - \frac{1}{8}H_+^2 - \frac{1}{8}H_+H_- + L_{\mathrm{inv}}(\chi).$$

Clearly, without loss of generality, we may set $H_- = 0$: since this field does not couple to F, its couplings must be themselves invariant (and indeed $H_+H_-$is an invariant) and can be absorbed in $L_{\mathrm{inv}}\,(\chi)$. Then (H, I) reduces to $(H, -jH)$, and we see that the correct transformation properties for L are obtained in the compact case by introducing a tensor $H = \frac{1}{2}jH_+$which transforms according to the same j-complex representation of the unitary group as does the field $F + jG$. In practice H is constructed from fermion fields ψ, e.g.

$$H^a_{\mu\nu} = C^a_{ij}\bar{\psi}^i\sigma_{\mu\nu}\psi^j$$

and the transformation (2.6b) is chiral. Finally we remark that by setting $H_- = 0$ we can rewrite the invariant

$$FI - GH = \frac{1}{2}\left(F^2 + G^2\right) = \frac{1}{2}(F - iG)(F + iG),$$

which is manifestly invariant under linear unitary transformations among F and G. The models of ref. [3] are examples of the developments of this subsection.

In order to generalize the above construction to the case of non-compact duality transformations without introducing ghosts, we must add scalar fields in the way recalled in the following section.

4.2.7 Non-linear realizations

(это не актуально, пока не смотрю)

Theory

In this section we recall the well-known description of non-linear models for scalars valued in the quotient space of a group by a subgroup [10,11]. Although the formalism is very general, and can be applied to a compact group as well, we have in mind the special case when the semisimple group $\mathcal{G}$ is non-compact but the subgroup $\mathcal{K}$ is its maximal compact subgroup. As pointed out in refs. [12-14], this case leads to a non-linear scalar lagrangian without ghosts.

The scalars can be described by a group element $g(x) \in \mathcal{G}$ (in some representation) and are therefore at first as many as the parameters of $\mathcal{G}$. However, we consider as equivalent all group elements which differ by multiplication (on the right) by an element of $\mathscr{K}$. This reduces the scalars to as many as the parameters of the quotient space (coset space). To implement the equivalence we require the lagrangian to be invariant under a gauge transformation which we choose to write as

$$g(x) \to g(x)[k(x)]^{-1}$$

We require it also to be invariant under the rigid transformation

$$g(x) \to g_0 g(x)$$

with $g_0 \in \mathcal{G}$.

In order to construct the lagrangian, we introduce a gauge field Q_μ which belongs to the Lie algebra of $\mathscr{H}$ and which transforms as

$$Q_\mu \to k Q_\mu k^{-1} - \partial_\mu k k^{-1}$$

under the gauge transformation (4.1). With it one can construct a covariant derivative

$$D_\mu g = \partial_\mu g - g Q_\mu$$

which transforms as

$$D_\mu g \to (D_\mu g)\,[k(x)]^{-1}$$

so that

$$g^{-1} D_\mu g \to k g^{-1} D_\mu g k^{-1}$$

Under (4.2), $g^{-1} D_\mu g$ is invariant. Therefore the lagrangian

$$L = -\frac{1}{2} \operatorname{Tr} \left(g^{-1} D_\mu g\right)^2$$

is invariant under both (4.1) and (4.2). Here Tr can be defined as the trace in some representation, suitably normalized. The field Q_μ enters the lagrangian without derivatives. If one varies Q_μ keeping g fixed, one obtains the equation of motion

$$\delta L = \operatorname{Tr} \delta Q_\mu \left(g^{-1} \partial_\mu g - Q_\mu\right) = 0$$

Since δQ_μ, like Q_μ, is in the Lie algebra of $\mathscr{K}$, this means that $P_\mu \equiv g^{-1}\partial_\mu g - Q_\mu$ is in the orthogonal complement. The element $g^{-1}\partial_\mu g$ of the Lie algebra of $\mathcal{G}$ is decomposed as

$$g^{-1}\partial_\mu g = Q_\mu + P_\mu$$

and Q_μ is its part in the Lie algebra of $\mathscr{K}$. This equation can be thought of as giving Q_μ and P_μ in terms of g and its derivatives. One can rewrite (4.9) as

$$P_\mu = g^{-1}D_\mu g$$

and give (4.7) the form

$$L = -\frac{1}{2}\,\mathrm{Tr}\,P_\mu^2$$

Substitution into (4.7) of Q_μ expressed in terms of g from (4.9) gives a lagrangian which depends only on g and its derivatives. Clearly it is given by (4.11) in which P_μ is expressed in terms of g from (4.9).

To obtain the equations of motion for the scalar fields one must vary g in the lagrangian keeping Q_μ fixed. With a little algebra and up to a total derivative one finds

$$\delta L = \mathrm{Tr}\,\delta g g^{-1}\partial_\mu\left(D_\mu g g^{-1}\right) = 0$$

Now, $\delta g g^{-1}$ is an arbitrary element of the Lie algebra of $\mathcal{G}$; therefore, one has the equations of motion

$$\partial_\mu\left(D_\mu g g^{-1}\right) = 0$$

or

$$\partial_\mu\left(g P_\mu g^{-1}\right) = 0$$

These equations can also be written as

$$D_\mu P_\mu \equiv \partial_\mu P_\mu - [P_\mu, Q_\mu] = 0$$

One could have obtained these equations of motion also by varying g in the lagrangian (4.11) where Q_μ no longer appears as an independent field.

Our lagrangian is invariant under (4.1) and (4.2). In particular it is invariant under the rigid transformation

$$g \to g k^{-1},$$

with k independent of x. According to a general argument due to Emmy Noether, the current corresponding to (4.16) vanishes identically as a consequence of the gauge invariance (4.1). The current corresponding to (4.2) does not vanish, however, and the current corresponding to the subtransformation

$$g \to k g$$

is the same as that for

$$g \to k g k^{-1}.$$

According to the Noether prescription, we construct the currents by replacing in the lagrangian the derivatives of the scalar fields by the variation of the same fields under the corresponding transformations. This amounts to varying the lagrangian (4.11),

$$\delta L = -\operatorname{Tr} P_\mu \delta P_\mu$$

and replacing δP_μ by δP (and δQ_μ by δQ when it appears in the lagrangian, as below) defined from

$$g^{-1}\delta g = \delta Q + \delta P$$

with the same decomposition of the Lie algebra. So the current is

$$J_\mu = -\operatorname{Tr} P_\mu \delta P$$

where δ is the infinitesimal transformation in question. This equation can also be written

$$\begin{aligned} J_\mu &= -\operatorname{Tr} P_\mu(\delta Q + \delta P) \\ &= -\operatorname{Tr} P_\mu g^{-1}\delta g \end{aligned}$$

since the term we have added vanishes by orthogonality. For (4.2) let

$$\delta g = (q+p)g$$

where q and p are infinitesimals in the Lie algebra of $\mathscr{K}$ and perpendicular to it, respectively. We find

$$J_\mu = -\operatorname{Tr} g P_\mu g^{-1}(q+p)$$

So the current of (4.2) is given by the operator

$$J_\mu = -g P_\mu g^{-1}$$

which is in the Lie algebra of $\mathcal{G}$. Its part in the Lie algebra of $\mathscr{K}$ gives the current of (4.17) and also of (4.18). The equations of motion, in the form (4.14), state just the conservation of the current in agreement with Noether's theorem. We can also check directly that the current of (4.16) vanishes identically. Indeed we must now use (4.20) and (4.21) with

$$\delta g = -gq$$

This gives

$$\delta Q = -q, \quad \delta P = 0.$$

Coupling to other fields, which we denote by ψ, can be introduced as follows. Let ψ be invariant under (4.2) and let it belong to some unitary representation of $\mathcal{H}$ so that it transforms under (4.1) as

$$\psi(x) \to k(x)\psi(x).$$

It is easy to construct lagrangians invariant under (4.1), (4.28) by means of the covariant derivative

$$D_\mu\psi = \partial_\mu\psi + Q_\mu\psi$$

where $k(x)$ and Q_μ are matrices in the appropriate representation. For instance, the kinetic term of the Dirac lagrangian would be

$$-\frac{1}{2}i\bar{\psi}\gamma^\mu\left(\vec{D}_\mu - \overleftarrow{D_\mu}\right)\psi$$

One can also have derivative coupling interaction terms such as

$$\bar{\psi}_1 \gamma^\mu P_\mu \psi_2$$

where P_μ is now a matrix in the appropriate representation of the Lie algebra of $\mathscr{K}$.

In presence of other fields, the derivatives of the scalars enter also in terms like (4.30) and (4.31) through Q_μ and P_μ. Therefore, the current receives additional contributions. For instance (4.30), which contains

$$-i\bar{\psi}\gamma^\mu Q_\mu \psi$$

gives, from (4.20) and (4.23),

$$J'_\mu = -i\bar{\psi}\gamma_\mu \delta Q \psi = -i\bar{\psi}\gamma_\mu q \psi + \cdots$$

where the dots represent terms of higher order in the scalar fields. Similarly, (4.31) gives

$$J''_\mu = \bar{\psi}_1 \gamma_\mu P \psi_2 + \cdots$$

We have assumed that the fields ψ are invariant under (4.2) because the group $\mathcal{G}$ is non-compact: if we had attributed them to a linear representation of $\mathcal{G}$ we would have obtained a lagrangian with ghosts.

The gauge field Q_μ in (4.29) can be taken as given by (4.9), a function of the scalar fields and their derivatives. On the other hand, Q_μ could be introduced as an independent field both in the scalar lagrangian and in the lagrangian of the other fields ψ. The field Q_μ can still be determined from its own equation of motion but it has now additional terms which are functions of the fields ψ.

It is often convenient to use the gauge equivalence (4.1) to choose a special gauge. For instance, one can always transform any group element g by (4.1) to the form

$$g^{\cdot} = \mathrm{e}^{p} = gk^{-1},$$

where P is an element of the Lie algebra of $\mathcal{G}$ perpendicular to those in the Lie algebra of $\mathcal{K}$. The scalars are now described by $P(x)$, the coset space $\mathcal{G}/\mathcal{K}$ is parametrized by P and g^* represents the equivalence class. A transformation (4.2) is now represented by

$$(g_0 g)^{\cdot} = \left(g_0 \mathrm{e}^P\right)^{\cdot} = \mathrm{e}^{P'} = g_0 \mathrm{e}^P k^{-1},$$

where $P'(g_0, P)$ and $k(g_0, P)$ are functions of the variables indicated and are determined by the group structure. The transformation

$$P \to P'(g_0, P)$$

is a non-linear realization of (4.2). Correspondingly one must transform the other fields as

$$\psi \to k(g_0, P)\,\psi$$

in order to maintain the special gauge. Similarly Q_μ transforms as in (4.3), but with the above $k(g_0, P)$. Clearly the lagrangian is invariant under the non-linear transformations (4.37) and (4.38), while the gauge invariance (4.1) is no longer apparent; it has been used to establish and maintain the special gauge.

4.2.8 Non-compact duality transformations

Theory

In sect. 2 we derived general formulae in terms of the vector field strength F and the remaining, unspecified fields χ. Here we wish to study in more detail the case of non-compact groups, requiring the introducing of scalar fields which are valued in the coset space $\mathcal{G}/\mathcal{K}$ as described in sect. 4. We saw that the most general non-compact group of dual rotations for n vectors is $\mathcal{G} = \mathrm{Sp}(2n, \mathrm{R})$ with $\mathfrak{K} = \mathrm{U}(n)$ as its maximal compact subgroup. In this case the scalar fields can be represented by an $\mathrm{Sp}(2n, \mathbf{R})$ matrix which is most easily expressed in the complex basis used in eq. (3.2) (see appendix A) as

$$g = \begin{pmatrix} \phi_0 & \phi_1^* \\ \phi_1 & \phi_0^* \end{pmatrix}$$

where ϕ_0 and ϕ_1 are complex $n \times n$ matrices which satisfy the constraint

$$\phi_0^\dagger \phi_0 - \phi_1^\dagger \phi_1 = 1$$

The transformation law of the scalar fields under $\mathrm{Sp}(2n, \mathbf{R})$ is

$$\delta g = \begin{pmatrix} T & V^* \\ V & T^* \end{pmatrix} g$$

where

$$T = -T^\dagger = M - iN,$$
$$V = V^{\mathrm{T}} = R - iS,$$

and the real matrices M, N, R, S are related to those of eq. (2.6a) by

$$A = M + R, \quad B = S + N, \quad C = S - N, \quad D = M - R.$$

Then eq. (2.51) for the transformation law of the matrix K can be written in the form

$$\delta K = [M, K] - \{R, K\} - jK(S + N)K - j(S - N)$$

It is easy to verify that this equation is solved by

$$K = \left(\phi_0^\dagger + \phi_1^\dagger\right)^{-1} \left(\phi_0^\dagger - \phi_1^\dagger\right)$$

where complex numbers are to be interpreted as having i replaced by the operator j defined in eq. (2.47), so that, for instance

$$\phi_i = \mathrm{Re}\, \phi_i + j\, \mathrm{Im}\, \phi_i$$

This trick [1, 7] allows one to write the equations for K in a very compact form.

Next we introduce couplings to fermions fields ψ in a way which is a straightforward generalization of the compact case. As discussed in sect. 4, ψ belongs to some representation of the compact gauge group $\mathcal{K}$. We form an antisymmetric Lorentz tensor $H^0_{\mu\nu}(\psi)$ which transforms under the gauge group $\mathscr{K}$ in the same way as $F + iG$ (or, in terms of the transformation matrices defined with $i \to j, F + jG$) transforms under the subgroup of the rigid non-compact symmetry which is isomorphic to $\mathscr{K}$. Then the tensors H and I entering eq. (2.44) may, without loss of generality, be constructed as:

$$\begin{pmatrix} H + jI \\ H - jI \end{pmatrix} = g \begin{pmatrix} jH^0(\psi) \\ 0 \end{pmatrix}$$

H and I are invariant under the gauge transformation of $\mathscr{K}$. We could, of course, replace the zero in (5.8) by a tensor H' which transforms according to the representation of $\mathscr{K}$ conjugate to that of H^0. However, as in the compact case described in sect. 3, H' does not couple to F so that its couplings to the remaining fields must be themselves invariant and can be absorbed in $L_{\text{inv}}(\chi)$. We have verified explicitly this invariance. Eq. (5.8) gives two equations for H and I in terms of the scalar and fermion fields:

$$I - jKH = \left(\phi_0^\dagger + \phi_1^\dagger\right)^{-1} H^0(\psi)$$
$$I - jK^* H = 0$$

Once one has specified $H^0(\psi)$ and $L_{\text{inv}}(\psi, \phi)$, the lagrangian is uniquely constructed as given by eqs. (2.53), (5.6) and (5.9). The compact case of sect. 3 is recovered by setting $\phi_0 = 1$ and $\phi_1 = 0$.

We would like to observe that the solution (2.49) of the differential equation (2.48) must be invariant under $\text{Sp}(2n, \text{R})$ transformations. Since H and I transform, respectively, in the same way as F and G, an invariant linear relation among them can be obtained by writing

$$g^{-1}\left[\begin{pmatrix} F + jG \\ F - jG \end{pmatrix} + \lambda j \begin{pmatrix} H + jI \\ H - jI \end{pmatrix}\right] = \begin{pmatrix} 0 \\ \cdots \end{pmatrix}$$

where

$$g^{-1} = \begin{pmatrix} \phi_0^\dagger & -\phi_1^\dagger \\ -\phi_1^{\text{T}} & \phi_0^{\text{T}} \end{pmatrix}$$

is the inverse of (5.1). The coefficient λ can be determined by considering the limit $K \to 1, \phi_0 \to 1, \phi_1 \to 0$, which gives $\lambda = 1$. Then one finds that in the general case, the relation given by the upper component of (5.10) is identical with (2.49) when K is given by eq. (5.6). In fact, this is how the expression for K was found. K is invariant under gauge transformations of the local $\text{U}(n)$ which operates on the matrix g from the right. It can also be expressed in the form

$$K = \frac{1 - Z^*}{1 + Z^*},$$

where

$$Z = \phi_1 \left(\phi_0\right)^{-1} = Z^T$$

In the special gauge in which

$$g = \begin{pmatrix} \phi_0 & \phi_1^* \\ \phi_1 & \phi_0^* \end{pmatrix} = \exp\begin{pmatrix} 0 & P^* \\ P & 0 \end{pmatrix}$$

the symmetry of the non-compact generator matrix P implies that

$$\phi_0 = \phi_0^\ddagger, \quad \phi_1 = \phi_1^{\text{T}}$$

This construction can be applied to the $N = 8$ supergravity lagrangian of Cremmer and Julia, except that the specific expressions they obtained for $L_{\text{inv}}(\chi)$ and $H^0_{\mu\nu}(\psi)$ were determined by the additional constraints of supersymmetry. (The inclusion of the gravitational field presents no difficulty, since it is invariant under duality transformations.) The prescription given in subsect. 2.2 and sect. 4 for obtaining the conserved currents can be directly applied for the subgroups SU(8) and even E_7 using their lagrangian.

The conserved currents associated with the non-compact generators can be constructed only when the interaction with the scalar fields is present. To understand what happens in the limit in which the scalar coupling constant tends to zero, let us recall that this coupling constant κ has the dimension of a reciprocal mass (in supergravity κ is the gravitational constant itself). In the special gauge of sect. 4, the coupling constant κ can be introduced explicitly in the lagrangians by rescaling $P \to \kappa P$ after which the lagrangian (4.11) must be divided by κ^2. The currents associated with the non-compact generators become

$$J_\mu = -\frac{1}{\kappa}\partial_\mu P + \cdots ,$$

where the dots denote terms containing the scalar fields which vanish with κ and terms containing the vector field strengths and other fields which have a finite limit as $\kappa \to 0$. To go to the limit we multiply (5.16) by κ and obtain

$$\lim_{\kappa\to 0} \kappa J_\mu = -\partial_\mu P.$$

All terms involving fields other than the scalars drop out, the scalars become free massless fields and the non-compact part of the group becomes abelian, corresponding to a contraction of the original non-compact group.

4.2.9 Invariance of a special derivative of a lagrangian

We have emphasized in the text that the lagrangian is not invariant under duality rotations; rather it transforms as given in (2.20). Here we will show that a suitably defined derivative of the lagrangian with respect to an invariant parameter is invariant. The invariant parameter could be a coupling constant or it could represent an invariant external field like the gravitational field in the matter lagrangian. An important consequence of this result is that the symmetric energy-momentum tensor is invariant under duality rotations, since it is obtained by taking the functional derivative of the matter action with respect to the gravitational field.

Let us assume that the lagrangian depends upon an invariant parameter λ. If ξ^i is independent of λ, differentiating (2.7) with respect to λ we obtain

$$\frac{\partial}{\partial\lambda}\delta L = \delta\frac{\partial L}{\partial\lambda} + \frac{\partial G}{\partial\lambda}B^{\mathrm{T}}\frac{\partial L}{\partial F}$$

Using (2.4) and (2.11), this equation can be written

$$\delta\frac{\partial L}{\partial\lambda} = \frac{\partial}{\partial\lambda}\left(\delta L - \frac{1}{4}GB\tilde{G} - \frac{1}{4}FC\tilde{F}\right).$$

We are free to add the last term in the right-hand side since the field F is independent of λ. We see that, if L satisfies our basic equation (2.20), then

$$\delta\frac{\partial L}{\partial\lambda} = 0$$

This observation provides a method for constructing invariant lagrangians, by switching on couplings in an invariant way, or for checking if a lagrangian is invariant.

In deriving (B.1) we assumed that the ξ^i are independent of the parameter λ. When the transformation is non-linear, if it is expressed in terms of scalar fields ϕ with canonical dimensions, the ξ^i will in general have an explicit dependence on some dimensional coupling constant κ. If we wish to apply the above argument to the parameter κ, we must first re-express the lagrangian and the transformation laws in terms of the dimensionless fields $\kappa\phi$. The partial derivative $\partial L/\partial\kappa$, taken keeping the fields $\kappa\phi$ constant instead of ϕ, is invariant.

We further wish to point out that the formalism developed in this paper applies directly to the case in which the fields F and χ of sect. 2 interact with an external gravitational field, described by a tensor $g_{\mu\nu}$ or by a vierbein $e_\mu{}^a$. Now we must distinguish between tensors (denoted by ordinary letters) and densities (denoted by script letters). The field $F_{\mu\nu}$ is an antisymmetric tensor with lower indices as in (2.2). The lagrangian is a density $\mathcal{E}$ and (2.4) must be written more precisely as

$$\tilde{\mathscr{G}}^{\mu\nu} = 2\frac{\partial \mathcal{L}}{\partial F_{\mu\nu}}$$

If we use the tensor

$$G_{\mu\nu} = -\frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}\tilde{\mathcal{G}}^{\rho\sigma}$$

and the density

$$\tilde{\mathscr{F}}^{\mu\nu} = \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}F_{\rho\sigma}$$

we can write (2.20) as

$$\delta\mathfrak{E} = \frac{1}{4}\left(F_{\mu\nu}C^{\tilde{\mathcal{G}}^{\mu\nu}} + G_{\mu\nu}B\tilde{\mathscr{G}}^{\mu\nu}\right).$$

Eq. (2.6a) remains correct as written, using tensors with lower indices. Remember that the numerical antisymmetric symbol $\varepsilon^{\mu\nu\rho\sigma}$ transforms a tensor into a density, while $\varepsilon_{\mu\nu\rho\sigma}$ transforms a density into a tensor.

As already emphasized, the gravitational field does not transform under the duality rotations discussed in this paper. We have not considered here those duality rotations which act on the gravitational field, because they are an invariance only of the linearized gravitational equations.

4.2.10 Conclusion and examples

Conclusion by GZ

We have shown how a general theory invariant under dual rotations may be constructed. These theories still have a considerable degree of arbitrariness, namely in the choice of the tensor $H^0_{\mu\nu}(\psi)$ and of $L_{\text{inv}}(\chi)$ (see sects. 4, 5). In supergravity theories these quantities, and in fact the field content itself, are fixed by supersymmetry. It appears that the duality invariance of supergravity theories is implied by supersymmetry, a fact which still remains very mysterious.

There has recently been an attempt [15, 16] to connect $N = 8$ supergravity with gauge theory phenomenology by assuming that the fields of grand unified theories are composites of the fundamental fields of supergravity. This means in particular that the gauge bosons should be zero-mass bound states. There are arguments which show that, in a theory where there is a vector current, no zero-mass spin-one state can exist which carries the associated charge [8]. A related argument [9] shows that a vector current operator applied to the vacuum state cannot create a massless spin-one state. In $N = 8$ supergravity, these arguments would forbid the existence of massless states associated with the composite SU(8) gauge fields Q_μ, since, as we have seen, a current associated with the SU(8) generators can indeed be constructed, and is even conserved. However, as we have pointed out, this current is not gauge invariant, and therefore is not a true Lorentz vector. To construct a current operator one has to choose a particular gauge, in which case the Lorentz transformation properties become rather complicated, and the above-mentioned arguments can no longer be applied. Supergravity escapes this difficulty in much the same way as do

Yang-Mills theories where one can define a conserved vector current carrying the same quantum numbers as the Yang-Mills field, but which is not gauge invariant.

Finally, we wish to clear up some confusion concerning the composite vector fields in the $N = 8$ supergravity theory. In the manifestly gauge-invariant formulation, the conserved current J_μ was constructed to be invariant under SU(8) gauge transformations, and to transform linearly under non-compact E_7. We emphasize that J_μ contains the vector potentials $\mathscr{A}_\mu$ and $\mathscr{B}_\mu$, and is therefore not invariant under their abelian gauge transformation; without these terms it would not integrate to the correct generator of E_7. On the other hand the composite gauge potential Q_μ was constructed to be invariant under E_7, and to transform like a gauge field under local SU(8); it does not contain $\mathscr{A}_\mu$ and $\mathscr{B}_\mu$. A source of confusion could be the fact that in the special gauge of sects. 4 and 5 these two operators appear similar at the bilinear level in the scalar fields and even in the spinor fields if these are included in Q_μ. However, even in the special gauge the higher non-linear terms are different. In this gauge the non-compact part of E_7 is realized non-linearly, while its SU(8) subgroup is realized linearly. The SU(8) part of J_μ gives a conserved SU(8) current in terms of the physical fields of the special gauge.

An unanswered question is the precise definition of the vector operator for which the spectral function may develop a zero mass pole in the spin-one channel. Cremmer and Julia [7] took this to be Q_μ as defined from the scalar fields alone. As discussed in sect. 4, other definitions, including fermion fields, are possible. Presumably the correct combination is one which belongs to an irreducible supermultiplet. The issue we believe to have clarified in this paper is simply that there is no inconsistency in the existence of zero-mass bound states which transform like gauge bosons under local SU(8) and therefore couple to the conserved current associated with the rigid SU(8) transformations defined in eq. (4.18).

Examples of applications of GZ methods (???)

(question: how is it applied???)

Part III
Questions and Problems in Duality

4.3 Typical questions and problems

(!!!! I want to create a lot of illustrative examples!!! but for now it is too time consuming task. Maybe later I'll collect them!)

(???? is there at least one example of dual currents and dual TEM????? I don't know it)

4.3.1 Questions about understanding duality

4.3.2 Problems about duality in electrodynamics

Solve some problem in ED with duality (?????)

(подумаю потом, интересно было бы так сделать!! пока доучиваю известную теорию.)

4.3.3 Problems about duality in general field theories

Duality in $\mathcal{N}=2$ supergravity

The example is (a part of) a theory in $\mathcal{N}=2$ supergravity with one vector multiplet. [4] Hence there are two vectors, one from the gravity multiplet and one from the vector multiplet, but they will be mixed. We thus use $A=1,2$. We will neglect the frame field and the gravitinos, but still consider the complex scalar z, and two fermions $\chi^i, i=1,2$. The Lagrangian is

$$\begin{aligned}\mathcal{L} =&\frac{3}{(z-\bar{z})^2}\left[\partial_\mu z\partial^\mu\bar{z}+\frac{1}{2}\left(\bar{\chi}^1 P_L \partial\!\!\!/ \chi^1+\bar{\chi}^2 P_L \partial\!\!\!/ \chi^2\right)\right]\\ &+\left\{-\frac{1}{4}\mathrm{i}\left[\frac{1}{2}z^2(3\bar{z}+z)F^{+1}_{\mu\nu}F^{+1\mu\nu}-3z(z+\bar{z})F^{+1}_{\mu\nu}F^{+2\mu\nu}+\frac{3}{2}(3z+\bar{z})F^{+2}_{\mu\nu}F^{+2\mu\nu}\right]\right.\\ &\left.-\frac{3}{8}\bar{\chi}^1 P_R\gamma^{\mu\nu}\chi^2\bar{y}\left(-zF^{+1}_{\mu\nu}+F^{+2}_{\mu\nu}\right)\right\}+\text{ h.c. }+\dots\end{aligned}$$

Here appears also a variable y. In terms of z only its modulus is determined by

$$|y|^2=(\mathrm{i}(z-\bar{z}))^{-3}.$$

Choosing the phase of y is in fact a choice of a phase symmetry, that could also act on the fermions, but we will fix this below by the form of the transformations of y. First look at the scalar part, which determines the isometries.

Obtain all the quantities of the main text, where the matrices in the duality transformations are in terms of the parameters for the isometries:

$$A=\begin{pmatrix}-\frac{3}{2}\theta^2 & -3\theta^3\\ \theta^1 & -\frac{1}{2}\theta^2\end{pmatrix},\quad B=\begin{pmatrix}0 & 0\\ 0 & -\frac{2}{3}\theta^3\end{pmatrix},\quad C=\begin{pmatrix}0 & 0\\ 0 & 6\theta^1\end{pmatrix}$$

The variable y transforms consistently with (A.26) as (here a phase has been chosen for y)

$$\delta_d\bar{y}=-3\left(\frac{1}{2}\theta^2+\bar{z}\theta^3\right)\bar{y}.$$

Since $P_L\chi^i$ is the supersymmetry transform of z under the two supersymmetries:

$$\delta_d z=k(z)\rightarrow\delta_d P_L\chi^i=(\partial_z k)\,P_L\chi^i,$$

and the complex conjugate for $\delta_d P_R\chi^i$. Are the... terms in (A.25) invariant, or what should we still add?

Part IV
Other topics

5 Duality Rotations in Nonlinear Electrodynamics and in Extended Supergravity by Aschieri, Ferrara, Zumino

Abstract

We review the general theory of duality rotations which, in four dimensions, exchange electric with magnetic fields. Necessary and sufficient conditions in order for a theory to have duality symmetry are established. A nontrivial example is Born-Infeld theory with n abelian gauge fields and with $Sp(2n, \mathbb{R})$ self-duality. We then review duality symmetry in supergravity theories. In the case of $N = 2$ supergravity duality rotations are in general not a symmetry of the theory but a key ingredient in order to formulate the theory itself. This is due to the beautiful relation between the geometry of special Kähler manifolds and duality rotations.

5.1 Introduction

It has long been known that the free Maxwell's equations are invariant under the rotation of the electric field into the magnetic fields; this is also the case if electric and magnetic charges are present. In 1935, Schrödinger [2] showed that the nonlinear electrodynamics of Born and Infeld [1], (then proposed as a new fundamental theory of the electromagnetic field and presently relevant in describing the low energy effective action of D-branes in open string theory), has also, quite remarkably, this property. Extended supergravity theories too, as first pointed out in [3, 5] exhibit electric-magnetic duality symmetry. Duality symmetry thus encompasses photons self-interactions, gravity interactions and couplings to spinors (of the magnetic moment type, not minimal couplings).

Shortly after [3, 4, 5] the general theory of duality invariance with abelian gauge fields coupled to fermionic and bosonic matter was developped in [6, 39]. Since then the duality symmetry of extended supergravity theories has been extensively investigated [8, 9, 10, 11], and examples of Born-Infeld type lagrangians with electric-magnetic duality have been presented, in the case of one abelian gauge field [12, 13, 14, 15, 16] and in the case of many abelian gauge fields [17, 18, 19, 20]. Their supersymmetric generalizations have been considered in [21, 22] and with different scalar couplings and noncompact duality group in [17, 18, 23, 24, 25].

We also mention that duality symmetry can be generalized to arbitrary even dimensions by using antisymmetric tensor fields such that the rank of their field strengths equals half the dimension of space-time, see [26, 27], and [30, 11, 31, 28, 16, 18, 24, 25].

We provide a rigorous formulation of the general theory of four-dimensional electric-magnetc duality in lagrangian field theories where many abelian vector fields are coupled to scalars, fermions and to gravity. When the scalar fields lagrangian is described by a non-linear sigma model with a symmetric space G/H where G is noncompact and H is its maximal compact subgroup, the coupling of the scalars with the vector fields is uniquely determined by a symplectic representation of G (i.e. where the representation space is equipped with an invariant antisymmetric product). Moreover fermions coupled to the sigma model, which lie in representations of H, must also be coupled to vectors through particular Pauli terms as implied by electric-magnetic duality.

This formalism is realized in an elegant way in extended supergravity theories in four dimensions and can be generalized to dyons [32] in D-dimensions, which exist when D is even and the dyon is a p-brane with $p = D/2 - 2$. In the context of superstring theory or M theory electric-magnetic dualities can arise from many sources, namely S-duality, T-duality or a combination thereof called U-duality [29]. From the point of view of a four dimensional observer such dualities manifest as some global symmetries of the lowest order Euler-Lagrange equations of the underlying four dimensional effective theory.

The study of the relations between the symmetries of higher dimensional theories and their realization in four dimension is rich and fruitful, and duality rotations are an essential ingredient. Seemingly different lagrangians with different elementary dynamical fields can be shown to describe equivalent equation of motions by using duality. An interesting example is provided by the $N = 8$, $D = 4$ supergravity lagrangian whose duality group is $G = E_{7,(7)}$, this is the formulation of Cremmer and Julia [5]. An alternative formulation obtained from dimensional reduction of the $D = 5$ supergravity, exhibits an action that is invariant under a different group of symmetries. These two theories can be related only after a proper duality rotation of electric and magentic fields which involves a suitable Legendre transformation (a duality rotation that is not a symmetry transformation).

Let us also recall that duality rotation symmetries can be further enhanced to local symmetries (gauging of duality groups). The corresponding gauged supergravities appear as string compactifications in the presence of fluxes and as generalized compactifications of (ungauged) higher dimensional supergravities.

As a main example consider again the $N = 8$, $D = 4$ supergravity lagrangian of Cremmer and Julia, it is invariant under $SO(8)$ (compact subgroup of $E_{7,(7)}$). The gauging of $SO(8)$ corresponds to the gauged $N = 8$ supergravity of De Witt and Nicolai [33]. As shown in [34] the gauging of a different subgroup, that is the natural choice in the equivalent formulation of the theory obtained from dimensional reduction of $D = 5$ supergravity, corresponds to the gauging of a flat group in the sense of Scherk and Schwarz dimensional reduction [35], and gives the massive deformation of the $N = 8$ supergravity as obtained by Cremmer, Scherk and Schwarz [36].

Electric-Magnetic duality is also the underlying symmetry which encompasses the physics of extremal black holes and of the "attractor mechanism" [37, 38, 39], for recent reviews on the attractor mechanism see [40, 41, 42]. Here the Bekenstein-Hawking entropy-area formula

$$S = \frac{1}{4}A$$

is directly derived by the evaluation of a certain black hole potential $\mathscr{V}_{BH}$ at its attractive critical points [43]

$$S = \pi\, \mathscr{V}_{BH}\big|_C$$

where the critical points C satisfy $\partial \mathscr{V}_{BH}|_C = 0$. The potential $\mathscr{V}_{BH}$ is a quadratic invariant of the duality group; it depends on both the matter and the gauge fields configuration. In all extended supersymmetries with $N > 2$, the entropy S can also be computed via a certain duality invariant combination of the magnetic and electric charges p, q of the fields configuration [44, 45]

$$S = \pi \mathscr{S}(p, q)\ .$$

In the remaining part of this introduction we present the structure of the paper by summarizing its different sections.

In Section 2 we give a pedagogical introduction to $U(1)$ duality rotations in nonlinear theories of electromagnetism. The basic aspects of duality symmetry are already present in this simple case with just one abelian gauge field: the hamiltonian is invariant (duality rotations are canonical transformations that commute with the hamiltonian); the lagrangian is not invariant but must transform in a well defined way. The Born-Infeld theory is the main example of duality invariant nonlinear theory.

In Section 3 the general theory is formulated with many abelian gauge fields interacting with bosonic and fermionic matter. Necessary and for the first time sufficient conditions in order for a theory to have duality symmetry are established. The maximal symmetry group in a theory with n abelian gauge fields includes $Sp(2n, \mathbb{R})$. If there are no scalar fields the maximal symmetry group is $U(n)$. The geometry of the symmetry transformations on the scalar fields is that of the coset space $Sp(2n, \mathbb{R})/U(n)$ that we study in detail. The kinetic term for the scalar fields is constructed by using this coset space geometry. In Subsection 3.6 we present the Born-Infeld lagrangian with n abelian gauge fields and $Sp(2n, \mathbb{R})$ duality symmetry [18]. The self-duality of this lagrangian is proven by studying another example: the Born-Infeld lagrangian with n complex gauge fields and $U(n, n)$ duality symmetry. Here $U(n, n)$ is the group of holomorphic duality rotations. We briefly develop the theory of holomorphic duality rotations.

The Born-Infeld lagrangian with $U(n, n)$ self-duality is per se interesting, the scalar fields span the coset space $\frac{U(n,n)}{U(n)\times U(n)}$, in the case $n = 3$ this is the coset space of the scalars of $N = 3$ supergravity with 3 vector multiplets. This Born-Infeld lagrangian is then a natural candidate for the nonlinear generalization of $N = 3$ supergravity.

We close this sections by presenting, in a formulation with auxiliary fields, the supersymmetric version of this Born-Infeld Lagrangian [17, 18]. We also present the form without auxiliary fields of the supersymmetric Born-Infeld Lagrangian with a single gauge field and a scalar field; this theory is invariant under $SL(2, \mathbb{R})$ duality, which reduces to $U(1)$ duality if the value of the scalar field is suitably fixed. Versions of this theory without the scalar field were presented in [46, 47, 48].

In Section 4 we apply the general theory of duality rotation to supergravity theories with $N > 2$ supersymmetries. In these supersymmetric theories the duality group is always a subgroup G of $Sp(2n, \mathbb{R})$, where G is the isometry group of the sigma model G/H of the scalar fields. Much of the geometry underlying these theories is in the (local) embedding of G in $Sp(2n, \mathbb{R})$. The supersymmetry transformation rules, the structure of the central and matter charges and the duality invariants associated to the entropy and the potential of extremal black holes configurations are all expressed in terms of the embedding of G in $Sp(2n, \mathbb{R})$ [11]. We thus present a unifying formalims. We also explicitly construct the symplectic bundles (vector bundles with a symplectic product on the fibers) associated to these theories, and prove that they are topologically trivial; this is no more the case for generic $N = 2$ supergravities.

In Section 5 we introduce special Kgeometry as studied in differential geometry, we follow in particular the work of Freed [49], see also [50] (and [51]) and then develop the mathematical definition up to the construction of those explicit flat symplectic sections used in $N = 2$ supergravity. We thus see for example that the flat symplectic bundle of a rigid special Kmanifold M is just the tangent bundle TM with symplectic product given by the Kform. A similar construction applies in the case of local special geometry (there the flat tangent bundle is not of the Kmanifold M but is essentially the tangent bundle of a complex line bundle $L \to M$). This clarifies the global aspects of special geometry and the key role played by duality rotations in the formulation of $N = 2$ supergravity with scalar fields taking value in the target space M. Duality rotations are needed for the theory to be globally well defined.

In Section 6 duality rotations in nonlinear electromagnetism are considered on a noncommutative spacetime, $[x^\mu, x^\nu] = i\Theta^{\mu\nu}$. The noncommutativity tensor $\Theta^{\mu\nu}$ must be light-like. A nontrivial example of nonlinear electrodynamics on commutative spacetime is presented and using Seiberg-Witten map between commutative and noncommutative gauge theories noncommutative $U(1)$ Yang Mills theory is shown to have duality symmetry. This theory formally is nonabelian, $\widehat{F}_{\mu\nu} = \partial_\mu \widehat{A}_\mu - \partial_\nu \widehat{A}_\mu - i[\widehat{A}_\mu, \widehat{A}_\nu]$, its self-duality is in this respect remarkable. One can also enhance the duality group to $Sp(2,\mathbb{R})$ and couple this noncommutative theory to axion, dilaton and Higgs fields, these latter via minimal couplings. Duality in noncommutative spacetime allows to relate space-noncommutative magnetic monopoles to space-noncommutative electric monopoles [52, 53].

A special kind of noncommutative spacetime is a lattice space (it can be studied with noncommutative geometry techniques). Duality rotations on a lattice have been studied in [54].

In Appendix 7 we prove some fundamental properties of the symplectic group $Sp(2n,\mathbb{R})$ and of the coset space $Sp(2n,\mathbb{R})/U(n)$. We also collect for reference some main formulae and definitions.

In Appendix 8 a symmetry property of the trace of a solution of a polynomial matrix equation is proven. This allows the explicit formulation of the Born-Infeld lagrangian with $Sp(2n,\mathbb{R})$ duality symmetry presented in Section 3.7.

5.2 $U(1)$ gauge theory and duality symmetry

Maxwell theory is the prototype of electric-magnetic duality invariant theories. In vacuum the equations of motion are

$$\begin{aligned} \partial_\mu F^{\mu\nu} &= 0 \ , \\ \partial_\mu \tilde{F}^{\mu\nu} &= 0 \ , \end{aligned} \tag{5.1}$$

where $\tilde{F}^{\mu\nu} \equiv \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}F_{\rho\sigma}$. They are invariant under rotations $\binom{F}{\tilde{F}} \mapsto \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}\binom{F}{\tilde{F}}$, or using vector notation under rotations $\binom{\boldsymbol{E}}{\boldsymbol{B}} \mapsto \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}\binom{\boldsymbol{E}}{\boldsymbol{B}}$. This rotational symmetry, called duality symmetry, and also duality invariance or self-duality, is reflected in the invariance of the hamiltonian $\boldsymbol{\mathcal{H}} = \frac{1}{2}(\boldsymbol{E}^2 + \boldsymbol{B}^2)$, notice however that the lagrangian $\boldsymbol{\mathcal{L}} = \frac{1}{2}(\boldsymbol{E}^2 - \boldsymbol{B}^2)$ is not invariant. This symmetry is not an internal symmetry because it rotates a tensor into a pseudotensor.

We study this symmetry for more general electromagnetic theories. In this section and the next one conditions on the lagrangians of (nonlinear) elecromagnetic theories will be found that guarantee the duality symmetry (self-duality) of the equations of motion.

The key mathematical point that allows to establish criteria for self-duality, thus avoiding the explicit check of the symmetry at the level of the equation of motions, is that the equations of motion (a system of PDEs) can be conveniently split in a set of equations that is of degree 0 (no derivatives on the field strengths F), the so-called constitutive relations (see e.g. (5.5), or (5.8)), and another set of degree 1 (see e.g. (5.2), (5.3) or (5.9), (5.10)). Duality rotations act as an obvious symmetry of the set of equations of degree 1, so all what is left is to check that they act as a symmetry on the set of equations of degree 0. It is therefore plausible that this check can be equivalently formulated as a specific transformation property of the lagrangian under duality rotations (and independent from the spacetime dependence $F_{\mu\nu}(x)$ of the fields), indeed both the lagrangian and the equations of motions of degree 0 are functions of the field strength F and not of its derivatives.

5.2.1 Duality symmetry in nonlinear electromagnetism

Maxwell equations read

$$\partial_t \boldsymbol{B} = -\nabla \times \boldsymbol{E} \quad , \quad \nabla \cdot \boldsymbol{B} = 0 \tag{5.2}$$
$$\partial_t \boldsymbol{D} = \nabla \times \boldsymbol{H} \quad , \quad \nabla \cdot \boldsymbol{D} = 0 \tag{5.3}$$

they are complemented by the relations between the electric field $\boldsymbol{E}$, the magnetic field $\boldsymbol{H}$, the electric displacement $\boldsymbol{D}$ and the magnetic induction $\boldsymbol{B}$. In vacuum we have

$$\boldsymbol{D} = \boldsymbol{E} \; , \quad \boldsymbol{H} = \boldsymbol{B} \; . \tag{5.4}$$

In a nonlinear theory we still have the equations (5.2), (5.3), but the relations $\boldsymbol{D} = \boldsymbol{E}$, $\boldsymbol{H} = \boldsymbol{B}$ are replaced by the nonlinear constitutive relations

$$\boldsymbol{D} = \boldsymbol{D}(\boldsymbol{E}, \boldsymbol{B}) \quad , \quad \boldsymbol{H} = \boldsymbol{H}(\boldsymbol{E}, \boldsymbol{B}) \tag{5.5}$$

(if we consider a material medium with electric and magnetic properties then these equations are the constitutive relations of the material, and (5.2) and (5.3) are the macroscopic Maxwell equations).

Equations (5.2), (5.3), (5.4) are invariant under the group of general linear transformations

$$\begin{pmatrix} \boldsymbol{B}' \\ \boldsymbol{D}' \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \boldsymbol{B} \\ \boldsymbol{D} \end{pmatrix} \quad , \quad \begin{pmatrix} \boldsymbol{E}' \\ \boldsymbol{H}' \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{H} \end{pmatrix} \quad . \tag{5.6}$$

We study under which conditions also the nonlinear constitutive relations (5.5) are invariant. We find constraints on the relations (5.5) as well as on the transformations (5.6).

We are interested in nonlinear theories that admit a lagrangian formulation so that relativistic covariance of the equations (5.2), (5.3), (5.5) and their inner consistency is automatically ensured. This requirement is fulfilled if the constitutive relations (5.5) are of the form

$$\boldsymbol{D} = \frac{\partial \mathcal{L}(\boldsymbol{E}, \boldsymbol{B})}{\partial \boldsymbol{E}} \quad , \quad \boldsymbol{H} = -\frac{\partial \mathcal{L}(\boldsymbol{E}, \boldsymbol{B})}{\partial \boldsymbol{B}} \quad , \tag{5.7}$$

where $\mathcal{L}(\boldsymbol{E}, \boldsymbol{B})$ is a Poincaré invariant function of $\boldsymbol{E}$ and $\boldsymbol{B}$. Indeed if we consider $\boldsymbol{E}$ and $\boldsymbol{B}$ depending on a gauge potential A_μ and vary the lagrangian $\mathcal{L}(\boldsymbol{E}, \boldsymbol{B})$ with respect to A_μ, we recover (5.2), (5.3) and (5.7). This property is most easily shown by using four component notation. We group the constitutive relations (5.7) in the constitutive relation*

$$\tilde{G}^{\mu\nu} \;=\; 2\frac{\partial \mathcal{L}(F)}{\partial F_{\mu\nu}} \; ; \tag{5.8}$$

we also define $G_{\mu\nu} = -\frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\tilde{G}^{\rho\sigma}$, so that $\tilde{G}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}G_{\rho\sigma}$ ($\epsilon^{0123} = -\epsilon_{0123} = 1$). If we consider the field strength $F_{\mu\nu}$ as a function of a (locally defined) gauge potential A_μ, then equations (5.2) and (5.3) are respectively the Bianchi identities for $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and the equations of motion for $\mathcal{L}(F(A))$,

$$\partial_\mu \tilde{F}^{\mu\nu} = 0 \; , \tag{5.9}$$
$$\partial_\mu \tilde{G}^{\mu\nu} = 0 \quad . \tag{5.10}$$

*a practical convention is to define $\frac{\partial F_{\rho\sigma}}{\partial F_{\mu\nu}} = \delta^\mu_\rho \delta^\nu_\sigma$ rather than $\frac{\partial F_{\rho\sigma}}{\partial F_{\mu\nu}} = \delta^\mu_\rho \delta^\nu_\sigma - \delta^\nu_\rho \delta^\mu_\sigma$. This explains the factor 2 in (5.8).

In our treatment of duality rotations we study the symmetries of the equations (5.9), (5.10) and (5.8). The lagrangian $\mathcal{L}(F)$ is always a function of the field strength F; it is not seen as a function of the gauge potential A_μ; accordingly the Bianchi identities for F are considered part of the equations of motions for F.

Finally we consider an action $S = \int \mathcal{L}\, d^4x$ with lagrangian density $\mathcal{L} = \mathcal{L}(F)$ that depends on F but not on its partial derivatives; it also depends on a spacetime metric $g_{\mu\nu}$ that we generally omit writing explicitly*, and on at least one dimensionful constant in order to allow for nonlinearity in the constitutive relations (5.8) (i.e. (5.5)). We set this dimensionful constant to 1.

The duality rotations (5.6) read

$$\begin{pmatrix} F' \\ G' \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} F \\ G \end{pmatrix} . \tag{5.11}$$

Since by construction equations (5.9) and (5.10) are invariant under (5.11), these duality rotations are a symmetry of the system of equations (5.9), (5.10), (5.8) (or (5.2), (5.3), (5.5)), iff on shell the constitutive relations are invariant in form, i.e., iff the functional dependence of $\tilde{G}'$ from F' is the same as that of $\tilde{G}$ from F, i.e. iff

$$\tilde{G}'^{\mu\nu} \;=\; 2\frac{\partial \mathcal{L}(F')}{\partial F'_{\mu\nu}} , \tag{5.12}$$

where $F'_{\mu\nu}$ and $G'_{\mu\nu}$ are given in (5.11). This is the condition that constrains the lagrangian $\mathcal{L}(F)$ and the rotation parameters in (5.11). This condition has to hold on shell of (5.8)-(5.10); however (5.12) is not a differential equation and therefore has to hold just using (5.8), i.e., off shell of (5.9) and (5.10) (indeed if it holds for constant field strengths F then it holds for any F).

In order to study the duality symmetry condition (5.12) let $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) + \epsilon \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) + \ldots$, and consider infinitesimal $GL(2,\mathbb{R})$ rotations $G \to G + \epsilon\Delta G,\;\; F \to F + \epsilon\Delta F$,

$$\Delta \begin{pmatrix} F \\ G \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} F \\ G \end{pmatrix} , \tag{5.13}$$

so that the duality condition reads

$$\tilde{G} + \Delta\tilde{G} = 2\frac{\partial \mathcal{L}(F + \Delta F)}{\partial (F + \Delta F)} . \tag{5.14}$$

The right hand side simplifies to†

$$\begin{aligned} \frac{\partial \mathcal{L}(F+\Delta F)}{\partial (F+\Delta F)} &= \frac{\partial \mathcal{L}(F+\Delta F)}{\partial F}\frac{\partial F}{\partial (F+\Delta F)} \\ &= \frac{\partial \mathcal{L}(F+\Delta F)}{\partial F} - \frac{\partial \mathcal{L}(F)}{\partial F}\frac{\partial (\Delta F)}{\partial F} \end{aligned}$$

then, using (5.13) and (5.8), condition (5.14) reads

$$c\tilde{F} + d\tilde{G} = 2\frac{\partial(\mathcal{L}(F+\Delta F) - \mathcal{L}(F))}{\partial F} - 2a\frac{\partial \mathcal{L}(F)}{\partial F} - b\tilde{G}\frac{\delta G}{\partial F} . \tag{5.15}$$

*Notice that (5.9), (5.10) are also the equation of motions in the presence of a nontrivial metric. Indeed $S = \int \mathcal{L}\, d^4x = \int L\,\sqrt{g} d^4x$. The equation of motions are $\partial_\mu(\sqrt{g}\,F^{*\,\mu\nu}) = \partial_\mu \tilde{F}^{\mu\nu} = 0$, $\partial_\mu(\sqrt{g}\,G^{*\,\mu\nu}) = \partial_\mu \tilde{G}^{\mu\nu} = 0$, where the Hodge dual of a two form $\Omega_{\mu\nu}$ is defined by $\Omega^*_{\mu\nu} \equiv \frac{1}{2}\sqrt{g}\,\epsilon_{\mu\nu\rho\sigma}\Omega^{\rho\sigma}$.

†here and in the following we suppress the spacetime indices so that for example $F\tilde{G} = F_{\mu\nu}\tilde{G}^{\mu\nu}$; notice that $F\tilde{G} = \tilde{F}G$, $\tilde{\tilde{F}} = -F$, and $\tilde{F}\tilde{G} = -FG$ where $FG = F^{\mu\nu}G_{\mu\nu}$.

In order to further simplify this expression we write $2\tilde{F} = \frac{\partial}{\partial F} F\tilde{F}$ and we factorize out the partial derivative $\frac{\partial}{\partial F}$. We thus arrive at the equivalent condition

$$\mathcal{L}(F + \Delta F) - \mathcal{L}(F) - \frac{c}{4}F\tilde{F} - \frac{b}{4}G\tilde{G} \;=\; (a+d)(\mathcal{L}(F) - \mathcal{L}_{F=0}) \; . \tag{5.16}$$

The constant term $(a+d)\mathcal{L}_{F=0}$, nonvanishing for example in D-brane lagrangians, is obtained by observing that when $F = 0$ also $G = 0$.

Next use $\mathcal{L}(F + \Delta F) - \mathcal{L}(F) = \frac{\partial \mathcal{L}(F)}{\partial F}\Delta F = \frac{1}{2}aF\tilde{G} + \frac{1}{2}bG\tilde{G}$ in order to rewrite expression (5.16) as

$$\frac{b}{4}G\tilde{G} - \frac{c}{4}F\tilde{F} = (a+d)(\mathcal{L}(F) - \mathcal{L}_{F=0}) - \frac{a}{2}F\tilde{G} \; . \tag{5.17}$$

If we require the nonlinear lagrangian $\mathcal{L}(F)$ to reduce to the usual Maxwell lagrangian in the weak field limit, $F^4 << F^2$, i.e., $\mathcal{L}(F) = \mathcal{L}_{F=0} - 1/4\int FF d^4x + O(F^4)$, then $\tilde{G} = -F + O(F^3)$, and we obtain the constraint (recall that $\tilde{\tilde{G}} = -G$)

$$b = -c \quad , \qquad a = d \; ,$$

the duality group can be at most $SO(2)$ rotations times dilatations. Condition (5.17) becomes

$$\frac{b}{4}\Big(G\tilde{G} + F\tilde{F}\Big) \;=\; 2a\Big(\mathcal{L}(F) - \mathcal{L}_{F=0} - \frac{1}{2}F\frac{\partial \mathcal{L}}{\partial F}\Big) \; . \tag{5.18}$$

The vanishing of the right hand side holds only if either $\mathcal{L}(F) - \mathcal{L}_{F=0}$ is quadratic in F (usual electromagnetism) or $a = 0$. We are interested in nonlinear theories; by definition in a nonlinear theory $\mathcal{L}(F)$ is not quadratic in F. This shows that dilatations alone cannot be a duality symmetry. If we require the duality group to contain at least $SO(2)$ rotations then

$$G\tilde{G} + F\tilde{F} = 0 \; , \tag{5.19}$$

and $SO(2)$ is the maximal duality group. Relation (5.18) is nontrivially satisfied iff

$$a = d = 0 \; ,$$

and (5.19) hold.

In conclusion equation (5.19) is a necessary and sufficient condition for a nonlinear electromagnetic theory to be symmetric under $SO(2)$ duality rotations, and $SO(2) \subset GL(2,\mathbb{R})$ is the maximal connected Lie group of duality rotations of pure nonlinear electromagnetism*.

This conclusion still holds if we consider a nonlinear lagrangian $\mathcal{L}(F)$ that in the weak field limit $F^4 << F^2$ (up to an overall normalization factor) reduces to the most general linear lagrangian

$$\mathcal{L}(F) = \mathcal{L}_{F=0} - \frac{1}{4}FF + \frac{1}{4}\Theta F\tilde{F} + O(F^4) \; .$$

In this case $G = \tilde{F} + \Theta F + O(F^3)$. We substitute in (5.17) and obtain the two conditions (the coefficients of the scalar F^2 and of the speudoscalar $F\tilde{F}$ have to vanish separately)

$$c = -b(1+\Theta^2) \quad , \qquad d - a = 2\Theta b \; . \tag{5.20}$$

*This symmetry cannot even extend to $O(2)$ because already in the case of usual electromagnetism the finite rotation $\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ does not satisfy the duality condition (5.12). It is instructive to see the obstruction at the hamiltonian level. The hamitonian itself is invariant under $D \to D$, $B \to -B$, but this transformation is not a canonical transformation: the Poisson bracket (5.33) is not invariant.

The most general infinitesimal duality transformation is therefore

$$\begin{pmatrix} a & b \\ -b(1+\Theta^2) & a+2\Theta b \end{pmatrix} = \begin{pmatrix} a+\Theta b & 0 \\ 0 & a+\Theta b \end{pmatrix} + \Theta \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} \Theta^{-1} \tag{5.21}$$

where $\Theta = \begin{pmatrix} 1 & 0 \\ \Theta & 1 \end{pmatrix}$. We have dilatations and $SO(2)$ rotations, they act on the vector $\begin{pmatrix} F \\ G \end{pmatrix}$ via the conjugate representation given by the matrix Θ. Let's now remove the weak field limit assumtion $F^4 << F^2$. We proceed as before. >From (5.12) (or from (5.17)) we immediately obtain that dilatations alone are not a duality symmetry of the nonlinear equations of motion. Then if $SO(2)$ rotations are a duality symmetry we have that they are the maximal duality symmetry group. This happens if

$$G\tilde{G} + (1+\Theta)^2 F\tilde{F} = 2\Theta F\tilde{G} \ . \tag{5.22}$$

Finally we note that the necessary and sufficient conditions for $SO(2)$ duality rotations (5.22) (or (5.19)) can be equivalently expressed as invariance of

$$\mathcal{L}(F) - \frac{1}{4} F\tilde{G} \ . \tag{5.23}$$

Proof: the variation of expression (5.23) under $F \to F + \Delta F$ is given by $\mathcal{L}(F+\Delta F) - \mathcal{L}(F) - \frac{1}{4}\Delta F\,\tilde{G} - \frac{1}{4}F\Delta\tilde{G}$. Use of (5.16) with $a+d=0$ (no dilatation) shows that this variation vanishes.

5.2.2 Legendre Transformations

In the literature on gauge theories of abelian p-form potentials, the term duality transformation denotes a different transformation from the one we have introduced, a Legendre transformation, that is not a symmetry transformation. In this section we relate these two different notions, see [15] for further applications and examples.

Consider a theory of nonlinear electrodynamics ($p=1$) with lagrangian $\mathcal{L}(F)$. The equations of motion and the Bianchi identity for F can be derived from the Lagrangian $\mathcal{L}(F, F_{\rm D})$ defined by

$$\mathcal{L}(F, F_{\rm D}) = \mathcal{L}(F) - \frac{1}{2}\, F\tilde{F}_{\rm D} \ , \qquad F_{\rm D}{}^{\mu\nu} = \partial^\mu A_{\rm D}{}^\nu - \partial^\nu A_{\rm D}{}^\mu \ , \tag{5.24}$$

where F is now an unconstrained antisymmetric tensor field, $A_{\rm D}$ a Lagrange multiplier field and $F_{\rm D}$ its electromagnetic field. [Hint: varying with respect to $A_{\rm D}$ gives the Bianchi identity for F, varying with respect to F gives $G^{\mu\nu} = F_{\rm D}{}^{\mu\nu}$ that is equivalent to the initial equations of motion $\partial_\mu \tilde{G}^{\mu\nu} = 0$ because $F_{\rm D}{}^{\mu\nu} = \partial^\mu A_{\rm D}{}^\nu - \partial^\nu A_{\rm D}{}^\mu$ (Poincaré lemma)].

Given the lagrangian (5.24) one can also first consider the equation of motion for F,

$$G(F) \ = \ F_{\rm D} \ , \tag{5.25}$$

that is solved by expressing F as a function of the dual field strength, $F = F(F_{\rm D})$. Then inserting this solution into $\mathcal{L}(F, F_{\rm D})$, one gets the dual model

$$\mathcal{L}_{\rm D}(F_{\rm D}) \equiv \mathcal{L}(F(F_{\rm D})) - \frac{1}{2} F(F_{\rm D}) \cdot \tilde{F}_{\rm D} \ . \tag{5.26}$$

Solutions of the (5.26) equations of motion are, tothether with (5.25), solutions of the (5.24) equations of motion. Therefore solutions to the (5.26) equations of motion are via (5.25) in 1-1 correspondence with solutions of the $\mathcal{L}(F)$ equations of motion.

One can always perform a Legendre transformation and describe the physical system with the new dynamical variables $A_{\rm D}$ and the new lagrangian $\mathcal{L}_{\rm D}$ rather than A and $\mathcal{L}$.

The relation with the duality rotation symmetry (self-duality) of the previous section is that if the system admits duality rotations then the solution $F_{\rm D}$ of the $\mathcal{L}_{\rm D}$ equations of motion is also a solution of the $\mathcal{L}$ equations of motion, we have a symmetry because the dual field $F_{\rm D}$ is a solution of the original system. This is the case because for any solution $\mathcal{L}$ of the self-duality equation, its Legendre transform $\mathcal{L}_{\rm D}$ satisfies:

$$\mathcal{L}_{\rm D}(F) \;=\; \mathcal{L}(F)\;. \tag{5.27}$$

This follows from considering a finite $SO(2)$ duality rotation with angle $\pi/2$. Then $F \to F' = G(F) = F_{\rm D}$, and invariance of (5.23), i.e. $\mathcal{L}(F') - \frac{1}{4}F'\tilde{G}' = \mathcal{L}(F) - \frac{1}{4}F\tilde{G}$, implies $\mathcal{L}_{\rm D}(F_{\rm D}) \;=\; \mathcal{L}(F_{\rm D})$, i.e., (5.27).

In summary, a Legendre transformation is a duality rotation only if the symmetry condition (5.8) is met. If the self-duality condition (5.8) does not hold, a Legendre transformation leads to a dual formulation of the theory in terms of a dual Lagrangian $\mathcal{L}_{\rm D}$, not to a symmetry of the theory.

5.2.3 Hamiltonian theory

The symmetric energy monentum tensor of a nonlinear theory of electromagnetism (obtained via Belinfante procedure or by varying with respect to the metric) is given by*

$$T^{\mu}_{\ \nu} = \tilde{G}^{\mu\lambda}F_{\nu\lambda} + \partial^{\mu}_{\ \nu}\mathcal{L}\;. \tag{5.28}$$

The equations of motion (5.10) and (5.9) imply its conservation, $\partial_\mu T^{\mu}_{\ \nu} = 0$. Invariance of the energy momentum tensor under duality rotations is easily proven by observing that for a generic antisymmetric tensor $F_{\mu\nu}$

$$\tilde{F}^{\mu\lambda}F_{\nu\lambda} = -\frac{1}{4}\partial^{\mu}_{\ \lambda}\tilde{F}^{\rho\sigma}F_{\rho\sigma}\;, \tag{5.29}$$

and then by recalling the duality symmetry condition (5.19).

In particular the hamiltonian

$$\mathcal{H} = T^{00} = \boldsymbol{D}\cdot\boldsymbol{E} - \mathcal{L} \tag{5.30}$$

of a theory that has duality rotation symmetry is invariant.

In the hamiltonian formalism duality rotations are canonical transformations, since they leave the hamiltonian invariant they are usual symmetry transformations. We briefly describe the hamiltonian formalism of (nonlinear) electromagnetism by avoiding to introduce the vector potential A_μ; this is appropriate since duality rotations are formulated independently from the notion of vector potential. Maxwell equations (5.2), (5.3) and the expression of the hamiltonian suggest to consider $\boldsymbol{B}$ and $\boldsymbol{D}$ as the analogue of canonical coordinates and momenta q and p, while $\boldsymbol{E}$, that enters the lagrangian togheter with $\boldsymbol{B}$, is the analogue of $\dot{q}$.

*symmetry of $T^{\mu\nu}$ follows immediately by observing that the tensor structure of $\tilde{G}^{\mu\nu}$ implies $\tilde{G}^{\mu\nu} = f_s(F)F^{\mu\nu} + f_p(F)\tilde{F}^{\mu\nu}$ with scalars $f_s(F)$ and $f_p(F)$ depending on F, the metric $\eta = diag(-1,1,1,1)$ and the completely antisymmetric tensor density $\epsilon_{\mu\nu\rho\sigma}$. (Actually, if the lagrangian is parity even, f_s is a scalar function while f_p is a pseudoscalar function).

Recalling the constitutive relations in the lagrangian form (5.7) we obtain that the hamiltonian $\mathcal{H} = \mathcal{H}(\boldsymbol{D}, \boldsymbol{B})$ is just given by the Legendre transformation (5.30). Moreover $\boldsymbol{H} = \frac{\partial \mathcal{H}}{\partial \boldsymbol{B}}$ and $\boldsymbol{E} = \frac{\partial \mathcal{H}}{\partial \boldsymbol{D}}$. The equations of motion are

$$\partial_t \boldsymbol{B} = -\nabla \times \frac{\delta \mathcal{H}}{\delta \boldsymbol{D}} \ , \tag{5.31}$$

$$\partial_t \boldsymbol{D} = \nabla \times \frac{\delta \mathcal{H}}{\delta \boldsymbol{B}} \quad . \tag{5.32}$$

The remaning equations $\nabla \cdot \boldsymbol{B} = 0$, $\nabla \cdot \boldsymbol{D} = 0$ are constraints that imposed at a given time are satisfied at any other time. The Poisson bracket between two arbitrary functionals U, V of the canonical variables is

$$\{U, V\} = \int \frac{\partial U}{\partial \boldsymbol{D}} \cdot \left(\nabla \times \frac{\partial V}{\partial \boldsymbol{B}} \right) - \frac{\partial V}{\partial \boldsymbol{D}} \cdot \left(\nabla \times \frac{\partial U}{\partial \boldsymbol{B}} \right) \, d^3 r \ , \tag{5.33}$$

in particular the only nonvanishing parenthesis between the canonical variables $\boldsymbol{B}$ and $\boldsymbol{D}$ are $\{\boldsymbol{B}^i(r), \boldsymbol{D}^j(r')\} = \epsilon^{ijk} \partial_k \partial^3(r - r')$. The equations of motion (5.31) and (5.32) assume then the canonical form $\partial_t \boldsymbol{B} = -\{\boldsymbol{B}, H\}$, $\partial_t \boldsymbol{D} = \{\boldsymbol{D}, H\}$ where $H = \int \mathcal{H} \, d^3 r$ is the hamiltonian ($\mathcal{H}$ being the hamiltonian density). We see that H as usual is the generator of time evolution. The consitency and the hidden Poincaré invariance of the present formalism is proven in [55].

In the canonical formalism the generator of duality rotations is the following nonlocal integral [57], [56]

$$\Lambda = \frac{1}{8\pi} \int \!\! \int \frac{\boldsymbol{D}_1 \cdot (\nabla \times \boldsymbol{D}_2) + \boldsymbol{B}_1 \cdot (\nabla \times \boldsymbol{B}_2)}{|r_1 - r_2|} \, d^3 r_1 d^3 r_2 \tag{5.34}$$

where the subscripts indicate that the fields are taken at the points r_1 and r_2. We have $\{\boldsymbol{D}, \Lambda\} = \boldsymbol{B}$ and $\{\boldsymbol{B}, \Lambda\} = -\boldsymbol{D}$.

Finally we remark that it is straighforward to establish duality symmetry in the hamiltonian formalism. Indeed there are three independent scalar combinations of the canonical fields $\boldsymbol{B}$ and $\boldsymbol{D}$, they can be taken to be: $\boldsymbol{D}^2 + \boldsymbol{B}^2$, $\boldsymbol{D}^2 - \boldsymbol{B}^2$ and $(\boldsymbol{D} \times \boldsymbol{B})^2$. The last two scalars are duality invariant and therefore any hamiltonian that depends just on them leads to a theory with duality symmetry. The nontrivial problem in this approach in now to constrain the hamiltonian so that the theory is Lorentz invariant [58], [57]. The condition is again (5.19) i.e., $\boldsymbol{D} \cdot \boldsymbol{H} = \boldsymbol{E} \cdot \boldsymbol{B}$.

5.2.4 Born-Infeld lagrangian

A notable example of a lagrangian whose equations of motion are invariant under duality rotations is given by the Born-Infeld one [1]

$$\mathcal{L}_{\mathrm{BI}} = 1 - \sqrt{-\det(\eta + F)} \tag{5.35}$$

$$= 1 - \sqrt{1 + \frac{1}{2} F^2 - \frac{1}{16} (F \tilde{F})^2} \tag{5.36}$$

$$= 1 - \sqrt{1 - \boldsymbol{E}^2 + \boldsymbol{B}^2 - (\boldsymbol{E} \cdot \boldsymbol{B})^2} \ . \tag{5.37}$$

In the second line we have simply expanded the 4x4 determinant and espressed the lagrangian in terms of the only two independent Lorentz invariants associated to the electromagnetic field: $F^2 \equiv F_{\mu\nu} F^{\mu\nu}$, $F \tilde{F} \equiv F_{\mu\nu} \tilde{F}^{\mu\nu}$.

The explicit expression of G is

$$G_{\mu\nu} = \frac{\tilde{F}_{\mu\nu} + \frac{1}{4} F\tilde{F}\, F_{\mu\nu}}{\sqrt{1 + \frac{1}{2}F^2 - \frac{1}{16}(F\tilde{F})^2}} \ , \tag{5.38}$$

and the duality condition (5.19) is readily seen to hold. The hamiltonian is

$$\mathcal{H}_{\rm BI} = \sqrt{1 + \boldsymbol{D}^2 + \boldsymbol{B}^2 + (\boldsymbol{D} \times \boldsymbol{B})^2} \ - 1 \ . \tag{5.39}$$

Notice that while the $\boldsymbol{E}$ and $\boldsymbol{B}$ variables are constrained by the reality of the square root in the lagrangian, the hamiltonian variables $\boldsymbol{D}, \boldsymbol{B}$ are unconstrained. By using the equations of motion and (5.19) it can be explicitly verified that the generator of duality rotations is time independent, $\{\Lambda, H\} = 0$.

5.2.5 Extended duality rotations

The duality symmetry of the equations of motion of nonlinear electromagnetism can be extended to $SL(2,\mathbb{R})$. We observe that the definition of duality symmetry we used can be relaxed by allowing the F dependence of G to change by a linear term: $G = 2\frac{\partial \mathcal{L}}{\partial F}$ and $G = 2\frac{\partial \mathcal{L}}{\partial F} + \vartheta F$ togheter with the Bianchi identities for F give equivalent equations of motions for F. Therefore the transformation

$$\begin{pmatrix} F' \\ G' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \vartheta & 1 \end{pmatrix} \begin{pmatrix} F \\ G \end{pmatrix} \tag{5.40}$$

is a symmetry of any nonlinear electromagnetism. It corresponds to the lagrangian change $\mathcal{L} \to \mathcal{L} + \frac{1}{4}\vartheta F\tilde{F}$. This symmetry alone does not act on F, but it is useful if the nonlinear theory has $SO(2)$ duality symmetry. In this case (5.40) extends duality symmetry from $SO(2)$ to $SL(2,\mathbb{R})$ (i.e. $Sp(2,\mathbb{R})$). Notice however that the $SL(2,\mathbb{R})$ transformed solution, contrary to the $SO(2)$ one, has a different energy and energy momentum tensor (recall (5.28)). On the other hand, as we show in Section 5.3.6, if the constant ϑ is promoted to a dynamical field we have invariance of the energy momentum tensor under $SL(2,\mathbb{R})$ duality.

5.3 General theory of duality rotations

We study in full generality the conditions in order to have theories with duality rotation symmetry. By properly introducing scalar fields (sigma model on coset space) we enhance theories with a compact duality group to theories with an extended noncompact duality group. A Born-Infeld lagrangian with n abelian field strengths and $U(n)$ duality group (or $Sp(2n,\mathbb{R})$ in the presence of scalars) is constructed.

5.3.1 General nonlinear theory

We consider a theory of n abelian gauge fields possibly coupled to other bosonic and fermionic fields that we denote φ^α, $(\alpha = 1,...p)$. We assume that the $U(1)$ gauge potentials enter the action $S = S[F,\varphi]$ only trough the field strengths $F^\Lambda_{\mu\nu}$ $(\Lambda = 1,\ldots,n)$, and that the action does not depend on partial derivatives of the field strengths. Define $\tilde{G}_\Lambda{}^{\mu\nu} = 2\frac{\partial \mathcal{L}}{\partial F^\Lambda_{\mu\nu}}$, i.e,

$$\tilde{G}_\Lambda{}^{\mu\nu} = 2\frac{\delta S[F,\varphi]}{\delta F^\Lambda_{\mu\nu}} \ ; \tag{5.41}$$

then the Bianchi identities and the equations of motions for $S[F,\varphi]$ are

$$\partial_\mu \tilde{F}^{\Lambda\,\mu\nu} = 0 \;, \tag{5.42}$$

$$\partial_\mu \tilde{G}_\Lambda{}^{\mu\nu} = 0 \;, \tag{5.43}$$

$$\frac{\delta S[F,\varphi]}{\delta \varphi^\alpha} = 0 \;. \tag{5.44}$$

The field theory is described by the system of equations (5.41)-(5.44). Consider the duality transformations

$$\begin{pmatrix} F' \\ G' \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} F \\ G \end{pmatrix} \tag{5.45}$$

$$\varphi'^\alpha = \Xi^\alpha(\varphi) \tag{5.46}$$

where $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ is a generic constant $L(2n,\mathbb{R})$ matrix and the φ^α fields transformation in full detail reads $\varphi'^\alpha = \Xi^\alpha(\varphi, \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right))$, with no partial derivative of φ appearing in Ξ^α.

These duality rotations are a symmetry of the system of equations (5.41)-(5.44) iff, given F, G, and φ solution of (5.41)-(5.44) then F', G' and φ', that by construction satisfy $\partial_\mu \tilde{F}'^{\Lambda\,\mu\nu} = 0$ and $\partial_\mu \tilde{G}'_\Lambda{}^{\mu\nu} = 0$, satisfy also

$$\tilde{G}'_\Lambda{}^{\mu\nu} = 2\frac{\delta S[F',\varphi']}{\delta F'^\Lambda_{\mu\nu}} \;, \tag{5.47}$$

$$\frac{\delta S[F',\varphi']}{\delta \varphi'^\alpha} = 0 \;. \tag{5.48}$$

We study these on shell conditions in the case of infinitesimal $GL(2n,\mathbb{R})$ rotations

$$F \to F' = F + \Delta F \quad , \quad G \to G' = G + \Delta G \;,$$

$$\Delta \begin{pmatrix} F \\ G \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} F \\ G \end{pmatrix} \;, \tag{5.49}$$

$$\Delta\varphi^\alpha = \xi^\alpha(\varphi) \;. \tag{5.50}$$

The right hand side of (5.47) can be rewritten as

$$\frac{\delta S[F',\varphi']}{\delta F'^\Lambda} = \int_y \frac{\delta S[F',\varphi']}{\delta F^\Sigma(y)} \frac{\delta F^\Sigma(y)}{\delta F'^\Lambda} \;. \tag{5.51}$$

We now invert the matrix $\begin{pmatrix} \frac{\delta F'}{\delta F} & \frac{\delta F'}{\delta \varphi} \\ \frac{\delta \varphi'}{\delta F} & \frac{\delta \varphi'}{\delta \varphi} \end{pmatrix}$, recall that $F_1 \tilde{F}_2 = \tilde{F}_1 F_2$ and observe that

$$\int_y \frac{\delta S[F,\varphi]}{\delta F(y)}\, b\, \frac{\delta G(y)}{\delta F^\Lambda} = \frac{1}{4}\frac{\delta}{\delta F^\Lambda}\int_y \tilde{G}\, b\, G + \frac{1}{4}\int_y \tilde{G}(b - b^t)\frac{\delta G}{\delta F^\Lambda} \;.$$

We thus obtain

$$\frac{\delta S[F',\varphi']}{\delta F'^\Lambda} = \frac{\delta S[F',\varphi']}{\delta F^\Lambda} - a^{\Sigma\Lambda}\frac{\delta S[F',\varphi']}{\delta F^\Sigma} - \frac{1}{4}\frac{\delta}{\delta F^\Lambda}\int_y \tilde{G}\, b\, G - \frac{1}{4}\int_y \tilde{G}(b - b^t)\frac{\delta G}{\delta F^\Lambda} \;. \tag{5.52}$$

Since the left hand side of (5.47) is $\tilde{G}_\Lambda + \frac{1}{2}\frac{\delta}{\delta F^\Lambda}\int_y \tilde{F}\, c\, F + \frac{1}{2}(c-c^t)_{\Lambda\Sigma}\tilde{F}^\Sigma + 2d_\Lambda^{\ \Sigma}\frac{\delta S[F,\varphi]}{\delta F^\Sigma}$, we rewrite (5.47) as

$$\begin{aligned}\frac{\delta}{\delta F^\Lambda}&\left(S[F',\varphi'] - S[F,\varphi] - \frac{1}{4}\int_y (\tilde{F}\, c\, F + \tilde{G}\, b\, G)\right) \\ &= (a^t+d)_\Lambda^{\ \Sigma}\frac{\delta}{\delta F^\Sigma}S[F,\varphi] + \frac{1}{4}(c-c^t)_{\Lambda\Sigma}\tilde{F}^\Sigma + \frac{1}{4}\int_y \tilde{G}(b-b^t)\frac{\delta G}{\delta F^\Lambda}\ .\end{aligned} \tag{5.53}$$

Since this expression does not contain derivatives of F, the functional variation becomes just a partial derivative, and (5.53) is equivalent to

$$\begin{aligned}\frac{\partial}{\partial F^\Lambda}&\left(\mathcal{L}(F',\varphi') - \mathcal{L}(F,\varphi) - \frac{1}{4}\tilde{F}\, c\, F - \frac{1}{4}\tilde{G}\, b\, G\right) \\ &= (a^t+d)_\Lambda^{\ \Sigma}\frac{\partial}{\partial F^\Sigma}\mathcal{L}(F,\varphi) + \frac{1}{4}(c-c^t)_{\Lambda\Sigma}\tilde{F}^\Sigma + \frac{1}{4}\tilde{G}(b-b^t)\frac{\partial G}{\partial F^\Lambda}\ .\end{aligned} \tag{5.54}$$

Here $\mathcal{L}(F,\varphi)$ is a shorthand notation for a lagrangian that depends on F, φ^α, $\partial\varphi^\alpha$ and eventually higher partial derivatives of the fields φ^α, say up to order ℓ. Equation (5.54) has to hold on shell of (5.41)-(5.44). Since this equation has no partial derivative of F and at most derivatives of φ^α up to order ℓ, if it holds on shell of (5.41)-(5.44) then it holds just on shell of (5.41), and of the fermions fields equations, the scalar and vector partial differential equations being of higher order in derivatives of F or φ^α fields. In particular if no fermion is present (5.54) holds just on shell of (5.41).

Since the left hand side of (5.54) is a derivative with respect to F^Λ so must be the right hand side. This holds if we consider infinitesimal dilatations, parametrized by $\frac{\kappa}{2}\in\mathbb{R}$, and infinitesimal $Sp(2n,\mathbb{R})$ transformations

$$a^t + d = \kappa\mathbb{1}\quad ,\qquad b^t = b\quad ,\quad c^t = c\ . \tag{5.55}$$

We can then remove the derivative $\frac{\partial}{\partial F^\Lambda}$ and obtain the equivalent condition

$$\mathcal{L}(F',\varphi') - \mathcal{L}(F,\varphi) - \kappa\mathcal{L}(F,\varphi) - \frac{1}{4}\tilde{F}\, c\, F - \frac{1}{4}\tilde{G}\, b\, G = f(\varphi) \tag{5.56}$$

where $f(\varphi)$ can contain partial derivatives of φ up to the same order as in the lagrangian.

We now show that $f(\varphi)$ in (5.56) is independent from φ. Consider the φ-equations of motion (5.48),

$$\begin{aligned}\frac{\delta S[F',\varphi']}{\delta\varphi'^\alpha} &= \int_y \frac{\delta S[F',\varphi']}{\delta\varphi^\beta(y)}\frac{\delta\varphi^\beta(y)}{\delta\varphi'^\alpha} + \int_y \frac{\delta S[F,\varphi]}{\delta F(y)}\frac{\delta F(y)}{\delta\varphi'^\alpha} \\ &= \frac{\delta S[F',\varphi']}{\delta\varphi^\alpha} - \frac{\delta S[F,\varphi]}{\delta\varphi^\beta}\frac{\partial\xi^\beta}{\partial\varphi^\alpha} - \frac{1}{4}\frac{\delta}{\delta\varphi^\alpha}\int_y \tilde{G}\, b\, G \\ &= \frac{\delta S[F,\varphi]}{\delta\varphi^\alpha} - \frac{\delta S[F,\varphi]}{\delta\varphi^\beta}\frac{\partial\xi^\beta}{\partial\varphi^\alpha} + \frac{\delta}{\delta\varphi^\alpha}\left(S[F',\varphi'] - S[F,\varphi] - \frac{1}{4}\int_y \tilde{G}\, b\, G\right)\end{aligned}$$

where only first order infinitesimals have been retained, and where techniques similar to those used in the study of (5.51) have been applied. On shell the left hand side has to vanish; since the first two addends on the right hand side are proportional to the φ-equations of motion, this happens iff on shell

$$\frac{\delta}{\delta\varphi^\alpha}\left(S[F',\varphi'] - S[F,\varphi] - \kappa S[F,\varphi] - \frac{1}{4}\int_y (\tilde{G}\, b\, G + \tilde{F}\, c\, F)\right) \ = \ 0\ . \tag{5.57}$$

Comparison with (5.56) shows that on shell

$$\frac{\delta}{\delta\varphi^\alpha} f(\varphi) = 0 \ . \tag{5.58}$$

In this expression no field strength F is present and therefore the equations of motion of our interacting system are of no use; equation (5.58) holds also off shell and we conclude that $f(\varphi)$ is φ independent, it is just a constant depending on the parameters a, b, c, d (it usually vanishes). We thus have the condition

$$\mathcal{L}(F', \varphi') - \mathcal{L}(F, \varphi) - \kappa\mathcal{L}(F, \varphi) - \frac{1}{4}\tilde{F}\, c\, F - \frac{1}{4}\tilde{G}\, b\, G \;=\; const_{a,b,c,d} \tag{5.59}$$

If we expand F' in terms of F and G, we obtain the equivalent condition

$$\Delta_\varphi\mathcal{L}(F, \varphi) = \frac{1}{4}\tilde{F}\, c\, F - \frac{1}{4}\tilde{G}\, b\, G + \kappa\mathcal{L}(F, \varphi) - \frac{1}{2}\tilde{G}\, a\, F + const_{a,b,c,d} \tag{5.60}$$

where $\Delta_\varphi\mathcal{L}(F, \varphi) = \mathcal{L}(F, \varphi') - \mathcal{L}(F, \varphi)$.

Equation (5.60), where $\tilde{G}_\Lambda^{\mu\nu} = 2\partial\mathcal{L}/\partial F^\Lambda_{\mu\nu}$, is a necessary and sufficient condition in order to have duality symmetry. This condition is on shell of the fermions equations of motion, in particular if no fermion is present this condition is off shell. In the presence of fermions, equation (5.60) off shell is a sufficient condition for duality symmetry.

The duality symmetry group is

$$\mathbb{R}^{>0} \times SL(2n, \mathbb{R}) \ , \tag{5.61}$$

the group of dilatations times symplectic transformation; it is the connected Lie group generated by the Lie algebra (5.55). It is also the maximal group of duality rotations as the example (or better, the limiting case) studied in the next section shows.

We have considered dynamical fermionic and bosonic fields φ^α. If a subset χ^r of these fields is not dynamical the corresponding equations of motion are of the same order as those defining G, and thus (5.54) and (5.60) hold on shell of all these equations. Moreover since no $\partial\chi^r$ appears in the lagrangian, the duality transformations for these fields can include the field strength F, i.e., $\chi^r \to \chi'^r = \Xi^r(F, \chi)$. In this case there is an extra addend in (5.51). The necessary and sufficient duality condition (5.60) does not change.

We also notice that condition (5.59) in the absence of dilatations ($\kappa = 0$), and for $const_{a,b,c,d} = 0$ is equivalent to the invariance of

$$\mathcal{L} - \frac{1}{4}\tilde{F}G \ . \tag{5.62}$$

5.3.2 The main example and the scalar fields fractional transformations

Consider the Lagrangian

$$\frac{1}{4}\mathcal{N}_\in \mathcal{F}\mathcal{F} + \frac{\infty}{\triangle}\mathcal{N}_\infty \mathcal{F}\tilde{\mathcal{F}} + \mathscr{L}(\phi) \tag{5.63}$$

where the real symmetric matrices $\mathcal{N}_\infty(\phi)$ and $\mathcal{N}_\in(\phi)$ and the lagrangian $\mathscr{L}(\phi)$ are just functions of the bosonic fields ϕ^i, $i = 1, \ldots m$, (and their partial derivatives).

Any nonlinear lagrangian in the limit of vanishing fermionic fields and of weak field strengths $F^4 << F^2$ reduces to the one in (5.63). A straighforward calculation shows that this lagrangian has $\mathbb{R}^{>0} \times SL(2n, \mathbb{R})$ duality symmetry if the matrices $\mathcal{N}_\infty$ and $\mathcal{N}_\in$ of the scalar fields transform as

$$\Delta\,\mathcal{N}_\infty = c + d\,\mathcal{N}_\infty - \mathcal{N}_\infty \dashv - \mathcal{N}_\infty \lfloor \mathcal{N}_\infty + \mathcal{N}_\in \lfloor \mathcal{N}_\in \quad , \tag{5.64}$$

$$\Delta\,\mathcal{N}_\in = d\,\mathcal{N}_\in - \mathcal{N}_\in \dashv - \mathcal{N}_\infty \lfloor \mathcal{N}_\in - \mathcal{N}_\in \lfloor \mathcal{N}_\infty \quad , \tag{5.65}$$

and

$$\Delta\,\mathscr{L}(\phi) = \kappa \mathscr{L}(\phi) \ . \tag{5.66}$$

If we define

$$\mathcal{N} = \mathcal{N}_\infty + \rangle \mathcal{N}_\in \ ,$$

i.e., $\mathcal{N}_\infty = \mathrm{Re}\,\mathcal{N}$, $\mathcal{N}_\in = \mathrm{Im}\,\mathcal{N}$, the transformations (5.64), (5.65) read

$$\Delta\,\mathcal{N} \ = \ \rfloor + \lceil \mathcal{N} - \mathcal{N} \dashv - \mathcal{N} \lfloor \mathcal{N} \ , \tag{5.67}$$

the finite version is the fractional transformation

$$\mathcal{N}' = (\mathcal{C} + \mathcal{D}\mathcal{N})\,(\mathcal{A} + \mathcal{B}\mathcal{N})^{-\infty} \ . \tag{5.68}$$

Under (5.68) the imaginary part of $\mathcal{N}$ transforms as

$$\mathcal{N}'_\in = (\mathcal{A} + \mathcal{B}\mathcal{N})^{-\dagger} \mathcal{N}_\in (\mathcal{A} + \mathcal{B}\mathcal{N})^{-\infty} \tag{5.69}$$

where $-\dagger$ is a shorthand notation for the hermitian conjugate of the inverse matrix.

The kinetic term $\frac{1}{4}\mathcal{N}_\in \mathcal{F}\mathcal{F}$ is positive definite if the symmetric matrix $\mathcal{N}_\in$ is negative definite. In Appendix 7.2 we show that the matrices $\mathcal{N} = \mathcal{N}_\infty + \rangle \mathcal{N}_\in$ with $\mathcal{N}_\infty$ and $\mathcal{N}_\in$ real and symmetric, and $\mathcal{N}_\in$ positive definite, are the coset space $\frac{Sp(2n,\mathbb{R})}{U(n)}$.

A scalar lagrangian that satisfies the variation (5.66) can always be constructed using the geometry of the coset space $\frac{Sp(2n,\mathbb{R})}{U(n)}$, see Section 5.3.4.

This example also clarifies the condition (5.55) that we have imposed on the $GL(2n, \mathbb{R})$ generators. It is a straighfoward calculation to check that the equations (5.42), (5.43) and

$$\tilde{G} = \mathcal{N}_\in \mathcal{F} + \mathcal{N}_\infty \tilde{\mathcal{F}} \tag{5.70}$$

have duality symmetry under $GL(2n, \mathbb{R})$ transformations with $\Delta\mathcal{N}$ given in (5.67). However it is easy to see that equation (5.54) implies, for the lagrangian (5.63), that condition (5.55) must hold. The point is that we want the constitutive relations $G = G[F, \varphi]$ to follow from a lagrangian. Those following from the lagrangian (5.63) are (5.70) with $\mathcal{N}_\infty$ and $\mathcal{N}_\in$ necessarily *symmetric* matrices. Only if the transformed matrices $\mathcal{N}'_\infty$ and $\mathcal{N}'_\in$ are again symmetric we can have $\tilde{G}' = \frac{\partial \mathcal{L}(F',\varphi')}{\partial F'}$ as in (5.47), (or more generally $\tilde{G}' = \frac{\partial \mathcal{L}'(F',\varphi')}{\partial F'}$). The constraints $\mathcal{N}'_\infty = \mathcal{N}'^{\sqcup}_\infty$, $\mathcal{N}'_\in = \mathcal{N}'^{\sqcup}_\in$, reduce the duality group to $\mathbb{R}^{>0} \times SL(2n, \mathbb{R})$.

In conclusion equation (5.60) is a necessary and sufficient condition for a theory of n abelian gauge fields coupled to bosonic matter to be symmetric under $\mathbb{R}^{>0} \times SL(2n, \mathbb{R})$ duality rotations, and $\mathbb{R}^{>0} \times SL(2n, \mathbb{R})$ is the maximal connected Lie group of duality rotations.

5.3.3 A basic example with fermi fields

Consider the Lagrangian with Pauli coupling

$$\mathcal{L}_0 = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2}\overline{\psi} \partial\!\!\!/ \psi - \frac{1}{2}\overline{\xi} \partial\!\!\!/ \xi + \frac{1}{2}\lambda F^{\mu\nu} \overline{\psi} \sigma_{\mu\nu} \xi \tag{5.71}$$

where $\sigma^{\mu\nu} = \frac{1}{4}[\gamma^\mu, \gamma^\nu]$ and ψ, ξ are two Majorana spinors. We have

$$\tilde{G}^{\mu\nu} = 2\frac{\partial \mathcal{L}_0}{\partial F_{\mu\nu}} = -F^{\mu\nu} + \lambda\overline{\psi}\sigma^{\mu\nu}\xi \tag{5.72}$$

and the duality condition (5.60) for an infinitesimal $U(1)$ duality rotation $\left(\begin{smallmatrix} 0 & b \\ -b & 0 \end{smallmatrix}\right)$ reads

$$\Delta_\psi \mathcal{L}_0 + \Delta_\xi \mathcal{L}_0 = -\frac{b}{4}\lambda \tilde{F}\overline{\psi}\sigma\xi + \frac{b}{4}\lambda^2\, \overline{\psi}\sigma^{\mu\nu}\xi\, \overline{\psi}\tilde{\sigma}_{\mu\nu}\xi \ . \tag{5.73}$$

It is natural to assume that the kinetic terms of the fermion fields are invariant under this duality rotation (this is also the case for the scalar lagrangian $\mathscr{L}(\phi)$ in (5.66)), then using $\gamma_5\sigma^{\mu\nu} = i\tilde{\sigma}^{\mu\nu}$ we see that the coupling of the fermions with the field strength is reproduced if the fermions rotate according to

$$\Delta\psi = \frac{i}{2}b\gamma_5\psi \ , \tag{5.74}$$

$$\Delta\xi = \frac{i}{2}b\gamma_5\xi \ ; \tag{5.75}$$

we also see that we have to add to the lagrangian $\mathcal{L}_0$ a new interaction term quartic in the fermion fields. Its coupling is also fixed by duality symmetry to be $-\lambda^2/8$.

The theory with $U(1)$ duality symmetry is therefore given by the lagrangian [3]

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}\overline{\psi}\partial\!\!\!/\psi - \frac{1}{2}\overline{\xi}\partial\!\!\!/\xi + \frac{1}{2}\lambda F^{\mu\nu}\psi\sigma_{\mu\nu}\xi - \frac{1}{8}\lambda^2\, \overline{\psi}\sigma_{\mu\nu}\xi\, \overline{\psi}\sigma^{\mu\nu}\xi \ . \tag{5.76}$$

Notice that fermions transform under the double cover of $U(1)$ indeed under a rotation of angle $b = 2\pi$ we have $\psi \to -\psi$, $\xi \to -\xi$, this is a typical feature of fermions transformations under duality rotations, they transform under the double cover of the maximal compact subgroup of the duality group. This is so because the interaction with the gauge field is via fermions bilinear terms.

5.3.4 Compact and noncompact duality rotations

Compact duality rotations

The fractional transformation (5.68) is also characteristic of nonlinear theories. The subgroup of $Sp(2n, \mathbb{R})$ that leaves invariant a fixed value of the scalar fields $\mathcal{N}$ is $U(n)$. This is easily seen by setting $\mathcal{N} = -\rangle\infty\infty$. Then infinitesimally we have relations (5.55) with $\kappa = 0$ and $b = -c$, $a = -a^t$, i.e. we have the antisymmetric matrix

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} ,$$

$a = -a^t$, $b = b^t$. For finite transformations the $Sp(2n, \mathbb{R})$ relations (5.391) are complemented by

$$A = D \quad , \quad B = -C \ . \tag{5.77}$$

Thus $A - iB$ is a unitary matrix (see also (5.397)). $U(n)$ is the maximal compact subgroup of $Sp(2n, \mathbb{R})$, it is the group of orthogonal and symplectic $2n \times 2n$ matrices.

More in general from Section 3.1 we easily conclude that a necessary and sufficient condition for a theory with just n abelian gauge fields to have $U(n)$ duality symmetry is (cf. (5.60))

$$\tilde{F}^\Lambda F^\Sigma + \tilde{G}^\Lambda G^\Sigma = 0 \tag{5.78}$$

$$\tilde{G}^\Lambda F^\Sigma - \tilde{G}^\Sigma F^\Lambda = 0 \tag{5.79}$$

for all Λ, Σ. Moreover since any nonlinear lagrangian in the limit of weak field strengths $F^4 << F^2$ reduces to the one in (5.63) (with a fixed value of $\mathcal{N}$), we conclude that $U(n)$ is the maximal duality group for a theory with only gauge fields.

Condition (5.79) is equivalent to

$$(F^\Sigma \frac{\partial}{\partial F^\Lambda} - F^\Lambda \frac{\partial}{\partial F^\Sigma})\mathcal{L} = 0 \ , \tag{5.80}$$

i.e. to the invariance of the Lagrangian under $SO(n)$ rotations of the n field strengths F^Σ. Condition (5.78) concerns on the other hand the invariance of the equations of motion under transformation of the electric field strengths into the magnetic field strengths.

In a theory with just n abelian gauge fields the field strengths appear in the Lagrangian only through the Lorentz invariant combinations

$$\alpha^{\Lambda\Sigma} \equiv \frac{1}{4}F^\Lambda F^\Sigma, \quad \beta^{\Lambda\Sigma} \equiv \frac{1}{4}\widetilde{F}^\Lambda F^\Sigma, \tag{5.81}$$

and equation (5.80), tell us that $\mathcal{L}$ is a scalar under $SO(n)$ rotations; e.g. $\mathcal{L}$ is a sum of traces, or of products of traces, of monomials in α and β (we implicitly use the metric $\delta_{\Lambda\Sigma}$ in the α and β products).

If we define

$$\mathcal{L}_\alpha \equiv \frac{\partial \mathcal{L}}{\partial \alpha^t} \ , \quad \mathcal{L}_\beta \equiv \frac{\partial \mathcal{L}}{\partial \beta^t} \ , \tag{5.82}$$

then using the chain rule and the definitions (5.81) we obtain that (5.78) is equivalent to

$$\mathcal{L}_\beta \beta \mathcal{L}_\beta - \mathcal{L}_\alpha \beta \mathcal{L}_\alpha + \mathcal{L}_\alpha \alpha \mathcal{L}_\beta + \mathcal{L}_\beta \alpha \mathcal{L}_\alpha + \beta = 0 \ . \tag{5.83}$$

If we define

$$p \equiv -\frac{1}{2}(\alpha + i\beta) \ , \quad q \equiv -\frac{1}{2}(\alpha - i\beta) \ , \tag{5.84}$$

then (5.83) simplifies and reads

$$p - \mathcal{L}_p \, p \, \mathcal{L}_p = q - \mathcal{L}_q \, q \, \mathcal{L}_q \ . \tag{5.85}$$

Condition (5.83) in the case of a single gauge field was considered in [15] togheter with other equivalent conditions, in particular $\mathcal{L}_u \mathcal{L}_v = 1$, where $u = \frac{1}{2}(\alpha + (\alpha^2 + \beta^2)^{\frac{1}{2}})$, $v = \frac{1}{2}(\alpha - (\alpha^2 + \beta^2)^{\frac{1}{2}})$, see also [20].

Coupling to scalar fields and noncompact duality rotations

By freezing the values of the scalar fields $\mathcal{N}$ we have obtained a theory with only gauge fields and with $U(n)$ duality symmetry. Vice versa (following [16] that extends to $U(n)$ the $U(1)$ interacting theory discussed in [14, 15]) we show that given a theory invariant under $U(n)$ duality rotations it is possible to extend it via $n(n+1)$ scalar fields $\mathcal{N}$ to a theory invariant under $Sp(2n, \mathbb{R})$. Let $\mathcal{L}(F)$ be the lagrangian of the theory with $U(n)$ duality. From (5.59) we see that under a $U(n)$ duality rotation

$$\mathcal{L}(F') - \mathcal{L}(F) = -\frac{1}{4}\tilde{F}\, b\, F + \frac{1}{4}\tilde{G}\, b\, G \ . \tag{5.86}$$

In particular $\mathcal{L}(F)$ is invariant under the orthogonal subgroup $SO(n) \subset U(n)$ given by the matrix $\left(\begin{smallmatrix} A & 0 \\ 0 & -A^t \end{smallmatrix}\right)$. This is the so-called electric subgroup of the duality rotation group $U(n)$ because it does not mix the electric fields F with the dual fields G.

Define the new lagrangian

$$\mathfrak{L}(F, R, \mathcal{N}_\infty) = \mathcal{L}(\mathcal{RF}) + \frac{\infty}{\triangle}\tilde{\mathcal{F}}\mathcal{N}_\infty\mathcal{F} \tag{5.87}$$

where $R = (R^\Lambda{}_\Sigma)_{\Lambda,\Sigma=1,\dots n}$ is an arbitrary nondegenerate real matrix and $\mathcal{N}_\infty$ is a real symmetric matrix. Because of the $O(n)$ symmetry the new lagrangian depends only on the combination

$$\mathcal{N}_\in = -\mathcal{R}^\sqcup\mathcal{R} \ , \tag{5.88}$$

rather than on R. Thus $\mathfrak{L}(F, R, \mathcal{N}_\infty) = \mathfrak{L}(\mathcal{F}, \mathcal{N})$ where $\mathcal{N} = \mathcal{N}_\infty + \rangle\mathcal{N}_\in$.

We show that $\mathfrak{L}$ satisfies the duality condition (5.60),

$$(\Delta_F + \Delta_R + \Delta_{\mathcal{N}_\infty})\mathfrak{L}(F, R, \mathcal{N}_\infty) = \frac{\infty}{\triangle}\tilde{\mathcal{F}}\,\rfloor\,\mathcal{F} + \frac{\infty}{\triangle}\tilde{\mathcal{G}}\,\lfloor\,\mathcal{G} \tag{5.89}$$

where as always $\tilde{G} = 2\frac{\partial\mathfrak{L}}{\partial F}$, and where $\mathcal{N}_\infty$ transforms as in (5.64) and

$$\Delta R = -R(a + b\mathcal{N}_\infty) \ \ , \tag{5.90}$$

so that $\mathcal{N}_\in = -\mathcal{R}^\sqcup\mathcal{R}$ transforms as in (5.65). Notice that we could also have chosen the transformation $\Delta R = \Lambda R - R(a + b\mathcal{N}_\infty)$ with Λ an infinitesimal $SO(n)$ rotation.

We first immediately check (5.89) in the case of the rotation $\left(\begin{smallmatrix} 0 & 0 \\ c & 0 \end{smallmatrix}\right)$. Then in the case $\left(\begin{smallmatrix} a & 0 \\ 0 & d \end{smallmatrix}\right)$, where $a = -d^t$. Finally we consider the duality rotation $\left(\begin{smallmatrix} 0 & b \\ 0 & 0 \end{smallmatrix}\right)$. It is convenient to introduce the notation

$$\mathcal{F} = RF \quad , \quad \mathcal{G} = 2\frac{\partial\mathcal{L}(\mathcal{F})}{\partial\mathcal{F}} \ . \tag{5.91}$$

We observe that $\mathcal{L}(\mathcal{F})$ satisfies the $U(n)$ duality conditions (5.78), (5.79) with $F \to \mathcal{F}$, $G \to \mathcal{G}$. Equation (5.89) holds because of (5.78) and proves $Sp(2n, \mathbb{R})$ duality invariance of the theory with lagrangian $\mathfrak{L}$.

We end this subsection with few comments. We notice that (5.79) is equivalent to the invariance of the lagrangian under the infinitesimal $SO(n)$ transformation $R \to \Lambda R$.

We also observe that under an $Sp(2n, \mathbb{R})$ duality transformation $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, the dressed fields $\mathcal{F}$ and $\mathcal{G}$ transform via the field dependent rotation $\left(\begin{smallmatrix} 0 & b' \\ -b' & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & RbR^t \\ -RbR^t & 0 \end{smallmatrix}\right)$,

$$\Delta\mathcal{F} = \ R\,b\,R^t\,\mathcal{G} \ , \tag{5.92}$$

$$\Delta\mathcal{G} = -R\,b\,R^t\,\mathcal{F} \ . \tag{5.93}$$

The geometry underlying the construction of $Sp(2n, \mathbb{R})$ duality invariant theories from $U(n)$ ones is that of coset spaces. The scalar fields $\mathcal{N}$ parametrize the coset space ${}^{Sp(2n,\mathbb{R})}/_{U(n)}$ (see proof in...). We also have ${}^{Sp(2n,\mathbb{R})}/_{U(n)} = {}_{SO(n)}\backslash^{GL^+(n)} \times \mathbb{R}^{\frac{n(n+1)}{2}}$ where $GL^+(n)$ is the connected component of $GL(n)$ and the equivalence classes $[R] = \{R' \in GL^+(n);\ R'R^{-1} = e^\Lambda \in SO(n)\}$ parametrize the coset space ${}_{SO(n)}\backslash^{GL^+(n)}$.

The proof of $Sp(2n, \mathbb{R})$ duality symmetry for the theory described by the lagrangian $\mathfrak{L}$ holds also if we add to $\mathfrak{L}$ an $Sp(2n, \mathbb{R})$ invariant lagrangian for the fields $\mathcal{N}$ like the lagrangian $\mathscr{L}$ in (5.105). Of course we can also consider initial lagrangians in (5.86) that depend on matter fields invariant under the $U(n)$ rotation, they will be $Sp(2n, \mathbb{R})$ invariant in the corresponding lagrangian $\mathfrak{L}$. Moreover, by considering an extra scalar field Φ, we can always extend an $Sp(2n, \mathbb{R})$ duality theory to an $\mathcal{R}^{>0} \times Sp(2n, \mathbb{R})$ one.

5.3.5 Nonlinear sigma models on G/H

In this section we briefly consider the geometry of coset spaces G/H. This is the geometry underlying the scalar fields and needed to formulate their dynamics [59, 60].

We study in particular the case $G = Sp(2n.\mathbb{R})$, $H = U(n)$ [6] and give a kinetic term for the scalar fields $\mathcal{N}$.

The geometry of the coset space G/H is conveniently described in terms of coset representatives, local sections L of the bundle $G \to G/H$. A point ϕ in G/H is an equivalence class $gH = \{\tilde{g} \,|\, g^{-1}\tilde{g} \in H\}$. We denote by ϕ^i $(i = 1, 2 \ldots m)$ its coordinates (the scalar fields of the theory). The left action of G on G/H is inherited from that of G on G, it is given by $gH \mapsto g'gH$, that we rewrite $\phi \mapsto g'\phi = \phi'$. Concerning the coset representatives we then have

$$g'L(\phi) = L(\phi')h \ , \tag{5.94}$$

because both the left and the right hand side are representatives of ϕ'. The geometry of G/H and the corresponding physics can be constructed in terms of coset representatives. Of course the construction must be insensitive to the particular representative choice, we have a gauge symmetry with gauge group H.

When H is compact the Lie algebra of G splits in the direct sum $\mathbb{G} = \mathbb{H} + \mathbb{K}$, where

$$[\mathbb{H}, \mathbb{H}] \subset \mathbb{H} \quad , \quad [\mathbb{K}, \mathbb{K}] \subset \mathbb{H} + \mathbb{K} \quad , \quad [\mathbb{H}, \mathbb{K}] \subset \mathbb{K} \ . \tag{5.95}$$

The last expression defines the coset space representation of $\mathbb{H}$. The representations of the compact Lie algebra $\mathbb{H}$ are equivalent to unitary ones, and therefore there exists a basis (H_α, K_a), where $[H_\alpha, K_a] = C^b_{\alpha a} K_b$ with $C_\alpha = (C^b_{\alpha a})_{a,b=1,\ldots m=dimG/H}$ antihermitian matrices. Since the coset representation is a real representation then these matrices C_α belong to the Lie algebra of $SO(m)$.

Given a coset representative $L(\phi)$, the pull back on G/H of the $\mathbb{G}$ Lie algebra left invariant 1-form $\Gamma = L^{-1}dL$ is decomposed as

$$\Gamma = L^{-1}dL = P^a(\phi)K_a + \omega^\alpha(\phi)H_\alpha \ .$$

Γ and therefore $P = P^a(\phi)K_a$ and $\omega = \omega^\alpha(\phi)H_\alpha$ are invariant under diffeomorphisms generated by the left G action. Under the local right H action of an element $h(\phi)$ (or under change of coset representative $L'(\phi) = L(\phi)h(\phi)$) we have

$$P \to h^{-1}Ph \quad , \quad \omega \to h^{-1}\omega h + h^{-1}dh \ . \tag{5.96}$$

The 1-forms $P^a(\phi) = P^a(\phi)_i d\phi^i$ are therefore vielbain on G/H transforming in the fundamental of $SO(m)$, while $\omega = \omega(\phi)_i d\phi^i$ is an $\mathbb{H}$-valued connection 1-form on G/H. We can then define the covariant derivative $\nabla P^a = [P, \omega]^a = P^b \otimes -C^a_{\alpha b}\omega^\alpha$.

There is a natural metric on G/H,

$$g = \delta_{ab} P^a \otimes P^b \ , \tag{5.97}$$

(this definition is well given because we have shown that the coset representation is via infiniesimal $SO(m)$ rotations). It is easy to see that the connection ∇ is metric compatible, $\nabla g = 0$.

If the coset is furthermore a symmetric coset we have

$$[\mathbb{K}, \mathbb{K}] \subset \mathbb{H} \ ,$$

then the identity $d\Gamma+\Gamma\wedge\Gamma=0$, that is (the pull-back on G/H of) the Maurer-Cartan equation, in terms of P and ω reads

$$R + P \wedge P = 0 \; , \tag{5.98}$$

$$dP + P \wedge \omega + \omega \wedge P = 0 \; . \tag{5.99}$$

This last relation shows that ω is torsionfree. Since it is metric compatible it is therefore the Riemannian connection on G/H. Equation (5.98) then relates the Riemannian curvature to the square of the vielbeins.

By using the connection ω and the vierbein P we can construct couplings and actions invariant under the rigid G and the local H transformations, i.e. sigma models on the coset space G/H.

For example a kinetic term for the scalar fields, which are maps from spacetime to G/H, is given by pulling back to spacetime the invariant metric (5.97) and then contracting it with the spacetime metric

$$\mathscr{L}_{\mathrm{kin}}(\phi) = \frac{1}{2} P^a_\mu P^\mu_a = \frac{1}{2} P^a_i \partial_\mu \phi^i P_{aj} \partial^\mu \phi^j \; . \tag{5.100}$$

By construction the lagrangian $\mathscr{L}_{\mathrm{kin}}(\phi)$ is invariant under G and local H transformations; it depends only on the coordinates of the coset space G/H.

The case $G = Sp(2n, \mathbb{R})$, $H = U(n)$

A kinetic term for the $\frac{Sp(2n,\mathbb{R})}{U(n)}$ valued scalar fields is given by (5.100). This lagrangian is invariant under $Sp(2n, \mathbb{R})$ and therefore satisfies the duality condition (5.66) with $G = Sp(2n, \mathbb{R})$ and $\kappa = 0$. We can also write

$$\mathscr{L}_{\mathrm{kin}}(\phi) = \frac{1}{2} P^a_\mu P^\mu_a = \frac{1}{2} \mathrm{Tr}(P_\mu P^\mu) \; ; \tag{5.101}$$

where in the last passage we have considered generators K_a so that $\mathrm{Tr}(K_a K_b) = \delta_{ab}$ (this is doable since $U(n)$ is the maximal compact subgroup of $Sp(2n, \mathbb{R})$).

We now recall the representation of the group $Sp(2n.\mathbb{R})$ and of the associated coset $\frac{Sp(2n,\mathbb{R})}{U(n)}$ in the complex basis discussed in the appendix (and frequently used in the later sections) and we give a more explicit expression for the lagrangian (5.101).

Rather than using the symplectic matrix $S = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ of the fundamental representation of $Sp(2n, \mathbb{R})$, we consider the conjugate matrix $\mathcal{A}^{-1} S \mathcal{A}$ where $\mathcal{A} = \frac{1}{\sqrt{2}} \left(\begin{smallmatrix} \mathbb{1} & \mathbb{1} \\ -i\mathbb{1} & i\mathbb{1} \end{smallmatrix}\right)$. In this complex basis the subgroup $U(n) \subset Sp(2n, \mathbb{R})$ is simply given by the block diagonal matrices $\left(\begin{smallmatrix} u & 0 \\ 0 & \bar{u} \end{smallmatrix}\right)$. We also define the $n \times 2n$ matrix

$$\begin{pmatrix} f \\ h \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} A - iB \\ C - iD \end{pmatrix} \tag{5.102}$$

and the matrix

$$V = \begin{pmatrix} f & \bar{f} \\ h & \bar{h} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \mathcal{A} \; . \tag{5.103}$$

Then (cf. (5.398), (5.399)),

$$V^{-1} dV = \begin{pmatrix} i(f^\dagger dh - h^\dagger df) & i(f^\dagger d\bar{h} - h^\dagger d\bar{f}) \\ -i(f^t dh - h^t df) & -i(f^t d\bar{h} - h^t d\bar{f}) \end{pmatrix} \equiv \begin{pmatrix} \omega & \bar{\mathcal{P}} \\ \mathcal{P} & \bar{\omega} \end{pmatrix} \; , \tag{5.104}$$

where in the last passage we have defined the $n \times n$ sub-blocks ω and $\mathcal{P}$ corresponding to the $U(n)$ connection and the vielbein of $Sp(2n, \mathbb{R})/U(n)$ in the complex basis, (with slight abuse of notation we use the same letter ω in this basis too).

We finally obtain the explicit expression

$$\mathscr{L}_{\rm kin}(\phi) = {\rm Tr}(\bar{\mathcal{P}}_\mu \mathcal{P}^\mu) = \frac{1}{4}{\rm Tr}(\mathcal{N}_\in^{-\infty}\partial_\mu\overline{\mathcal{N}}\ \mathcal{N}_\in^{-\infty}\partial^\mu\mathcal{N}) \tag{5.105}$$

where $\mathcal{P} = \mathcal{P}_\mu dx^\mu = \mathcal{P}_i\partial_\mu\phi^i dx^\mu$, $\overline{\mathcal{N}} = \mathcal{N}_\infty - \rangle\mathcal{N}_\in$ and $\mathcal{N} = \mathcal{N}_\infty + \rangle\mathcal{N}_\in = {\rm Re}\mathcal{N} + \rangle{\rm Im}N$. The matrix of scalars $\mathcal{N}$ parametrizes the coset space $Sp(2n,\mathbb{R})/U(n)$ (see Appendix 7.2); in terms of the f and h matrices it is given by (cf. (5.408))

$$\mathcal{N} = \{\,\langle^{-\infty}\ ,\quad \mathcal{N}_\in^{-\infty} = -\in\{\,\{^\dagger\ . \tag{5.106}$$

Under the symplectic rotation $\left(\begin{smallmatrix}A\,B\\C\,D\end{smallmatrix}\right) \to \left(\begin{smallmatrix}A'\,B'\\C'\,D'\end{smallmatrix}\right)\left(\begin{smallmatrix}A\,B\\C\,D\end{smallmatrix}\right)$ the matrix $\mathcal{N}$ changes via the fractional transformation $\mathcal{N} \to (\mathcal{C}' + \mathcal{D}'\mathcal{N})\,(\mathcal{A}' + \mathcal{B}'\mathcal{N})^{-\infty}$, (cf. (5.68)).

Another proof of the invariance of the kinetic term (5.105) under the $Sp(2n,\mathbb{R})$ follows by observing that (5.105) is obtained from the pullback to the spacetime manifold of the metric associated to the $\frac{Sp(2n,\mathbb{R})}{U(n)}$ Kform ${\rm Tr}(\mathcal{N}_\in^{-\infty}{\rm d}\,\overline{\mathcal{N}}\ \mathcal{N}_\in^{-\infty}{\rm d}\,\mathcal{N})$ (here ${\rm d} = \partial + \overline{\partial}$ is the exterior derivative). This metric is obtained from the Kpotential

$$\mathcal{K} = -4\,{\rm Tr}\log i(\mathcal{N} - \overline{\mathcal{N}})\ . \tag{5.107}$$

Under the action of $Sp(2n,\mathbb{R})$, $\mathcal{N}$ and $\mathcal{N} - \overline{\mathcal{N}}$ change as in (5.68), (5.69) and the Kpotential changes by a Ktransformation, thus showing the invariance of the metric.

3.4.2 The case $G = \mathbb{R}^{>0} \times Sp(2n,\mathbb{R})$, $H = U(n)$

In this case the duality rotation matrix $\left(\begin{smallmatrix}a\,b\\c\,d\end{smallmatrix}\right)$ belongs to the Lie algebra of $\mathbb{R}^{>0} \times Sp(2n,\mathbb{R})$, as defined in (5.55). In particular infinitesimal dilatations are given by the matrix $\frac{\kappa}{2}\left(\begin{smallmatrix}\mathbb{1}\,0\\0\,\mathbb{1}\end{smallmatrix}\right)$. The coset space is

$$\frac{\mathbb{R}^{>0} \times Sp(2n,\mathbb{R})}{U(n)} = \mathbb{R}^{>0}\times\frac{Sp(2n,\mathbb{R})}{U(n)}\ , \tag{5.108}$$

there is no action of $U(n)$ on $\mathbb{R}^{>0}$. We consider a real positive scalar field $\Phi = e^\sigma$ invariant under $Sp(2n,\mathbb{R})$ transformations. The fields Φ and $\mathcal{N}$ parametrize the coset space (5.108).

Let's first consider the main example of Section 3.2. The duality symmetry conditions for the lagrangian (5.63) are (5.64)-(5.66). >From equations (5.64),(5.65) (that hold for $\left(\begin{smallmatrix}a\,b\\c\,d\end{smallmatrix}\right)$ in the Lie algebra of $\mathbb{R}^{>0} \times Sp(2n,\mathbb{R})$) we see that the fields $\mathcal{N}$, and henceforth the lagrangian $\mathscr{L}_{\rm kin}(\phi)$, are invariant under the $\mathbb{R}^{>0}$ action. It follows that the scalar lagrangian

$$\Phi^2\mathscr{L}_{\rm kin}(\phi) + \partial_\mu\Phi\partial^\mu\Phi \tag{5.109}$$

satisfies the duality condition (5.66). This shows that the lagrangian (5.63) with the scalar kinetic term given by (5.109) has $\mathbb{R}^{>0} \times Sp(2n,\mathbb{R})$ duality symmetry. We see that in the lagrangian (5.63) the scalar Φ does not couple to the field strenght F. The coupling of Φ to F is however present in lagrangians where higher powers of F are present.

More in general expression (5.109) is a scalar kinetic term for lagrangians that satisfy the $\mathbb{R}^{>0} \times Sp(2n,\mathbb{R})$ duality condition (5.60).

5.3.6 Invariance of energy momentum tensor

Duality rotation symmetry is a symmetry of the equations of motion that does not leave invariant the lagrangian. The total change $\Delta\mathcal{L} \equiv \mathcal{L}(F',\varphi') - \mathcal{L}(F,\varphi)$ of the lagrangian is given in equation (5.59). Even if $\kappa = 0$ this variation is not a total derivative because F and G are the curl of vector potentials A_F and A_G only on shell.

We show however that the variation of the action with respect to a duality rotation invariant parameter λ is invariant under $Sp(2n, \mathbb{R})$ rotations if the duality rotation (5.50) of the φ fields is λ independent.

Consider the λ-variation of $\Delta S[F, \varphi] \equiv S[F', \varphi'] - S[F, \varphi] = \int_y \frac{\partial \mathcal{L}}{\partial F} \Delta F + \Delta_\varphi S$,

$$\begin{aligned}\frac{\delta}{\delta\lambda}\Delta S &= \int_y \frac{\delta}{\delta\lambda}(\frac{\partial\mathcal{L}}{\partial F})\;\Delta F + \int_y \frac{\partial\mathcal{L}}{\partial F}\frac{\delta}{\delta\lambda}(\Delta F) + \frac{\delta}{\delta\lambda}(\Delta_\varphi S)\\ &= \int_y \frac{\partial}{\partial F}(\frac{\delta\mathcal{L}}{\delta\lambda})\;\Delta F + \frac{1}{2}\int_y \tilde{G}\frac{\delta}{\delta\lambda}(\Delta F) + \Delta_\varphi(\frac{\delta S}{\delta\lambda})\\ &= \Delta(\frac{\delta S}{\delta\lambda}) + \frac{1}{4}\frac{\delta}{\delta\lambda}\int_y \tilde{G}\, b\, G \end{aligned} \tag{5.110}$$

where in the second line we used that $\frac{\delta}{\delta\lambda}\Delta\varphi = 0$. Thus $\Delta(\frac{\delta S}{\delta\lambda}) = \frac{\delta}{\delta\lambda}(\Delta S - \frac{1}{4}\int_y \tilde{G}\, b\, G)$ and therefore from (5.59) we have,

$$\Delta(\frac{\delta S}{\delta\lambda}) = \kappa\,\frac{\delta S}{\delta\lambda} \tag{5.111}$$

thus showing invariance of $\frac{\delta S}{\delta\lambda}$ under $Sp(2n, \mathbb{R})$ rotations ($\kappa = 0$ rotations).

An important case is when λ is the metric $g_{\mu\nu}$, this is invariant under duality rotations. This shows that the energy momentum tensor $\frac{\delta S}{\delta g_{\mu\nu}}$ is invariant under $Sp(2n, \mathbb{R})$ duality rotations.

Another instance is when λ is the dimensional parameter typically present in a nonlinear theory. Provided the matter fields are properly rescaled $\varphi \to \hat{\varphi} = \lambda^s\varphi$, so that they become adimensional and therefore their transformation $\Delta\hat{\varphi}$, usually nonlinear, does not explicitly involve λ, then $\frac{\delta S}{\delta\lambda}$ is invariant, where it is understood that $\frac{\partial\hat{\varphi}}{\partial\lambda} = 0$.

For the action of the Born-Infeld theory coupled to the axion and dilaton fields, $\mathcal{L} = \frac{1}{\lambda}\big(1 - \sqrt{1 - \frac{1}{2}\lambda\mathcal{N}_\in\mathcal{F}^\in - \frac{\infty}{\infty/}\lambda^\in\mathcal{N}_\in(\mathcal{F}\tilde{\mathcal{F}})^\in}\,\big)$ we obtain the invariant $\frac{\partial\mathcal{L}}{\partial\lambda} = -\frac{1}{\lambda}(\mathcal{L} - \frac{1}{4}F\tilde{G})$; we already found this invariant in (5.62).

5.3.7 Generalized Born Infeld theory

In this section we present the Born-Infeld theory with n abelian gauge fields coupled to $n(n+1)/2$ scalar fields $\mathcal{N}$ and show that is has an $Sp(2n, \mathbb{R})$ duality symmetry. If we freeze the scalar fields $\mathcal{N}$ to the value $\mathcal{N} = -\rangle\infty\!\!\infty$ then the lagrangian has $U(n)$ duality symmetry and reads

$$\mathcal{L} = \mathrm{Tr}[\mathbb{1} - \mathcal{S}_{\alpha,\beta}\sqrt{\mathbb{1} + 2\alpha - \beta^2}\,]\ , \tag{5.112}$$

where as defined in (5.81), the componets of the $n \times n$ matrices α and β are $\alpha^{\Lambda\Sigma} = \frac{1}{4}F^\Lambda F^\Sigma$, $\beta^{\Lambda\Sigma} = \frac{1}{4}\widetilde{F}^\Lambda F^\Sigma$. The square root is to be understood in terms of its power series expansion, and the operator $\mathcal{S}_{\alpha,\beta}$ acts by symmetrizing each monomial in the α and β matrices. A world (monomial) in the letters α and β is symmetrized by averaging over all permutations of its letters. The normalization of $\mathcal{S}_{\alpha,\beta}$ is such that if α and β commute then $\mathcal{S}_{\alpha,\beta}$ acts as the identity. Therefore in the case of just one abelian gauge field (5.112) reduces to the usual Born-Infeld lagrangian.

The $Sp(2n, \mathbb{R})$ Born-Infeld lagrangian is obtained by coupling the lagrangian (5.112) to the scalar fields $\mathcal{N}$ as described in Subsection 5.3.4 and explicitly considered in (5.149).

Following [18] we prove the duality symmetry of the Born-Infeld theory (5.112) by first showing that a Born-Infeld theory with n *complex* abelian gauge fields written in an auxiliary field formulation has $U(n, n)$ duality symmetry. We then eliminate the auxiliary fields by proving a remarkable property of solutions of matrix equations [19]. Then we can consider real fields.

Duality rotations with complex field strengths

>From the general study of duality rotations we know that a theory with $2n$ real fields F_1^Λ and F_2^Λ ($\Lambda = 1, \dots n$) has at most $Sp(4n, \mathbb{R})$ duality if we consider duality rotations that leave invariant the energy-momentum tensor (and in particular the hamiltonian). We now consider the complex fields

$$F^\Lambda = F_1^\Lambda + iF_2^\Lambda \quad , \qquad \bar{F}^\Lambda = F_1^\Lambda - iF_2^\Lambda \ , \tag{5.113}$$

the corresponding dual fields

$$G = \frac{1}{2}(G_1 + iG_2) \quad , \qquad \bar{G} = \frac{1}{2}(G_1 - iG_2) \ , \tag{5.114}$$

and restrict the $Sp(4n, \mathbb{R})$ duality group to the subgroup of *holomorphic* transformations,

$$\Delta \begin{pmatrix} F \\ G \end{pmatrix} = \begin{pmatrix} \mathsf{a} & \mathsf{b} \\ \mathsf{c} & \mathsf{d} \end{pmatrix} \begin{pmatrix} F \\ G \end{pmatrix} \tag{5.115}$$

$$\Delta \begin{pmatrix} \bar{F} \\ \bar{G} \end{pmatrix} = \begin{pmatrix} \bar{\mathsf{a}} & \bar{\mathsf{b}} \\ \bar{\mathsf{c}} & \bar{\mathsf{d}} \end{pmatrix} \begin{pmatrix} \bar{F} \\ \bar{G} \end{pmatrix} \ . \tag{5.116}$$

This requirement singles out those matrices, acting on the vector $\begin{pmatrix} F_1 \\ F_2 \\ G_1 \\ G_2 \end{pmatrix}$, that belong to the Lie algebra of $Sp(4n, \mathbb{R})$ and have the form

$$\begin{pmatrix} \bar{A}\left(\begin{smallmatrix} \mathsf{a} & 0 \\ 0 & \bar{\mathsf{a}} \end{smallmatrix}\right)\bar{A}^{-1} & \frac{1}{2}\bar{A}\left(\begin{smallmatrix} \mathsf{b} & 0 \\ 0 & \bar{\mathsf{b}} \end{smallmatrix}\right)\bar{A}^{-1} \\ \\ 2\bar{A}\left(\begin{smallmatrix} \mathsf{c} & 0 \\ 0 & \bar{\mathsf{c}} \end{smallmatrix}\right)\bar{A}^{-1} & \bar{A}\left(\begin{smallmatrix} \mathsf{d} & 0 \\ 0 & \bar{\mathsf{d}} \end{smallmatrix}\right)\bar{A}^{-1} \end{pmatrix} \tag{5.117}$$

where $\mathcal{A} = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix} \mathbb{1} & \mathbb{1} \\ -i\mathbb{1} & i\mathbb{1} \end{smallmatrix}\right)$. The matrix (5.117) belongs to $Sp(4n, \mathbb{R})$ iff the $n \times n$ complex matrices $\mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{d}$ satisfy

$$\mathsf{a}^\dagger = -\mathsf{a} \ , \quad \mathsf{b}^\dagger = \mathsf{b} \ , \quad \mathsf{c}^\dagger = \mathsf{c} \ . \tag{5.118}$$

Matrices $\begin{pmatrix} \mathsf{a} & \mathsf{b} \\ \mathsf{c} & \mathsf{d} \end{pmatrix}$, that satisfy (5.118), define the Lie algebra of the real form $U(n, n)$. The group $U(n, n)$ is here the subgroup of $GL(2n, \mathbb{C})$ caracterized by the relations*

$$M^\dagger \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} M = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \ . \tag{5.119}$$

One can check that (5.119) implies the following relations for the block components of $M = \begin{pmatrix} \mathsf{A} & \mathsf{B} \\ \mathsf{C} & \mathsf{D} \end{pmatrix}$,

$$\mathsf{C}^\dagger \mathsf{A} = \mathsf{A}^\dagger \mathsf{C} \ , \quad \mathsf{B}^\dagger \mathsf{D} = \mathsf{D}^\dagger \mathsf{B} \ , \quad \mathsf{D}^\dagger \mathsf{A} - \mathsf{B}^\dagger \mathsf{C} = \mathbb{1} \ . \tag{5.120}$$

The Lie algebra relations (5.118) can be obtained from the Lie group relations (5.120) by writing $\left(\begin{smallmatrix} \mathsf{A} & \mathsf{C} \\ \mathsf{B} & \mathsf{D} \end{smallmatrix}\right) = \left(\begin{smallmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{smallmatrix}\right) + \epsilon\left(\begin{smallmatrix} \mathsf{a} & \mathsf{b} \\ \mathsf{c} & \mathsf{d} \end{smallmatrix}\right)$ with ϵ infinitesimal. Equation (5.117) gives the embedding of $U(n, n)$ in $Sp(4n, \mathbb{R})$.

The theory of holomorphic duality rotations can be seen as a special case of that of real duality rotations, but (as complex geometry versus real geometry) it deserves also an independent formulation based on the holomorphic variables $\begin{pmatrix} F \\ G \end{pmatrix}$ and maps $\begin{pmatrix} \mathsf{a} & \mathsf{b} \\ \mathsf{c} & \mathsf{d} \end{pmatrix}$.

*In Appendix 7.1 we define $U(n, n)$ as the group of complex matrices that satisfy the condition $U^\dagger\left(\begin{smallmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{smallmatrix}\right)U = \left(\begin{smallmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{smallmatrix}\right)$. The similarity transformation between these two definitions is $M = \bar{A}U\bar{A}^{-1}$.

The dual fields in (5.114), or rather the Hodge dual of the dual field strength, $\widetilde{G}_\Lambda{}^{\mu\nu} = \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}G_\Lambda{}^{\rho\sigma}$, is equivalently defined via

$$\tilde{G}_\Lambda{}^{\mu\nu} \equiv 2\frac{\partial\mathcal{L}}{\partial\bar{F}^\Lambda_{\mu\nu}}\ , \quad \tilde{\bar{G}}_\Lambda{}^{\mu\nu} \equiv 2\frac{\partial\mathcal{L}}{\partial F^\Lambda_{\mu\nu}}\ . \tag{5.121}$$

Repeating the passages of Section 3.1 we have that the Bianchi identities and equations of motion $\partial_\mu\tilde{F}^{\Lambda\,\mu\nu} = 0\ ,\ \ \partial_\mu\tilde{G}_\Lambda{}^{\mu\nu} = 0\ ,\ \ \frac{\delta S[F,\bar{F},\varphi]}{\delta\varphi^\alpha} = 0$ transform covariantly under the holomorphic infinitesimal transformations (5.115) if the lagrangian satisfies the condition (cf. (5.59))

$$\mathcal{L}(F+\Delta F,\bar{F}+\Delta\bar{F},\varphi+\Delta\varphi) - \mathcal{L}(F,\bar{F},\varphi) - \frac{1}{2}\tilde{F}\,\mathsf{c}\,\bar{F} - \frac{1}{2}\tilde{G}\,\mathsf{b}\,\bar{G} \ = \ const_{\mathsf{a,b,c,d}} \tag{5.122}$$

Of course we can also consider dilatations $\kappa \neq 0$, then in the left hand side of (5.122) we have to add the term $-\kappa\mathcal{L}(F,\bar{F},\varphi)$.

The maximal compact subgroup of $U(n,n)$ is $U(n)\times U(n)$ and is obtained by requiring (5.120) and

$$\mathsf{A} = \mathsf{D}\ ,\quad \mathsf{B} = -\mathsf{C}\ .$$

The corresponding infinitesimal relations are (5.118) and $\mathsf{a} = \mathsf{d}\ ,\ \mathsf{b} = -\mathsf{c}\ .$

The coset space $\frac{U(n,n)}{U(n)\times U(n)}$ is the space of all negative definite hermitian matrices M^* of $U(n,n)$, see for example [18] (the proof is similar to that for $Sp(2n,\mathbb{R})/U(n)$ in Appendix 7.2). All these matrices are for example of the form $M^* = -g^{\dagger^{-1}}g^{-1}$ with $g\in U(n,n)$. These matrices can be factorized as

$$\begin{aligned} M^* &= \begin{pmatrix} \mathbb{1} & -\mathcal{N}_\infty \\ 0 & \mathbb{1} \end{pmatrix}\begin{pmatrix} \mathcal{N}_\in & 0 \\ 0 & \mathcal{N}_\in^{-\infty} \end{pmatrix}\begin{pmatrix} \mathbb{1} & 0 \\ -\mathcal{N}_\infty^\dagger & \mathbb{1} \end{pmatrix} \\ &= \begin{pmatrix} \mathcal{N}_\in + \mathcal{N}_\infty\,\mathcal{N}_\in^{-\infty}\,\mathcal{N}_\infty^\dagger & -\mathcal{N}_\infty\,\mathcal{N}_\in^{-\infty} \\ -\mathcal{N}_\in^{-\infty}\,\mathcal{N}_\infty^\dagger & \mathcal{N}_\in^{-\infty} \end{pmatrix} \\ &= -i\begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} + \begin{pmatrix} \mathcal{N}\,\mathrm{Im}\,\mathcal{N}^{-\infty}\,\mathcal{N}^\dagger & -\mathcal{N}\,\mathrm{Im}\,\mathcal{N}^{-\infty} \\ -\mathrm{Im}\,\mathcal{N}^{-\infty}\,\mathcal{N}^\dagger & \mathrm{Im}\,\mathcal{N}^{-\infty} \end{pmatrix} \\ &= -i\begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} + \begin{pmatrix} \mathcal{N} & 0 \\ -\mathbb{1} & 0 \end{pmatrix}\begin{pmatrix} \mathcal{N}_\in & 0 \\ 0 & \mathcal{N}_\in^{-\infty} \end{pmatrix}\begin{pmatrix} \mathcal{N}^\dagger & -\mathbb{1} \\ 0 & 0 \end{pmatrix} \end{aligned} \tag{5.123}$$

where $\mathcal{N}_\infty$ is hermitian, $\mathcal{N}_\in$ is hermitian and negative definite, and

$$\mathcal{N} \equiv \mathcal{N}_\infty + \rangle\mathcal{N}_\in\ . \tag{5.124}$$

Since any complex matrix can always be decomposed into hermitian matrices as in (5.124), the only requirement on $\mathcal{N}$ is that $\mathcal{N}_\in$ is negative definite.

The left action of $U(n,n)$ on itself $g \to \left(\begin{smallmatrix} \mathsf{A} & \mathsf{B} \\ \mathsf{C} & \mathsf{D} \end{smallmatrix}\right)g$, induces the action on the coset space $M^* \to \left(\begin{smallmatrix} \mathsf{D} & -\mathsf{C} \\ -\mathsf{B} & \mathsf{A} \end{smallmatrix}\right)M^*\left(\begin{smallmatrix} \mathsf{D} & -\mathsf{C} \\ -\mathsf{B} & \mathsf{A} \end{smallmatrix}\right)^\dagger$ because $M^* = -g^{\dagger^{-1}}g^{-1}$. Expression (5.123) then immediately gives the action of $U(n,n)$ on the parametrization $\mathcal{N}$ of the coset space,

$$\mathcal{N} \to \mathcal{N}' = (\mathsf{C}+\mathsf{D}\mathcal{N})\,(\mathsf{A}+\mathsf{B}\mathcal{N})^{-\infty}\ , \tag{5.125}$$

$$\mathcal{N}_\in \to \mathcal{N}_\in' = (\mathsf{A}+\mathsf{B}\mathcal{N})^{-\dagger}\mathcal{N}_\in(\mathsf{A}+\mathsf{B}\mathcal{N})^{-\infty}\ . \tag{5.126}$$

As in Section 3.4, given a theory depending on n complex fields F^Λ and invariant under the maximal compact duality group $U(n)\times U(n)$ it is possible to extend it via the complex scalar fields $\mathcal{N}$, to a theory invariant under $U(n,n)$. The new lagrangian is

$$\mathfrak{L}(F,R,\mathcal{N}_\infty) = \mathcal{L}(\mathcal{R}\mathcal{F}) + \frac{\infty}{\in}\tilde{\mathcal{F}}\mathcal{N}_\infty\bar{\mathcal{F}} \tag{5.127}$$

where $R = (R^\Lambda{}_\Sigma)_{\Lambda,\Sigma=1,\dots n}$ is now an arbitrary nondegenerate complex matrix. Because of the $U(n)$ maximal compact electric subgroup this new lagrangian depends only on the combination

$$\mathcal{N}_\in = -\mathcal{R}^\dagger \mathcal{R} \ , \tag{5.128}$$

rather than on R. Thus $\mathfrak{L}(F, R, \mathcal{N}_\infty) = \mathfrak{L}(\mathcal{F}, \mathcal{N})$ where $\mathcal{N} = \mathcal{N}_\infty + \rangle \mathcal{N}_\in$. A transformation for R compatible with (5.125) is

$$R' = R(\mathsf{A} + \mathsf{B}\mathcal{N})^{-\infty}, \tag{5.129}$$

whose infinitesimal transformation is $\Delta R = -R(\mathsf{a} + \mathsf{b}\mathcal{N})$.

Conversely, if we are given a Lagrangian $\mathcal{L}$ with equations of motion invariant under $U(n, n)$ we can obtain a theory without the scalar field $\mathcal{N}$ by setting $\mathcal{N} = -\rangle\infty\!\!\infty$. Then the duality group is broken to the stability group of $\mathcal{N} = -\rangle\infty\!\!\infty$ which is $U(n) \times U(n)$, the maximal compact subgroup.

Similarly to Section 3.4.1 we define the Lorentz invariant combinations

$$\alpha^{ab} \equiv \frac{1}{2} F^a \bar{F}^b, \quad \beta^{ab} \equiv \frac{1}{2} \widetilde{F}^a \bar{F}^b. \tag{5.130}$$

If we consider lagrangians $\mathcal{L}(F, \bar{F})$ that depend only on gauge fields and only through sum of traces (or of products of traces) of monomials in α and β, then the necessary and sufficient condition for $U(n) \times U(n)$ holomorphic duality symmetry is still (5.83), where now α and β are as in (5.130).

Born-Infeld with auxiliary fields

A lagrangian that satisfies condition (5.122) is

$$\mathcal{L} = \mathrm{Re}\,\mathrm{Tr}\,[\, i(\mathcal{N} - \lambda)\chi - \frac{\rangle}{\in}\lambda\chi^\dagger \mathcal{N}_\in \chi - \rangle\lambda(\alpha + \rangle\beta)\,] \ , \tag{5.131}$$

The auxiliary fields χ and λ and the scalar field $\mathcal{N}$ are n dimensional complex matrices. We can also add to the lagrangian a duality invariant kinetic term for the scalar field $\mathcal{N}$, (cf (5.105))

$$\mathrm{Tr}(\mathcal{N}_\in^{-\infty}\partial_\mu \mathcal{N}^\dagger \ \mathcal{N}_\in^{-\infty}\partial^\mu \mathcal{N}) \ . \tag{5.132}$$

In order to prove the duality of (5.131) we first note that the last term in the Lagrangian can be written as

$$-\mathrm{Re}\,\mathrm{Tr}\,[\, i\lambda(\alpha + i\beta)\,] = -\mathrm{Tr}(\lambda_2\alpha + \lambda_1\beta) \ .$$

If the field λ transforms by fractional transformation and λ_1, λ_2 and the gauge fields are real this is the $U(1)^n$ Maxwell action (5.63), with the gauge fields interacting with the scalar field λ. This term by itself has the correct transformation properties under the duality group. Similarly for hermitian α, β, λ_1 and λ_2 this term by itself satisfies equation (5.122). It follows that the rest of the Lagrangian must be duality invariant. The duality transformations of the scalar and auxiliary fields are*

$$\lambda' = (\mathsf{C} + \mathsf{D}\lambda)\,(\mathsf{A} + \mathsf{B}\lambda)^{-1} \ , \tag{5.133}$$

$$\chi' = (\mathsf{A} + \mathsf{B}\mathcal{N})\chi(\mathsf{A} + \mathsf{B}\lambda^\dagger)^\dagger \ , \tag{5.134}$$

and (5.125). Invariance of $\mathrm{Tr}[i(\mathcal{N} - \lambda)\chi]$ is easily proven by using (5.120) and by rewriting (5.133) as

$$\lambda' = (\mathsf{A} + \mathsf{B}\lambda^\dagger)^{-\dagger}\,(\mathsf{C} + \mathsf{D}\lambda^\dagger)^\dagger \ . \tag{5.135}$$

Invariance of the remaining term which we write as $\mathrm{Re}\,\mathrm{Tr}\,[-\frac{i}{2}\lambda\chi^\dagger\mathcal{N}_\in\chi] = \mathrm{Tr}\,[\frac{\infty}{\in}\lambda_\in\chi^\dagger\mathcal{N}_\in\chi]$, is straightforward by using (5.126) and the following transformation obtained from (5.135),

$$\lambda_2' = (\mathsf{A} + \mathsf{B}\lambda^\dagger)^{-\dagger}\lambda_2(\mathsf{A} + \mathsf{B}\lambda^\dagger)^{-1} \ . \tag{5.136}$$

*In [18] we use different notations: $\mathcal{N} \to \mathcal{S}^\dagger, \lambda \to \lambda^\dagger, \chi \to \chi^\dagger, \left(\begin{smallmatrix} \mathsf{A} & \mathsf{B} \\ \mathsf{C} & \mathsf{D} \end{smallmatrix}\right) \to \left(\begin{smallmatrix} \mathsf{D} & \mathsf{C} \\ \mathsf{B} & \mathsf{A} \end{smallmatrix}\right)$.

Elimination of the Auxiliary Fields

The equation of motion obtained by varying λ gives an equation for χ,

$$\chi + \frac{1}{2}\chi^{\dagger}\mathcal{N}_{\epsilon}\chi + \alpha + \rangle\beta = \prime\ , \tag{5.137}$$

using this equation in the Lagrangian (5.131) we obtain

$$\begin{aligned} \mathcal{L} &= \mathrm{Re}\,\mathrm{Tr}\,(i\mathcal{N}\chi) & (5.138)\\ &= \mathrm{Re}\,\mathrm{Tr}\,(-\mathcal{N}_{\epsilon}\chi) + \mathrm{Tr}\,(\mathcal{N}_{\infty}\beta)\ , & (5.139) \end{aligned}$$

where χ is now a function of α, β and $\mathcal{N}_{\epsilon}$ that solves (5.137). In the second line we observed that the anti-hermitian part of (5.137) implies $\chi_2 = -\beta$.

In this subsection we give the explicit expression of $\mathcal{L}$ in terms of α, β and $\mathcal{N}$.

First notice that (5.137) can be simplified with the following field redefinitions

$$\begin{aligned} \widehat{\chi} &= R\chi R^{\dagger}\ ,\\ \widehat{\alpha} &= R\alpha R^{\dagger}\ ,\\ \widehat{\beta} &= R\beta R^{\dagger}\ , \end{aligned} \tag{5.140}$$

where, as in (5.128), $R^{\dagger}R = -\mathcal{N}_{\epsilon}$. The equation of motion for χ is then equivalent to

$$\widehat{\chi} - \frac{1}{2}\widehat{\chi}^{\dagger}\widehat{\chi} + \widehat{\alpha} - i\widehat{\beta} = 0\ . \tag{5.141}$$

The anti-hermitian part of (5.141) implies $\widehat{\chi}_2 = -\widehat{\beta}$, thus $\widehat{\chi}^{\dagger} = \widehat{\chi} - 2i\beta$. This can be used to eliminate $\widehat{\chi}^{\dagger}$ from (5.141) and obtain a quadratic equation for $\widehat{\chi}$. If we define $Q = \frac{1}{2}\widehat{\chi}$ this equation reads

$$Q = q + (p-q)Q + Q^2, \tag{5.142}$$

where

$$p \equiv -\frac{1}{2}(\alpha + i\beta)\ ,\quad q \equiv -\frac{1}{2}(\alpha - i\beta)\ .$$

The lagrangian is then

$$\mathcal{L} = 2\,\mathrm{Re}\,\mathrm{Tr}\,Q + \mathrm{Tr}\,(\mathcal{N}_{\infty}\beta)\ . \tag{5.143}$$

If the degree of the matrices is one, we can solve for Q in the quadratic equation (5.142). Apart from the fact that the gauge fields are complex, the result is the Born-Infeld Lagrangian coupled to the dilaton and axion fields $\mathcal{N}$,

$$\mathcal{L} = 1 - \sqrt{1 - 2\mathcal{N}_{\epsilon}\alpha + \mathcal{N}_{\epsilon}^{\epsilon}\beta^{\epsilon}} + \mathcal{N}_{\infty}\beta\ . \tag{5.144}$$

For matrices of higher degree, equation (5.142) can be solved perturbatively,

$$Q_0 = 0\ ,\quad Q_{k+1} = q + (p-q)Q_k + Q_k^2\ , \tag{5.145}$$

and by analyzing the first few terms in an expansion similar to (5.145) in [17, 18] it was conjectured that

$$\mathrm{Tr}\,Q = \frac{1}{2}\,\mathrm{Tr}\left[\mathbb{1} + q - p - \mathcal{S}_{p,q}\sqrt{\mathbb{1} - 2(p+q) + (p-q)^2}\,\right]\ , \tag{5.146}$$

The right hand side formula is understood this way: first expand the square root as a power series in p and q assuming that p and q commute. Then solve the ordering ambiguities arising from the noncommutativity of p and q by symmetrizing, with the operator $\mathcal{S}_{p,q}$, each monomial in the p and q matrices. A world (monomial) in the letters p and q is symmetrized by considering the sum of all the permutations of its letters, then normalize the sum by dividing by the number of permutations. This normalization of $\mathcal{S}_{p,q}$ is such that if p and q commute then $\mathcal{S}_{p,q}$ acts as the identity. Therefore in the case of just one abelian gauge field (5.112) reduces to the usual Born-Infeld lagrangian. An explicit formula for the coefficients of the expansion of the trace of Q is [19, 69]

$$\operatorname{Tr} Q = \operatorname{Tr}\left[q + \sum_{r,s\geq 1} \begin{pmatrix} r+s-2 \\ r-1 \end{pmatrix} \begin{pmatrix} r+s \\ r \end{pmatrix} \mathcal{S}(\, p^r q^s\,) \right] . \tag{5.147}$$

In Appendix 8, following [19], see also [70] and [71], we prove that the trace of Q is completely symmetrized in the matrix coefficients q and $p-q$. Since this is equivalent to symmetrization in q and p (5.146) follows. Since symmetrization in p and q is equivalent to symmetrization in $\widehat{\alpha}$ and $\widehat{\beta}$, the Born-Infeld lagrangian also reads

$$\mathcal{L} = \operatorname{Tr}[\mathbb{1} - \mathcal{S}_{\alpha,\beta}\sqrt{\mathbb{1} + 2\widehat{\alpha} - \widehat{\beta}^2} + \mathcal{N}_\infty \beta] \; . \tag{5.148}$$

In [69] the convergence of perturbative matrix solutions of (5.137), are studied. A sufficient condition for the convergence of the sequence (5.145) to a solution of (5.142) is that the norms of $p-q$ and q have to satisfy $(1-||p-q||)^2 > 4||q||$. Here $||\;||$ denotes any matrix norm with the Banach algebra property $||MM'|| \leq ||M||\,||M'||$ (e.g. the usual norm). This condition is surely met if the field strengths $F^\Lambda_{\mu\nu}$ are weak.

If equation (5.142) is written as $(\mathbb{1}+q-p)Q = q + Q^2$, then the sequence given by $Q_0 = 0\,,\;\; Q_{k+1} = (\mathbb{1}+q-p)^{-1}q + (\mathbb{1}+q-p)^{-1}Q_k^2$ converges and is a solution of equation (5.142) if $||(\mathbb{1}+q-p)^{-1}||\,||(\mathbb{1}+q-p)^{-1}q|| < 1/4$. Notice that the matrix $\mathbb{1}+q-p$ is always invertible, use $\frac{1}{2}(\mathbb{1}+q-p) + \frac{1}{2}(\mathbb{1}+q-p)^\dagger = \mathbb{1}$, and the same argument as in (5.407). Notice also that if p and q commute then $\sqrt{\mathbb{1} - 2(p+q) + (p-q)^2} = (\mathbb{1}+q-p)\sqrt{\mathbb{1} - 4(\mathbb{1}+q-p)^{-2}q}$ and convergence of the power series expansion of this latter square root holds if $||(\mathbb{1}+q-p)^{-2}q|| < 1/4$.

Real field Strengths

We here construct a Born-Infeld theory with n real field strengths which is duality invariant under the duality group $Sp(2n,\mathbb{R})$.

We first study the case without scalar fields, i.e. $\mathcal{N}_\infty = \prime$ and $-\mathcal{N}_\in = \mathcal{R} = \infty\!\!\infty$. Consider a Lagrangian $\mathcal{L} = \mathcal{L}(\alpha,\beta)$ with n complex gauge fields which describes a theory symmetric under the maximal compact group $U(N)\times U(N)$ of holomorphic duality rotations. Assume that the Lagrangian is a sum of traces (or of products of traces) of monomials in α and β . It follows that this Lagrangian satisfies the self-duality equations (5.83) with α and β complex (recall end of Section 3.7.1). This equation remains true in the special case that α and β assume real values. That is $\mathcal{L} = \mathcal{L}(\alpha,\beta)$ satisfies the self-duality equation (5.83) with $\alpha = \alpha^T = \bar{\alpha}$ and $\beta = \beta^T = \bar{\beta}$. We now recall that equation (5.83) is also the self-duality condition for Lagrangians with real gauge fields provided that α and β are defined as in (5.81) as functions of field strengths F^Λ that are real (cf. the different complex case definition (5.130)). This implies that the theory described by the lagrangian $\mathcal{L}(\alpha,\beta)$ that is now function of n real field strengths is self-dual with duality group $U(n)$, the maximal compact subgroup of $Sp(2n,\mathbb{R})$. The duality group can be extended to the full noncompact $Sp(2n,\mathbb{R})$, by introducing the symmetric matrix of scalar fields $\mathcal{N}$ via the prescription (5.87).

As a straighforward application we obtain the Born-Infeld Lagrangian with n real gauge fields describing an $Sp(2n,\mathbb{R})$ duality invariant theory

$$\mathcal{L} = \operatorname{Tr}\,[\,\mathbb{1} - \mathcal{S}_{\widehat{\alpha},\widehat{\beta}}\sqrt{1 + 2\widehat{\alpha} - \widehat{\beta}^2} + \mathcal{N}_\infty\beta\,]\;, \tag{5.149}$$

where $\widehat{\alpha} = R\alpha R^t$, $\widehat{\beta} = R\beta R^t$, $\mathcal{N}_\in = -\mathcal{R}^\sqcup\mathcal{R}$, and $\alpha^{\Lambda\Sigma} = \frac{1}{4}F^\Lambda F^\Sigma$, $\beta^{\Lambda\Sigma} = \frac{1}{4}\tilde{F}^\Lambda F^\Sigma$ as in (5.81).

Supersymmetric Theory

In this section we briefly discuss supersymmetric versions of some of the Lagrangians introduced. First we discuss the supersymmetric form of the Lagrangian (5.131). Consider the superfields $V^\Lambda = \frac{1}{\sqrt{2}}(V_1^\Lambda + iV_2^\Lambda)$ and $\check{V}^\Lambda = \frac{1}{\sqrt{2}}(V_1^\Lambda - iV_2^\Lambda)$ where V_1^Λ and V_2^Λ are real vector superfields, and define

$$W_\alpha^\Lambda = -\frac{1}{4}\bar{D}^2 D_\alpha V^\Lambda \ , \quad \check{W}_\alpha^\Lambda = -\frac{1}{4}\bar{D}^2 D_\alpha \check{V}^\Lambda \ .$$

Both W^Λ and $\check{W}^\Lambda$ are chiral superfields and can be used to construct a matrix of chiral superfields

$$M^{\Lambda\Sigma} \equiv W^\Lambda \check{W}^\Sigma \ .$$

The supersymmetric version of the Lagrangian (5.131) is then given by

$$\mathcal{L} = \mathrm{Re} \int d^2\Theta \left[\mathrm{Tr}\,(i(\mathcal{N} - \lambda)\chi - \frac{\rangle}{\in}\lambda \bar{\mathcal{D}}^\in(\chi^\dagger \mathcal{N}_\in \chi) + \rangle \lambda \mathcal{M}) \right] \ ,$$

where $\mathcal{N}$, λ and χ denote chiral superfields with the same symmetry properties as their corresponding bosonic fields. While the bosonic fields $\mathcal{N}$ and λ appearing in (5.131) are the lowest component of the superfields denoted by the same letter, the field χ in the action (5.131) is the highest component of the superfield χ. A supersymmetric kinetic term for the scalar field $\mathcal{N}$ can be written using the Kähler potential (5.107) as described in [72].

Just as in the bosonic Born-Infeld theory, one would like to eliminate the auxiliary fields. This is an open problem if $n \neq 1$. For $n = 1$ just as in the bosonic case the theory with auxiliary fields also admits both a real and a complex version, i.e. one can also consider a Lagrangian with a single real superfield. Then by integrating out the auxiliary superfields the supersymmetric version of the Born-Infeld lagrangian (5.144) is obtained

$$\mathcal{L} = \int d^4\Theta \frac{\mathcal{N}_\in^\in \mathcal{W}^\in \bar{\mathcal{W}}^\in}{1 + A + \sqrt{1 + 2A + B^2}} + \mathrm{Re}\left[\int d^2\Theta (\frac{i}{2}\mathcal{N}\mathcal{W}^\in) \right] \ , \tag{5.150}$$

where

$$A = \frac{1}{4}(D^2(\mathcal{N}_\in \mathcal{W}^\in) + \bar{\mathcal{D}}^\in(\mathcal{N}_\in \bar{\mathcal{W}}^\in)) \ , \quad \mathcal{B} = \frac{\infty}{\triangle}(\mathcal{D}^\in(\mathcal{N}_\in \mathcal{W}^\in) - \bar{\mathcal{D}}^\in(\mathcal{N}_\in \bar{\mathcal{W}}^\in)) \ .$$

If we only want a $U(1)$ duality invariance we can set $\mathcal{N} = -\rangle$ and then the lagrangian (5.150) reduces to the supersymmetric Born-Infeld lagrangian described in [46, 47, 48].

In the case of weak fields the first term of (5.150) can be neglected and the Lagrangian is quadratic in the field strengths. Under these conditions the combined requirements of supersymmetry and self duality can be used [73] to constrain the form of the weak coupling limit of the effective Lagrangian from string theory. Self-duality of Born-infeld theories with $N = 2$ supersymmetries is discussed in [24].

5.4 Dualities in $N > 2$ extended Supergravities

In this section we consider $N > 2$ supergravity theories in $D = 4$; in these theories the graviton is also coupled to gauge fields and scalars. We study the corresponding duality groups, that are subgroups of the symplectic group. It is via the geometry of these subgroups of the symplectic group that we can obtain the scalars kinetic terms, the supersymmetry transformation rules and the structure of the central and matter charges of the theory with their differential equations and their duality invariant combinations $\mathscr{V}_{BH}$ and $\mathscr{S}$ (that for extremal black holes are the effective potential and the entropy).

Four dimensional N-extended supergravities contain in the bosonic sector, besides the metric, a number n of vectors and m of (real) scalar fields. The relevant bosonic action is known to have the following general form:

$$\mathcal{S} = \frac{1}{4}\int \sqrt{-g}\, d^4x \left(-\frac{1}{2}\, R + \mathrm{Im}\, \mathcal{N}_{\Lambda\Gamma} F^{\Lambda}_{\mu\nu} F^{\Gamma\,\mu\nu} + \frac{1}{2\sqrt{-g}}\, \mathrm{Re}\, \mathcal{N}_{\Lambda\Gamma} \epsilon^{\mu\nu\rho\sigma}\, F^{\Lambda}_{\mu\nu} F^{\Gamma}_{\rho\sigma} + \right.$$
$$\left. + \frac{1}{2}\, g_{ij}(\phi)\partial_\mu \phi^i \partial^\mu \phi^j \right) , \tag{5.151}$$

where $g_{ij}(\phi)$ $(i, j, \cdots = 1, \cdots, m)$ is the scalar metric on the σ-model described by the scalar manifold M_{scalar} of real dimension m and the vectors kinetic matrix $\mathcal{N}_{\Lambda\Sigma}(\phi)$ is a complex, symmetric, $n \times n$ matrix depending on the scalar fields. The number of vectors and scalars, namely n and m, and the geometric properties of the scalar manifold M_{scalar} depend on the number N of supersymmetries and are summarized in Table 1.

The duality group of these theories is in general not the maximal one $Sp(2n, \mathbb{R})$ because the requirement of supersymmetry constraints the number and the geometry of the scalar fields in the theory. In this section we study the case where the scalar fields manifold is a coset space G/H, and we see that the duality group in this case is G.

In Section 5 we then study the general $N = 2$ case where the target space is a special Kähler manifold M and thus in general we do not have a coset space. There the $Sp(2n, \mathbb{R})$ transformations are needed in order to globally define the supergravity theory. We do not have a duality symmetry of the theory; $Sp(2n, \mathbb{R})$ is rather a gauge symmetry of the theory, in the sense that only $Sp(2n, \mathbb{R})$ invariant expressions are physical ones.

The case of duality rotations in $N = 1$ supergravity is considered in [9], [74], see also [25]. In this case there is no vector potential in the graviton multiplet hence no scalar central charge in the supersymmetry algebra. Duality symmetry is due to the number of matter vector multiplets in the theory, the coupling to eventual chiral multiplets must be via a kinetic matrix $\mathcal{N}$ holomorphic in the chiral fields. We see that the structure of duality rotations is similar to that of $N = 1$ rigid supersymmetry. For duality rotations in $N = 1$ and $N = 2$ rigid supersymmetry using superfields see the review [24].

5.4.1 Extended supergravities with target space G/H

In $N \geq 2$ supergravity theories where the scalars target space is a coset G/H, the scalar sector has a Lagrangian invariant under the global G rotations. Since the scalars appear in supersymmetry multiplets the symmetry G should be a symmetry of the whole theory. This is indeed the case and the symmetry on the vector potentials is duality symmetry.

Let's examine the gauge sector of the theory. We recall from Section 3.1 that we have an $Sp(2n, \mathbb{R})$ duality group if the vector $\binom{F}{G}$ transforms in the fundamental of $Sp(2n, \mathbb{R})$, and the gauge kinetic term $\mathcal{N}$ transforms via fractional transformations, if $\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in Sp(2n, \mathbb{R})$,

$$\mathcal{N} \to \mathcal{N}' = (\mathcal{C} + \mathcal{D}\mathcal{N})\,(\mathcal{A} + \mathcal{B}\mathcal{N})^{-\infty} . \tag{5.152}$$

Thus in order to have G duality symmetry, G needs to act on the vector $\binom{F}{G}$ via symplectic transformations, i.e. via matrices $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ in the fundamental of $Sp(2n, \mathbb{R})$. This requires a homomorphism

$$S : G \to Sp(2n, \mathbb{R}) . \tag{5.153}$$

Different infinitesimal G transformations should correspond to different infinitesimal symplectic rotations so that the induced map $\mathrm{Lie}(G) \to \mathrm{Lie}(Sp(2n, \mathbb{R}))$ is injective, and equivalently the homomorphism S is a local embedding (in general S it is not globally injective, the kernel of S may contain some discrete subgroups of G).

Since $U(n)$ is the maximal compact subgroup of $Sp(2n,\mathbb{R})$ and since H is compact, we have that the image of H under this local embedding is in $U(n)$. It follows that we have a G-equivariant map

$$\mathcal{N} : \mathcal{G}/\mathcal{H} \to \mathcal{S}_{\surd}(\in\backslash,\mathbb{R})/\mathcal{U}(\backslash) \ , \tag{5.154}$$

explicitly, for all $g \in G$,

$$\mathcal{N}(\}\phi) = (\mathcal{C} + \mathcal{D}\mathcal{N}(\phi))\,(\mathcal{A} + \mathcal{B}\mathcal{N}(\phi))^{-\infty} \ , \tag{5.155}$$

where with $g\phi$ we denote the action of G on G/H, while the action of G on $Sp(2n,\mathbb{R})/U(n)$ is given by fractional transformations. Notice that we have identified $Sp(2n,\mathbb{R})/U(n)$ with the space of complex symmetric matrices $\mathcal{N}$ that have imaginary part $\mathrm{Im}\,\mathcal{N} = -\rangle(\mathcal{N} - \overline{\mathcal{N}})$ negative definite (see Appendix 7.2).

The $D = 4$ supergravity theories with $N > 2$ have all target space G/H, they are characterized by the number n of total vectors, the number N of supersymmetries, and the coset space G/H, see Table 1[*].

Table 1: *Scalar Manifolds of $N > 2$ Extended Supergravities*

N	Duality group G	isotropy H	M_{scalar}	n	m
3	$SU(3,n')$	$S(U(3)\times U(n'))$	$\frac{SU(3,n')}{S(U(3)\times U(n'))}$	$3+n'$	$6n'$
4	$SU(1,1)\times SO(6,n')$	$U(1)\times S(O(6)\times O(n'))$	$\frac{SU(1,1)}{U(1)}\times\frac{SO(6,n')}{S(O(6)\times O(n'))}$	$6+n'$	$6n'+2$
5	$SU(5,1)$	$S(U(5)\times U(1))$	$\frac{SU(5,1)}{S(U(5)\times U(1))}$	10	10
6	$SO^\star(12)$	$U(6)$	$\frac{SO^\star(12)}{U(6)}$	16	30
7, 8	$E_{7(7)}$	$SU(8)/\mathbb{Z}_2$	$\frac{E_{7(7)}}{SU(8)/\mathbb{Z}_2}$	28	70

In the table, n stands for the number of vectors and $m = \dim M_{scalar}$ for the number of real scalar fields. In all the cases the duality group G is (locally) embedded in $Sp(2\,n,\mathbb{R})$. The number n of vector potentials of the theory is given by $n = n_g + n'$ where n' is the number of vectors potentials in the matter multiplet while n_g is the number of graviphotons (i.e. of vector potentials that belong to the graviton multiplet). We recall that $n_g = \frac{N(N-1)}{2}$ if $N \neq 6$; and $n_g = \frac{N(N-1)}{2} + 1 = 16$ if $N = 6$; we also have $n' = 0$ if $N > 4$. The scalar manifold of the $N = 4$ case is usually written as $SO_o(6,n')/SO(6)\times SO(n')$ where $SO_o(6,n')$ is the component of $SO(6,n')$ connected to the indentity. The duality group of the $N = 6$ theory is more precisely the double cover of $SO^*(12)$. Spinors fields transform according to H or its double cover.

In general the isotropy group H is the product

$$H = H_{\mathrm{Aut}} \times H_{\mathrm{matter}} \tag{5.156}$$

[*]In Table 1 the group $S(U(p)\times U(q))$ is the group of block diagonal matrices $\left(\begin{smallmatrix} P & 0 \\ 0 & Q \end{smallmatrix}\right)$ with $P \in U(p)$, $Q \in U(q)$ and $\det P \det Q = 1$. There is a local isomorphism between $S(U(p)\times U(q))$ and the direct product group $U(1)\times SU(p)\times SU(q)$, in particular the corresponding Lie algebras coincide. Globally these groups are not the same, for example $S(U(5)\times U(1)) = U(5) = U(1)\times PSU(5) \neq U(1)\times SU(5)$.

where $H_{\rm Aut}$ is the authomorphism group of the supersymmetry algebra, while $H_{\rm matter}$ depends on the matter vector multiplets, that are not present in $N > 4$ supergravities.

In Section 3.5 we have described the geometry of the coset space G/H in terms of coset representatives, local sections L of the bundle $G \to G/H$. Under a left action of G they transform as $gL(\phi) = L(\phi')h$, where the g action on $\phi \in G/H$ gives the point $\phi' \in G/H$.

We now recall that duality symmetry is implemented by the symplectic embeddings (5.153) and (5.154) and conclude that the embeddings of the coset representatives L in $Sp(2n, \mathbb{R})$ will play a central role. Recalling (5.102) these embeddings are determined by defining

$$L \to f(L) \quad \text{and} \quad L \to h(L) \ . \tag{5.157}$$

In the following we see that the matrices $f(L)$ and $h(L)$ determine the scalar kinetic term $\boldsymbol{\mathcal{N}}$, the supersymmetry transformation rules and the structure of the central and matter charges of the theory. We also derive the differential equations that these charges satisfy and consider their positive definite and duality invariant quadratic expression $\mathscr{V}_{BH}$. These relations are similar to the Special Geometry ones of $N = 2$ supergravity.

>From the equation of motion

$$dF^\Lambda = 4\pi j^\Lambda_m \tag{5.158}$$
$$dG^\lambda = 4\pi j_{e\Lambda} \tag{5.159}$$

we associate with a field strength 2-form F a magnetic charge p^Λ and an electric charge q_Λ given respectively by:

$$p^\Lambda = \frac{1}{4\pi}\int_{S^2} F^\Lambda \quad , \qquad q_\Lambda = \frac{1}{4\pi}\int_{S^2} G_\Lambda \tag{5.160}$$

where S^2 is a spatial two-sphere containing these electric and magnetic charges. These are not the only charges of the theory, in particular we are interested in the central charges of the supersymmetry algebra and other charges related to the vector multiplets. These latter charges result to be the electric and magnentic charges p^Λ and q_Λ dressed with the scalar fields of the theory. In particular these dressed charges are invariant under the duality group G and transform under the isotropy subgroup $H = H_{Aut} \times H_{matter}$.

While the index Λ is used for the fundamental representation of $Sp(2n; R)$ the index M is used for that of $U(n)$. According to the local embedding

$$H = H_{Aut} \times H_{matter} \to U(n) \tag{5.161}$$

the index M is further divided as $M = (AB, \bar{I})$ where $\bar{I}$ refers to H_{matter} and $AB = -BA$ ($A = 1, \dots, N$) labels the two-times antisymmetric representation of the R-symmetry group H_{Aut}. We can understand the appearence of this representation of H_{Aut} because this is a typical representation acting on the central charges. The index $\bar{I}$ rather than I is used because the image of H_{matter} in $U(n)$ will be the complex conjugate of the fundamental of H_{matter}, this agrees with the property that under Kähler transformations of the $U(1)$ bundle $Sp(2n, \mathbb{R})/SU(n) \to Sp(2n, \mathbb{R})/U(n)$ the coset representatives of the scalar fields in the gravitational and matter multiplets transform with opposite Kähler weights. This is also what happens in the generic $N = 2$ case (cf. (5.357)).

The dressed graviphotons field strength 2-forms T_{AB} may be identified from the supersymmetry transformation law of the gravitino field in the interacting theory, namely:

$$\delta\psi_A = \nabla\varepsilon_A + \alpha T_{AB\,\mu\nu}\gamma^a\gamma^{\mu\nu}\varepsilon^B V_a + \dots \tag{5.162}$$

Here ∇ is the covariant derivative in terms of the space-time spin connection and the composite connection of the automorphism group H_{Aut}, α is a coefficient fixed by supersymmetry, V^a is the space-time vielbein. Here and in the following the dots denote trilinear fermion terms which are characteristic of any supersymmetric theory but do not play any role in the following discussion. The 2-form field strength T_{AB} is constructed by dressing the bare field strengths F^Λ with the image $f(L(\phi))$, $h(L(\phi))$ in $Sp(2n;R)$ of the coset representative $L(\phi)$ of G/H. Note that the same field strengths T_{AB} which appear in the gravitino transformation law are also present in the dilatino transformation law in the following way:

$$\delta\chi_{ABC} = \mathcal{P}_{ABCD\,\ell}\partial_\mu\phi^\ell\gamma^\mu\varepsilon^D + \beta T_{[AB\,\mu\nu}\gamma^{\mu\nu}\varepsilon_{C]} \tag{5.163}$$

Analogously, when vector multiplets are present, the matter vector field strengths T_I appearing in the transformation laws of the gaugino fields, are linear combinations of the field strengths dressed with a different combination of the scalars:

$$\delta\lambda_{IA} = i\mathcal{P}_{IAB\,r}\partial_\mu\phi^r\gamma^\mu\varepsilon^B + \gamma T_{I\,\mu\nu}\gamma^{\mu\nu}\varepsilon_A + \dots \tag{5.164}$$

Here $\mathcal{P}_{ABCD} = \mathcal{P}_{ABCD\,\ell}\,d\phi^\ell$ and $\mathcal{P}^I_{AB} = \mathcal{P}^I_{AB\,r}\,d\phi^r$ are the vielbein of the scalar manifolds spanned by the scalar fields $\phi^i = (\phi^\ell, \phi^r)$ of the gravitational and vector multiplets respectively (more precise definitions are given below), and β and γ are constants fixed by supersymmetry.

According to the transformation of the coset representative $gL(\phi) = L(\phi')h$, under the action of $g \in G$ on G/H we have

$$S(\phi)\bar{A} \to S(\phi')\bar{A} = S(g)S(\phi)S(h^{-1})\bar{A} = S(g)S(\phi)\bar{A}U^{-1} \tag{5.165}$$

where $\mathcal{A} = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix}\mathbb{1} & \mathbb{1}\\ -i\mathbb{1} & i\mathbb{1}\end{smallmatrix}\right)$ is unitary and symplectic (cf. (5.394)), $S(g) = \left(\begin{smallmatrix}A & B\\ C & D\end{smallmatrix}\right)$ and $S(h)$ are the embeddings of g and h in the fundamental of $Sp(2n,\mathbb{R})$, while $U = \bar{A}^{-1}S(h)\bar{A}$ is the embedding of h in the complex basis of $Sp(2n,\mathbb{R})$. Explicitly $U = \left(\begin{smallmatrix}u & 0\\ 0 & \bar{u}\end{smallmatrix}\right)$, where u is in the fundamental of $U(n)$ (cf. (5.402) and (5.397)). Therefore the symplectic matrix

$$V = S\bar{A} = \begin{pmatrix} f & \bar{f} \\ h & \bar{h} \end{pmatrix} \tag{5.166}$$

transforms according to

$$V(\phi) \to V(\phi') = S(g)V(\phi)\begin{pmatrix} u^{-1} & 0 \\ 0 & \bar{u}^{-1} \end{pmatrix} . \tag{5.167}$$

The dressed field strengths transform only under a unitary representation of H and, in accordance with (5.167), are given by [11]

$$\begin{pmatrix} T \\ -\bar{T} \end{pmatrix} = -i\,\overline{V(\phi)}^{-1}\begin{pmatrix} F \\ G \end{pmatrix} ; \tag{5.168}$$

$$T \to \bar{u}T\ .\bar{u}T\ . \tag{5.169}$$

Explicitly, since

$$-i\,\bar{V}^{-1} = \begin{pmatrix} h^t & -f^t \\ -h^\dagger & f^\dagger \end{pmatrix} \tag{5.170}$$

we have

$$\begin{aligned} T_{AB} &= h_{\Lambda\,AB}F^\Lambda - f^\Lambda{}_{AB}G_\Lambda \\ \bar{T}_{\bar{I}} &= \bar{h}_{\Lambda\bar{I}}F^\Lambda - \bar{f}^\Lambda_{\bar{I}}G_\Lambda \end{aligned} \tag{5.171}$$

where we used the notation $T = (T^{\bar{M}}) = (T_M) = (T_{AB}, \bar{T}_{\bar{I}})$,

$$
\begin{aligned}
f &= (f^\Lambda{}_M) = (f^\Lambda{}_{AB}, \bar{f}^\Lambda{}_{\bar{I}}) \ , \\
h &= (h_{\Lambda M}) = (h_{\Lambda AB}, \bar{h}_{\Lambda \bar{I}}) \ ,
\end{aligned}
\tag{5.172}
$$

that enphasizes that (for every value of Λ) the sections $\begin{pmatrix} \bar{f}^\Lambda{}_{\bar{I}} \\ \bar{h}_{\Lambda\bar{I}} \end{pmatrix}$ have Kweight opposite to the $\begin{pmatrix} f^\Lambda{}_{AB} \\ h_{\Lambda AB} \end{pmatrix}$ ones. This may be seen from the supersymmetry transformation rules of the supergravity fields, in virtue of the fact that gravitinos and fotinos with the same chirality have opposite Kähler weight. Notice that this notation (as in [41]) differs from the one in [11], where $(f^\Lambda{}_M) = (f^\Lambda{}_{AB}, f^\Lambda{}_I)$, $(h_{\Lambda M}) = (h_{\Lambda AB}, h_{\Lambda I})$.

Consequently the central charges are

$$
Z_{AB} = -\frac{1}{4\pi} \int_{S^2_\infty} T_{AB} = f^\Lambda{}_{AB} q_\Lambda - h_{\Lambda\,AB} p^\Lambda \tag{5.173}
$$

$$
\bar{Z}_{\bar{I}} = -\frac{1}{4\pi} \int_{S^2_\infty} \bar{T}_{\bar{I}} = \bar{f}^\Lambda_{\bar{I}} q_\Lambda - \bar{h}_{\Lambda \bar{I}} p^\Lambda \tag{5.174}
$$

where the integral is considered at spatial infinity and, for spherically symmetric configurations, f and h in (5.173), (5.174) are $f(\phi_\infty)$ and $h(\phi_\infty)$ with ϕ_∞ the constant value assumed by the scalar fields at spatial infinity.

The integral of the graviphotons $T_{AB\,\mu\nu}$ gives the value of the central charges Z_{AB} of the supersymmetry algebra, while by integrating the matter field strengths $T_{I\,\mu\nu}$ one obtains the so called matter charges Z_I. The charges of these dressed field strength that appear in the supersymmetry transformations of the fermions have a profound meaning and play a key role in the physics of extremal black holes. In particular, recalling (5.167) the quadratic combination (black hole potential)

$$
\mathscr{V}_{BH} := \frac{1}{2} \bar{Z}^{AB} Z_{AB} + \bar{Z}^I Z_I \tag{5.175}
$$

(the factor 1/2 is due to our summation convention that treats the AB indices as independent) is invariant under the symmetry group G. In terms of the charge vector

$$
Q = \begin{pmatrix} p^\Lambda \\ q_\Lambda \end{pmatrix} \ , \tag{5.176}
$$

we have the formula for the potential (also called charges sum rule)

$$
\mathscr{V}_{BH} = \frac{1}{2} \bar{Z}^{AB} Z_{AB} + \bar{Z}^I Z_I = -\frac{1}{2} Q^t \mathcal{M}(\mathcal{N}) Q \tag{5.177}
$$

where

$$
\mathcal{M}(\mathcal{N}) = -(i\bar{V}^{-1})^\dagger i\bar{V}^{-1} = -(S^{-1})^t S^{-1} \tag{5.178}
$$

is a negative definite matrix, here depending on ϕ_∞. In Appendix 7.2 we show that the set of matrices of the kind SS^t with $S \in Sp(2n, \mathbb{R})$ are the coset space $Sp(2n, \mathbb{R})/U(n)$, hence the matrices $\mathcal{M}(\mathcal{N})$ parametrize $Sp(2n, \mathbb{R})/U(n)$. Also the matrices $\mathcal{N}$ parametrize $Sp(2n, \mathbb{R})/U(n)$. The relation between $\mathcal{M}(\mathcal{N})$ and $\mathcal{N}$ is

$$
M^*(\mathcal{N}) = \begin{pmatrix} \mathbb{1} & -\mathrm{Re}\,\mathcal{N} \\ 0 & \mathbb{1} \end{pmatrix} \begin{pmatrix} \mathrm{Im}\,\mathcal{N} & 0 \\ 0 & \mathrm{Im}\,\mathcal{N}^{-\infty} \end{pmatrix} \begin{pmatrix} \mathbb{1} & 0 \\ -\mathrm{Re}\,\mathcal{N} & \mathbb{1} \end{pmatrix} \ . \tag{5.179}
$$

This and further properties of the $\mathcal{M}(\mathcal{N})$ matrix are derived in Appendix 7.2.

For each of the supergravities with target space G/H there is another G invariant expression $\mathscr{S}$ quadratic in the charges [63]; the invariant $\mathscr{S}$ is independent from the scalar fields of the theory and thus depends only on the electric and magnetic charges p^Λ and q_Λ. In extremal black hole configurations $\pi\mathscr{S}$ is the entropy of the black hole. In the $N=3$ supergravity theory $\mathscr{S}$ is the absolute value of a quadratic combination of the charges, while for $N\geq 4$ it is the square root of the absolute value of a quartic combination of the charges. The positive or negative value of this quadratic combination is related to the different BPS properties of the black hole. It turns out that $\mathscr{S}$ coincides with the potential $\mathscr{V}_{BH}$ computed at its critical point (attractor point) [43, 45, 63]. In the next section we give the explicit expressions of the invariants $\mathscr{S}$. They are obtained by considering among the H invariant combination of the charges those that are also G invariant, i.e. those that do not depend on the scalar fields. This is equivalent to require invariance of $\mathscr{S}$ under the coset space covariant derivative ∇ defined in Section 3.5, see also (5.184).

We now derive some differential relations among the central and matter charges. We recall the symmetric coset space geometry G/H studied in Section 3.5, and in particular relations (5.98), (5.99) that express the Maurer-Cartan equation $d\Gamma+\Gamma\wedge\Gamma=0$ in terms of the vielbein P and of the Riemannian connection ω. Using the (local) embedding of G in $Sp(2n,\mathbb{R})$ we consider the pull back on G/H of the $Sp(2n,\mathbb{R})$ Lie algebra left invariant one form $V^{-1}dV$ given in (5.104), we have

$$V^{-1}dV=\begin{pmatrix} i(f^\dagger dh-h^\dagger df) & i(f^\dagger d\bar{h}-h^\dagger d\bar{f})\\ -i(f^t dh-h^t df) & -i(f^t d\bar{h}-h^t d\bar{f})\end{pmatrix}=\begin{pmatrix}\omega & \bar{\mathcal{P}}\\ \mathcal{P} & \bar{\omega}\end{pmatrix}\ , \tag{5.180}$$

where with slight abuse of notation we use the same letters V, $\mathcal{P}$ and ω for the pulled back forms (we also recall that $\mathcal{P}$ denotes P in the complex basis). Relation (5.180) equivalently reads

$$dV=V\begin{pmatrix}\omega & \bar{\mathcal{P}}\\ \mathcal{P} & \bar{\omega}\end{pmatrix}\ , \tag{5.181}$$

that is equivalent to the $n\times n$ matrix equations:

$$\nabla f=\bar{f}\mathcal{P}\ , \tag{5.182}$$

$$\nabla h=\bar{h}\mathcal{P}\ , \tag{5.183}$$

where

$$\nabla f=df-f\omega\quad ,\qquad \nabla h=dh-h\omega\ . \tag{5.184}$$

Recalling that $\mathcal{P}$ is symmetric (cf. (5.419)) we equivalently have $\nabla f=\mathcal{P}\bar{f}$, $\nabla h=\mathcal{P}\bar{h}$. In these equations we can now see ω and $\mathcal{P}$ as our data (vielbein and Riemannian connection) on a manifold M, while f and h are the unknowns. By construction these equations are automatically satisfied if $M=G/H$ and G is a Lie subgroup of $Sp(2n,\mathbb{R})$. More in general equations (5.182),(5.183) hold (with f and h invertible) iff the integrability condition, i.e. the Cartan-Maurer equation, $d\left(\begin{smallmatrix}\omega & \mathcal{P}\\ \bar{\mathcal{P}} & \omega\end{smallmatrix}\right)+\left(\begin{smallmatrix}\omega & \mathcal{P}\\ \bar{\mathcal{P}} & \omega\end{smallmatrix}\right)\wedge\left(\begin{smallmatrix}\omega & \mathcal{P}\\ \bar{\mathcal{P}} & \omega\end{smallmatrix}\right)=0$ holds. With abuse of terminology we sometimes call (5.182), (5.183) the Maurer-Cartan equations.

The differential relations among the charges Z_{AB} and $\bar{Z}_{\bar{I}}$ follow after rewriting (5.182), (5.183) with AB and $\bar{I}$ indices. The embedded connection ω and vielbein $\mathcal{P}$ are decomposed as follows:

$$\omega=(\omega^N{}_M)=\begin{pmatrix}\omega^{AB}{}_{CD} & 0\\ 0 & \omega^{\bar{I}}{}_{\bar{J}}\end{pmatrix}\ , \tag{5.185}$$

$$\mathcal{P}=(\mathcal{P}^{\bar{N}}{}_M)=(\mathcal{P}_{NM})=\begin{pmatrix}\mathcal{P}^{\bar{A}\bar{B}}{}_{CD} & \mathcal{P}^{\bar{A}\bar{B}}{}_{\bar{J}}\\ \mathcal{P}^{\bar{I}}{}_{CD} & \mathcal{P}^{\bar{I}}{}_{\bar{J}}\end{pmatrix}=\begin{pmatrix}\mathcal{P}_{ABCD} & \mathcal{P}_{AB\bar{J}}\\ \mathcal{P}_{ICD} & \mathcal{P}_{I\bar{J}}\end{pmatrix}\ , \tag{5.186}$$

the subblocks being related to the vielbein of G/H, written in terms of the indices of $H_{Aut} \times H_{matter}$. We used the following indices conventions:

$$f = \left(f^{\Lambda}{}_{M}\right) \quad , \quad f^{-1} = \left(f^{M}{}_{\Lambda}\right) = \left(f_{\bar{M}\Lambda}\right) \qquad \text{etc.} \tag{5.187}$$

where in the last passage, since we are in $U(n)$, we have lowered the index M with the $U(n)$ hermitian form $\eta = \left(\eta_{M\bar{N}}\right)_{M,N=1,...n} = diag(1,1,....1)$. Similar conventions hold for the AB and I indices, for example $\overline{f^{\Lambda}_{I}} = \bar{f}^{\Lambda}_{\bar{I}} = \bar{f}^{\Lambda I}$.

Using further the index decomposition $M = (AB, \bar{I})$, relations (5.182), (5.183) read (the factor 1/2 is due to our summation convention that treats the AB indices as independent):

$$\nabla f^{\Lambda}_{AB} = \frac{1}{2}\bar{f}^{\Lambda CD}\mathcal{P}_{CDAB} + f^{\Lambda}{}_{I}\mathcal{P}^{I}{}_{AB}\ , \tag{5.188}$$

$$\nabla h^{\Lambda}_{AB} = \frac{1}{2}\bar{h}^{\Lambda CD}\mathcal{P}_{CDAB} + h^{\Lambda}{}_{I}\mathcal{P}^{I}{}_{AB}\ , \tag{5.189}$$

$$\nabla f^{\Lambda}{}_{\bar{I}} = \frac{1}{2}\bar{f}^{\Lambda CD}\mathcal{P}_{CD\bar{I}} + f^{\Lambda\bar{J}}\mathcal{P}_{\bar{J}\bar{I}}\ , \tag{5.190}$$

$$\nabla h^{\Lambda}{}_{\bar{I}} = \frac{1}{2}\bar{h}^{\Lambda CD}\mathcal{P}_{CD\bar{I}} + h^{\Lambda\bar{J}}\mathcal{P}_{\bar{J}\bar{I}}\ . \tag{5.191}$$

As we will see, depending on the coset manifold, some of the sub-blocks of (5.186) can be actually zero. For $N > 4$ (no matter indices) we have that $\mathcal{P}$ coincides with the vielbein $\mathcal{P}_{ABCD}$ of the relevant G/H. Using the definition of the charges (21) we then get the differential relations among charges: $\nabla Z_M = \bar{Z}_{\bar{N}}\mathcal{P}^{\bar{N}}{}_{M}$, where $\nabla Z_M = \frac{\partial Z_M}{\partial\phi^i_\infty}d\phi^i_\infty - Z_N\omega^N{}_M$, with ϕ^i_∞ the value of the i-th coordinate of $\phi_\infty \in G/H$ and $\phi_\infty = \phi(r=\infty)$. Explicitly, using the AB and I indices,

$$\nabla Z_{AB} = Z_I\mathcal{P}^{I}{}_{AB} + \frac{1}{2}\bar{Z}^{CD}\mathcal{P}_{CDAB}\ , \tag{5.192}$$

$$\nabla\bar{Z}_{\bar{I}} = \frac{1}{2}\bar{Z}^{AB}\mathcal{P}_{AB\bar{I}} + Z^{\bar{J}}\mathcal{P}_{\bar{J}\bar{I}}\ . \tag{5.193}$$

The geometry underlying the differential equation (5.181) is that of a flat symplectic vector bundle of rank $2n$, a structure that appears also in the special Kmanifolds of scalars of $N = 2$ supergravities. Indeed if we are able to find $2n$ linearly independent row vectors $V^{\xi} = (V^{\xi}{}_{\zeta})_{\zeta=1,...2n}$ then the matrix V in (5.181) is invertible and therefore the connection $\left(\begin{smallmatrix}\omega & \bar{\mathcal{P}}\\ \mathcal{P} & \bar{\omega}\end{smallmatrix}\right)$ is flat. If these vectors are mutually symplectic then we have a symplectic frame, the transition functions are constant symplectic matrices, the connection is symplectic.

In the present case we naturally have a flat symplectic bundle,

$$G \times_H \mathbb{R}^{2n} \to G/H\,;$$

this bundle is the space of all equivalence classes $[g, v] = \{(gh, S(h)^{-1}v)\,,\ g \in G, v \in \mathbb{R}^{2n}, h \in H\}$. The symplectic structure on $\mathbb{R}^{2n}$ immediately extends to a well defined symplectic structure on the fibers of the bundle. Using the local sections of G/H and the usual basis $\{e_\xi\} = \{e_M, e^M\}$ of $\mathbb{R}^{2n}$ (e_1 is the column vector with with 1 as first and only nonvanishing entry, etc.) we obtain immediately the local sections $s_\xi = [L(\phi), e_\xi]$ of $G \times_H \mathbb{R}^{2n} \to G/H$. Since the action of H on $\mathbb{R}^{2n}$ extends to the action of G on $\mathbb{R}^{2n}$, we can consider the new sections $\mathsf{e}_\xi = s_\zeta S^{-1}(L(\phi))^{\zeta}{}_{\xi} = [L(\phi), S^{-1}(L(\phi))e_\xi]$, that are determined by the column vectors $S^{-1}(L(\phi))\ {}_{\xi} = (S^{-1}(L(\phi))^{\zeta}{}_{\xi})_{\zeta=1,...2n}$. These sections are globally defined and linearly independent. Therefore this bundle is not only flat, it is trivial. If we use the complex local frame $\mathcal{V}_\xi = \{s_\zeta\bar{A}^{\zeta}{}_{\xi}\}$ rather than the $\{s_\xi\}$ one (we recall that $\mathcal{A} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1\!\!1 & 1\!\!1 \\ -i1\!\!1 & i1\!\!1 \end{pmatrix}$, cf. (5.394)), then the global sections e_ξ are determined by the column vectors $V^{-1}(L(\phi))\ _\xi = (V^{-1}(L(\phi))^\zeta{}_\xi)_{\zeta=1,\dots 2n}$,

$$\mathsf{e}_\xi = \mathcal{V}_\eta V^{-1\eta}{}_\xi \ . \tag{5.194}$$

The sections $\mathcal{V}_\xi$ too form a symplectic frame (a symplectonormal basis, indeed $V^\rho_\xi \Omega_{\rho\sigma} V^\sigma_\zeta = \Omega_{\xi\zeta}$, where $\Omega = \begin{pmatrix} 0 & -1\!\!1 \\ 1\!\!1 & 0 \end{pmatrix}$)), and the last n sections are the complex conjugate of the first n ones, $\{\mathcal{V}_\xi\} = \{\mathcal{V}_M, \bar{\mathcal{V}}_{\bar M}\}$. Of course the column vectors $V_{\ \eta} = (V^\xi_{\ \eta})_{\xi=1,\dots 2n}$, are the coefficients of the sections $\mathcal{V}_\eta$ with respect to the flat basis $\{\mathsf{e}_\xi\}$.

Also the rows of the V matrix define global flat sections. Let's consider the dual bundle of the vector bundle $G \times_H \mathbb{R}^{2n} \to G/H$, i.e. the bundle with fiber the dual vector space. If $\{s_\zeta\}$ is a frame of local sections of $G \times_H \mathbb{R}^{2n} \to G/H$, then $\{s^\zeta\}$, with $\langle s^\zeta, s_\xi\rangle = \delta^\zeta_\xi$, is the dual frame of local sections of the dual bundle. Concerning the transition functions, if $s'_\zeta = s_\eta S^\eta{}_\zeta$ then $s'^\xi = S^{-1\xi}{}_\lambda s^\lambda$. This dual bundle is also a trivial bundle and a trivialization is given by the global symplectic sections $\mathsf{e}^\xi = V^\xi_{\ \eta} \mathcal{V}^\eta$, whose coefficients are the row vectors $V^\xi = (V^\xi_{\ \zeta})_{\zeta=1,\dots 2n}$ i.e., the rows of the symplectic matrix V defined in (5.166),

$$\begin{aligned} \left(V^\Lambda{}_\zeta\right)_{\zeta=1,\dots 2n} &= \left(f^\Lambda{}_M, \bar f^\Lambda{}_{\bar M}\right)_{M=1,\dots n} \ , \\ \left(V_{\Lambda\zeta}\right)_{\zeta=1,\dots 2n} &= \left(h_{\Lambda M}, \bar h_{\Lambda \bar M}\right)_{M=1,\dots n} \ . \end{aligned} \tag{5.195}$$

5.4.2 Specific cases

We now describe in more detail the supergravities of Table 1. The aim is to write down the group theoretical structure of each theory, their symplectic (local) embedding $S : G \to Sp(2n, \mathbb{R})$ and $\mathcal{N} : \mathcal{G}/\mathcal{H} \to \mathcal{S}_{\surd}(\in\backslash, \mathbb{R})/\mathcal{U}(\backslash)$, the vector kinetic matrix $\mathcal{N}$, the supersymmetric transformation laws, the structure of the central and matter charges, their differential relations originating from the Maurer-Cartan equations (5.98),(5.99), and the invariants $\mathscr{V}_{BH}$ and $\mathscr{S}$. As far as the boson transformation rules are concerned we prefer to write down the supercovariant definition of the field strengths (denoted by a superscript hat), from which the supersymmetry transformation laws are retrieved. As it has been mentioned in previous section it is here that the symplectic sections $(f^\Lambda{}_{AB}, \bar f^\Lambda{}_{\bar I}, \bar f^\Lambda{}_{AB}, f^\Lambda{}_I)$ appear as coefficients of the bilinear fermions in the supercovariant field strengths while the analogous symplectic section $(h_{\Lambda AB}, \bar h_{\Lambda \bar I}, \bar h_{\Lambda AB}, h_{\Lambda I})$ would appear in the dual magnetic theory. We include in the supercovariant field strengths also the supercovariant vielbein of the G/H manifolds. Again this is equivalent to giving the susy transformation laws of the scalar fields. The dressed field strengths from which the central and matter charges are constructed appear instead in the susy transformation laws of the fermions for which we give the expression up to trilinear fermion terms. We stress that the numerical coefficients in the aforementioned susy transformations and supercovariant field strengths are fixed by supersymmetry (or, equivalently, by Bianchi identities in superspace), but we have not worked out the relevant computations being interested in the general structure rather that in the precise numerical expressions. These numerical factors could also be retrieved by comparing our formulae with those written in the standard literature on supergravity and performing the necessary redefinitions. The same kind of considerations apply to the central and matter charges whose precise normalization has not been fixed.

Throughout this section we denote by $A, B, \dots$ indices of $SU(N)$, $SU(N) \times U(1)$, being H_{aut} the automorphism group of the N–extended supersymmetry algebra. Lower and upper $SU(N)$ indices on the fermion fields are related to their left or right chirality respectively. If some fermion is a $SU(N)$ singlet, chirality is denoted by the usual (L) or (R) suffixes.

Furthermore for any boson field v carrying $SU(N)$ indices we have that lower and upper indices are related by complex conjugation, namely: $\overline{(v_{AB\cdots})} = \bar{v}^{AB\cdots}$.

The $N=4$ theory

The field content is given by the

– Gravitational multiplet (vierbein for the graviton, gravitino, graviphoton, dilatino, dilaton):

$$(V^a_\mu, \psi_{A\mu}, A^{AB}_\mu, \chi_{ABC}, \mathcal{N}) \qquad (A,B=1,\cdots,4) \tag{5.196}$$

frequently the upper half plane parametrization $S=\bar{\mathcal{N}}$ is used for the axion-dilaton field.

– Vector multiplets:

$$(A_\mu, \lambda^A, 6\,\phi)^I \qquad (I=1,\cdots,n) \tag{5.197}$$

The coset space is the product

$$G/H = \frac{SU(1,1)}{U(1)} \times \frac{SO(6,n)}{S(O(6)\times O(n))} \tag{5.198}$$

We have to embed

$$Sp(2,\mathbb{R}) \times SO(6,n) \to Sp(2(6+n),\mathbb{R})\ . \tag{5.199}$$

We first consider the embedding of $SO(6,n)$,

$$\begin{aligned} S : SO(6,n) &\to Sp(2(6+n),\mathbb{R}) \\ L \mapsto S(L) &= \begin{pmatrix} L^{t-1} & 0 \\ 0 & L \end{pmatrix} \end{aligned} \tag{5.200}$$

we see that under this embedding $SO(6,n)$ is a symmetry of the action (not only of the equation of motions) that rotates electric fields into electric fields and magentic fields into magnetic fields. The natural embedding of $SU(1,1) \simeq SL(2,\mathbb{R}) \simeq Sp(2,\mathbb{R})$ into $Sp(2(6+n),\mathbb{R})$ is the S-duality that rotates each electric field in its corresponding magnetic field, we also want the image of $Sp(2,\mathbb{R})$ in $Sp(2(6+n),\mathbb{R})$ to commute with that of $SO(6,n)$ (since we are looking for a symplectic embedding of all $Sp(2,\mathbb{R})\times SO(6,n)$) and therefore we have

$$\begin{aligned} S : Sp(2,\mathbb{R}) &\to Sp(2(6+n),\mathbb{R}) \\ \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right) \mapsto S\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right) &= \begin{pmatrix} A\mathbb{1} & B\eta \\ C\eta & D\mathbb{1} \end{pmatrix} \end{aligned} \tag{5.201}$$

where $\eta = diag(1,1,...,-1,-1,...)$ is the $SO(6,n)$ metric.

Concerning the coset representatives, on one hand we denote by $L(t)$ the representative in $SO(6,n)$ of the point $t \in SO(6,n)/S(O(6)\times O(n))$. On the other hand we have that $SU(1,1)/U(1) \simeq Sp(2,\mathbb{R})/U(1)$ is the lower half plane (see appendix) and is spanned by the complex number $\mathcal{N}$ with $\mathrm{Im}\,\mathcal{N} < 0$, (frequently the upper half plane parametrization $S=\bar{\mathcal{N}}$ is used). A coset representative of $SU(1,1)/U(1)$ is

$$U(\mathcal{N}) = \frac{1}{n(\mathcal{N})}\begin{pmatrix} 1 & \frac{i-\bar{\mathcal{N}}}{i+\bar{\mathcal{N}}} \\ \frac{i+\mathcal{N}}{i-\mathcal{N}} & 1 \end{pmatrix} \quad , \quad n(\mathcal{N}) = \sqrt{\frac{-4\mathrm{Im}\,\mathcal{N}}{1+|\mathcal{N}|^2 - 2\mathrm{Im}\,\mathcal{N}}} \tag{5.202}$$

(In order to show that the $SU(1,1)$ matrix $U(\mathcal{N})$ projects to $\mathcal{N}$ use (5.402) and (5.408), that reads $\mathcal{N} = hf^{-1}$ with h and f complex numbers). The coset representative $U(\mathcal{N})$ is defined for any $\mathcal{N}$ in the lower complex plane and therefore $U(\mathcal{N})$ is a global section of the bundle

$SU(1,1) \to SU(1,1)/U(1)$. (The projection $SU(1,1) \to SU(1,1)/U(1)$ can be also obtained by extracting ${\scriptscriptstyle N}$ from $M^*({\scriptscriptstyle N}) = \left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)\mathcal{A}UU^\dagger\mathcal{A}^{-1}\left(\begin{smallmatrix}0 & 1\\ -1 & 0\end{smallmatrix}\right)$, cf. (5.415)).

With the given coset parametrizations the symplectic embedded section $\begin{pmatrix} f^\Lambda{}_\Sigma \\ h_{\Lambda\Sigma}\end{pmatrix}$ is

$$
\begin{aligned}
f^\Lambda_\Sigma &= (f^\Lambda{}_{AB}, \bar f^\Lambda{}_{\bar I}) = \frac{1}{n({\scriptscriptstyle N})}\Big(\frac{2}{1+i{\scriptscriptstyle N}} L^{t-1\,\Lambda}{}_{AB}, \frac{2}{1-i\bar{\scriptscriptstyle N}} L^{t-1\,\Lambda}{}_{\bar I}\Big) \\
h_{\Lambda\Sigma} &= (h_{\Lambda AB}, \bar h_{\Lambda\bar I}) = \frac{1}{n({\scriptscriptstyle N})}\Big(\frac{2{\scriptscriptstyle N}}{i{\scriptscriptstyle N}+1} L^\Lambda{}_{AB}, \frac{2\bar{\scriptscriptstyle N}}{i\bar{\scriptscriptstyle N}-1} L^\Lambda{}_{\bar I}\Big)
\end{aligned}
\tag{5.203}
$$

We now have all ingredients to compute the matrix $\boldsymbol{\mathcal{N}}$ in terms of ${\scriptscriptstyle N}$ and L. The coset representative in $Sp(2(6+m),\mathbb{R})$ of $({\scriptscriptstyle N}, L)$ is $S(\mathcal{A}U({\scriptscriptstyle N})\mathcal{A}^{-1})S(L)$, and recalling that $\boldsymbol{\mathcal{N}} = \langle\{^{-\infty}$ and (5.102), we obtain after elementary algebra the kinetic matrix

$$
\boldsymbol{\mathcal{N}} = \mathrm{Re}\,\boldsymbol{\mathcal{N}} + \rangle\mathrm{Im}\,\boldsymbol{\mathcal{N}} = \mathrm{Re}\,{\scriptscriptstyle N}\,\eta + \rangle\mathrm{Im}\,{\scriptscriptstyle N}\,\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{L}}^\sqcup\ . \tag{5.204}
$$

Table 2: Group assignments of the fields in $D=4$, $N=4$

	V^a_μ	$\psi_{A\vert\mu}$	A^Λ_μ	χ_{ABC}	λ_{IA}	$U({\scriptscriptstyle N})L^\Lambda_{AB}$	$U({\scriptscriptstyle N})L^\Lambda_I$	R_H
$SU(1,1)$	1	1	-	1	1	2×1	2×1	-
$SO(6,n')$	1	1	$6+n'$	1	1	$1\times(6+n')$	$1\times(6+n')$	-
$SO(6)$	1	4	1	$\bar 4$	$\bar 4$	1×6	1	6
$SO(n')$	1	1	1	1	n'	1	n'	n'
$U(1)$	0	$\frac{1}{2}$	0	$\frac{3}{2}$	$-\frac{1}{2}$	1	1	0

In this and in the following tables, R_H is the representation under which the scalar fields of the linearized theory, or the vielbein $\mathcal{P}$ of G/H of the full theory transform (recall text after (5.95) and that $\mathcal{P}$ is P in the complex basis). Only the left–handed fermions are quoted, right handed fermions transform in the complex conjugate representation of H. Care must be taken in the transformation properties under the H subgroups; indeed according to (5.167) the inverse right rep. of the one listed should really appear, i.e. since we are dealing with unitary rep., the complex conjugate

The supercovariant field strengths and the vielbein of the coset manifold are:

$$
\begin{aligned}
\hat F^\Lambda &= dA^\Lambda + \big[f^\Lambda_{AB}(c_1\bar\psi^A\psi^B + c_2\bar\psi_C\gamma_a\chi^{ABC}V^a) \\
&\quad + f^\Lambda_I(c_3\bar\psi^A\gamma_a\lambda^I_A V^a + c_4\bar\chi^{ABC}\gamma_{ab}\lambda^{ID}\epsilon_{ABCD}V^aV^b) + h.c.\big] &(5.205)\\
\hat{\mathcal{P}} &= \mathcal{P} - \bar\psi_A\chi_{BCD}\epsilon^{ABCD} &(5.206)\\
\hat{\mathcal{P}}^I_{AB} &= \mathcal{P}^I_{AB} - (\bar\psi_A\lambda^I_B + \epsilon_{ABCD}\bar\psi^C\lambda^{ID}) &(5.207)\\
& &(5.208)
\end{aligned}
$$

where $\mathcal{P} = \mathcal{P}_{\scriptscriptstyle N}\,d{\scriptscriptstyle N}$ and $\mathcal{P}^I_{AB} = \mathcal{P}^I_{AB\,i}\,d\phi^i$ are the vielbein of $\frac{SU(1,1)}{U(1)}$ and $\frac{SO(6,n')}{S(O(6)\times O(n'))}$ respectively. The fermion transformation laws are:

$$
\begin{aligned}
\delta\psi_A &= D\epsilon_A + a_1 T_{AB\,\mu\nu}\gamma^a\gamma^{\mu\nu}\epsilon^B V_a + \cdots &(5.209)\\
\delta\chi_{ABC} &= a_2\mathcal{P}_{\scriptscriptstyle N}\,\partial_\mu{\scriptscriptstyle N}\,\gamma^\mu\epsilon^D\epsilon_{ABCD} + a_3 T_{[AB\,\mu\nu}\gamma^{\mu\nu}\epsilon_{C]} + \cdots &(5.210)\\
\delta\lambda^I_A &= a_4\mathcal{P}^I_{AB\,i}\,\partial_a\phi^i\gamma^a\epsilon^B + a_5 T^{-I}_{\mu\nu}\gamma^{\mu\nu}\epsilon_A + \cdots &(5.211)
\end{aligned}
$$

where the 2–forms T_{AB} and T_I are defined in eq.(5.171). By integration of these two-forms we find the central and matter dyonic charges given in equations (5.173), (5.174). >From the equations (5.182),(5.183) for f, h and the definitions of the charges one easily finds:

$$\nabla^{SU(4)\times U(1)} Z_{AB} = \bar{Z}^I \mathcal{P}_{IAB} + \frac{1}{2}\epsilon_{ABCD}\bar{Z}^{CD}\mathcal{P} \tag{5.212}$$

$$\nabla^{SO(n')} Z_I = \frac{1}{2}\bar{Z}^{AB}\mathcal{P}_{IAB} + Z_I\bar{\mathcal{P}} \tag{5.213}$$

where $\frac{1}{2}\varepsilon_{ABCD}\bar{Z}^{CD} = \bar{Z}_{AB}$. In terms of the kinetic matrix (5.203) the invariant $\mathscr{V}_{BH}$ for the charges is given by, cf. (5.177),

$$\mathscr{V}_{BH} = \frac{1}{2}Z_{AB}\bar{Z}^{AB} + Z_I\bar{Z}^I = -\frac{1}{2}Q^t\mathcal{M}(\mathcal{N})Q \ . \tag{5.214}$$

The unique $SU(1,1)\times SO(6,n')$ invariant combination of the charges that is independent from the scalar fields is $I_1^2 - I_2\bar{I}_2$, so that

$$\mathscr{S} = \sqrt{|I_1^2 - I_2\bar{I}_2|} \ . \tag{5.215}$$

Here, I_1, I_2 and $\bar{I}_2$ are the three $SO(6,n')$ invariants given by

$$I_1 = \frac{1}{2}Z_{AB}\bar{Z}^{AB} - Z_I\bar{Z}^I \quad , \qquad I_2 = \frac{1}{4}\epsilon^{ABCD}Z_{AB}Z_{CD} - \bar{Z}_I\bar{Z}^I \ . \tag{5.216}$$

The $N = 3$ theory

In the $N = 3$ case [64] the coset space is:

$$G/H = \frac{SU(3,n')}{S(U(3)\times U(n'))} \tag{5.217}$$

and the field content is given by:

$$(V^a_\mu, \psi_{A\mu}, A^{AB}_\mu, \chi_{(L)}) \qquad A = 1,2,3 \qquad \text{(gravitational multiplet)} \tag{5.218}$$

$$(A_\mu, \lambda_A, \lambda_{(R)}, 3\,z)^I \qquad I = 1,\dots,n' \qquad \text{(vector multiplets)} \tag{5.219}$$

The transformation properties of the fields are given in Table 3. We consider the (local)

Table 3: Transformation properties of fields in $D = 4$, $N = 3$

	V^a_μ	$\psi_{A\mu}$	A^Λ_μ	$\chi_{(L)}$	λ^I_A	$\lambda^I_{(L)}$	L^Λ_{AB}	L^Λ_I	R_H
$SU(3,n')$	1	1	$3+n'$	1	1	1	$3+n'$	$3+n'$	-
$SU(3)$	1	3	1	1	3	1	$\bar{3}$	1	3
$SU(n')$	1	1	1	1	n'	n'	1	n'	n'
$U(1)$	0	$\frac{n'}{2}$	0	$3\frac{n'}{2}$	$3+\frac{n'}{2}$	$-3(1+\frac{n'}{2})$	n'	-3	$3+n'$

embedding of $SU(3,n')$ in $Sp(3+n',\mathbb{R})$ defined by the following dependence of the matrices f and h in terms of the G/H coset representative L,

$$f^\Lambda_{\ \Sigma} = \frac{1}{\sqrt{2}}(L^\Lambda_{\ AB}, \bar{L}^\Lambda_{\ I}) \tag{5.220}$$

$$h_{\Lambda\Sigma} = -i(\eta f\eta)_{\Lambda\Sigma} \qquad \eta = \begin{pmatrix} \mathbb{1}_{3\times 3} & 0 \\ 0 & -\mathbb{1}_{n'\times n'} \end{pmatrix} \tag{5.221}$$

where AB are antisymmetric $SU(3)$ indices, I is an index of $SU(n')$ and $\bar{L}^{\Lambda}{}_{I}$ denotes the complex conjugate of the coset representative. We have:

$$\mathcal{N}_{\Lambda\Sigma} = (hf^{-1})_{\Lambda\Sigma} = -i(\eta f\eta f^{-1})_{\Lambda\Sigma} \tag{5.222}$$

The supercovariant field strengths and the supercovariant scalar vielbein are:

$$\begin{aligned}
\hat{F}^{\Lambda} &= dA^{\Lambda} + \big[\frac{i}{2}f_I^{\Lambda}\bar{\lambda}_A^I\gamma_a\psi^A V^a - \frac{1}{2}f_{AB}^{\Lambda}\bar{\psi}^A\psi^B + if_{AB}^{\Lambda}\bar{\chi}_{(R)}\gamma_a\psi_C\epsilon^{ABC}V^a + h.c.\big]\\
\hat{\mathcal{P}}_I{}^A &= \mathcal{P}_I{}^A - \bar{\lambda}_B^I\psi_C\epsilon^{ABC} - \bar{\lambda}_{I(R)}\psi^A
\end{aligned} \tag{5.223}$$

where the only nonvanishing entries of the vierbein $\mathcal{P}$ are

$$\mathcal{P}_I^A = \frac{1}{2}\epsilon^{ABC}\mathcal{P}_{IBC} = \mathcal{P}_{I\,i}^A dz^i \tag{5.224}$$

z^i being the (complex) coordinates of G/H. The chiral fermions transformation laws are given by:

$$\delta\psi_A = D\epsilon_A + 2iT_{AB\,\mu\nu}\gamma^a\gamma^{\mu\nu}V_a\epsilon^B + \cdots \tag{5.225}$$
$$\delta\chi_{(L)} = 1/2\,T_{AB\,\mu\nu}\gamma^{\mu\nu}\epsilon_C\epsilon^{ABC} + \cdots \tag{5.226}$$
$$\delta\lambda_{IA} = -i\mathcal{P}_I{}^B{}_i\partial_\mu z^i\gamma^\mu\epsilon^C\epsilon_{ABC} + T_{I\,\mu\nu}\gamma^{\mu\nu}\epsilon_A + \cdots \tag{5.227}$$
$$\delta\lambda^I_{(L)} = i\mathcal{P}_I{}^A{}_i\partial_\mu z^i\gamma^\mu\epsilon_A + \cdots \tag{5.228}$$

where T_{AB} and T_I have the general form given in equation (5.171). >From the general form of the equations (5.182), (5.183) for f and h we find:

$$\nabla f^{\Lambda}_{AB} = f^{\Lambda}{}_I\mathcal{P}^I{}_{AB}\ , \tag{5.229}$$
$$\nabla h^{\Lambda}_{AB} = h^{\Lambda}{}_I\mathcal{P}^I{}_{AB}\ , \tag{5.230}$$
$$\nabla f^{\Lambda}_{\bar{I}} = \frac{1}{2}\bar{f}^{\Lambda CD}\mathcal{P}_{CD\bar{I}}\ , \tag{5.231}$$
$$\nabla h^{\Lambda}_{\bar{I}} = \frac{1}{2}\bar{h}^{\Lambda CD}\mathcal{P}_{CD\bar{I}}\ . \tag{5.232}$$

According to the general study of Section 4.1, using (5.173), (5.174) one finds

$$\nabla^{(H)}Z_{AB} = \bar{Z}^I\mathcal{P}_I{}^C\epsilon_{ABC} \tag{5.233}$$
$$\nabla^{(H)}Z_I = \frac{1}{2}\bar{Z}^{AB}\mathcal{P}_I{}^C\epsilon_{ABC} \tag{5.234}$$

and the formula for the potential, cf. (5.177),

$$\mathscr{V}_{BH} = \frac{1}{2}Z^{AB}\bar{Z}_{AB} + Z^I\bar{Z}_I = -\frac{1}{2}Q^t\mathcal{M}(\mathcal{N})Q \tag{5.235}$$

where the matrix $\mathcal{M}(\mathcal{N})$ has the same form as in equation (5.179) in terms of the kinetic matrix $\mathcal{N}$ of equation (5.222), and Q is the charge vector $Q = \binom{g}{e}$.

The $G = SU(3,n')$ invariant is $Z^A\bar{Z}_A - Z_I\bar{Z}^I$ (one can check that $\partial_i(Z^A\bar{Z}_A - Z_I\bar{Z}^I) = \nabla_i^{(H)}(Z^A\bar{Z}_A - Z_I\bar{Z}^I) = 0$) so that

$$\mathscr{S} = |Z^A\bar{Z}_A - Z_I\bar{Z}^I|\ . \tag{5.236}$$

The $N = 5$ theory

For $N > 4$ the only available supermultiplet is the gravitational one, so that $H_{\rm matter} = 1$. The coset manifold of the scalars of the $N = 5$ theory [33] is:

$$G/H = \frac{SU(5,1)}{U(5)} \tag{5.237}$$

The field content and the group assignments are displayed in Table 4.

Table 4: Transformation properties of fields in $D = 4$, $N = 5$

	V^a	ψ_A;	χ_{ABC}, χ_L	$A^{\Lambda\Sigma}$	L^x_A	R_H
$SU(5,1)$	1	1	1	-	6	-
$SU(5)$	1	5	$(10, 1)$	1	5	$\bar{5}$
$U(1)$	0	$\frac{1}{2}$	$(\frac{3}{2}, -\frac{5}{2})$	0	1	2

In Table 4 the incides $x, y, \dots = 1, \dots, 6$ and $A, B, C, \dots = 1, \dots, 5$ are indices of the fundamental representations of $SU(5,1)$ and $SU(5)$, respectively. L^x_A denotes as usual the coset representative in the fundamental representation of $SU(5,1)$. The antisymmetric couple $\Lambda\Sigma$, $\Lambda, \Sigma = 1, \dots, 5$, enumerates the ten vector potentials. The local embedding of $SU(5,1)$ into the Gaillard-Zumino group $Usp(10,10)$ is given in terms of the three-times antisymmetric representation of $SU(5,1)$, this is a 20 dimensional complex representation, we denote by t^{xyz} a generic element. This representation is reducible to a complex 10 dimensional one by imposing the self-duality condition

$$\bar{t}^{\bar{x}\bar{y}\bar{z}} = \frac{1}{3!}\epsilon^{\bar{x}\bar{y}\bar{z}}{}_{uvw}t^{uvw} \tag{5.238}$$

here indices are raised with the $SU(5,1)$ hermitian structure $\eta = diag(1,1,1,1,1,-1)$. The self duality condition (5.238) is compatible with the $SU(5,1)$ action (on $\bar{t}^{\bar{x}\bar{y}\bar{z}}$ acts the complex conjugate of the three-times antisymmetric of $SU(5,1)$). Due to the self-duality condition we can decompose t^{xyz} as follows:

$$t^{xyz} = \begin{pmatrix} t^{\Lambda\Sigma 6} \\ \bar{t}^{\bar{\Lambda}\bar{\Sigma}\bar{6}} \end{pmatrix} \tag{5.239}$$

where $(\Lambda, \Sigma, \dots = 1, \dots, 5)$. In the following we set $t^{\Lambda\Sigma} \equiv t^{\Lambda\Sigma 6}$, $\bar{t}^{\bar{\Lambda}\bar{\Sigma}} \equiv \bar{t}^{\bar{\Lambda}\bar{\Sigma}\bar{6}}$, $\bar{t}_{\Lambda\Sigma} \equiv \bar{t}_{\Lambda\Sigma 6} = -\bar{t}_{\Lambda\Sigma}{}^{\bar{6}}$. The symplectic structure in this complex basis is given by the matrix $\begin{pmatrix} 0 & -1\!\!1 \\ 1\!\!1 & 0 \end{pmatrix}$,

$$\langle t, \ell\rangle := \frac{1}{2}(t^{\Lambda\Sigma}, \bar{t}^{\bar{\Lambda}\bar{\Sigma}}) \begin{pmatrix} 0 & -\delta_{\Lambda\Sigma\,\bar{\Gamma}\bar{\Pi}} \\ \delta_{\bar{\Lambda}\bar{\Sigma}\,\Gamma\Pi} & 0 \end{pmatrix} \begin{pmatrix} \ell^{\Gamma\Pi} \\ \ell^{\bar{\Gamma}\bar{\Pi}} \end{pmatrix} \tag{5.240}$$

$$= \frac{1}{2}t^{\Lambda\Sigma}\bar{\ell}_{\Lambda\Sigma} - \frac{1}{2}\bar{t}_{\Lambda\Sigma}\ell^{\Lambda\Sigma}$$

$$= \frac{1}{3!3!}t^{xyz}\varepsilon_{xyzuvw}\ell^{uvw} \tag{5.241}$$

this last equality implies that the $SU(5,1)$ action preserves the symplectic structure. We have thus embedded* $SU(5,1)$ into $Sp(20,\mathbb{R})$ (in the complex basis).

*Strictly speaking we have immersed $SU(5,1)$ into $Sp(20,\mathbb{R})$, in fact this map is a local embedding but fails to be injective, indeed the three $SU(5,1)$ elements $\sqrt[3]{1}\,1\!\!1$ are all mapped into the identity element of $Sp(20,\mathbb{R})$.

The 20 dimensional real vector $(F^{\Lambda\Sigma}, G_{\Lambda\Sigma})$ transforms under the 20 of $SU(5,1)$, as well as, for fixed AB, each of the 20 dimensional vectors $\binom{f^{\Lambda\Sigma}{}_{AB}}{h_{\Lambda\Sigma AB}}$ of the embedding matrix:

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} f + ih & \bar{f} + i\bar{h} \\ f - ih & \bar{f} - i\bar{h} \end{pmatrix} . \tag{5.242}$$

The supercovariant field strengths and vielbein are:

$$\hat{F}^{\Lambda\Sigma} = dA^{\Lambda\Sigma} + \left[f^{\Lambda\Sigma}{}_{AB}(a_1 \bar{\psi}^A \psi^B + a_2 \bar{\psi}_C \gamma_a \chi^{ABC} V^a) + h.c. \right] \tag{5.243}$$

$$\hat{\mathcal{P}}_{ABCD} = \mathcal{P}_{ABCD} - \bar{\chi}_{[ABC}\psi_{D]} - \epsilon_{ABCDE}\bar{\chi}^{(R)}\psi^E \tag{5.244}$$

where $\mathcal{P}_{ABCD} = \epsilon_{ABCDF}\mathcal{P}^F$ is the complex vielbein, completely antisymmetric in $SU(5)$ indices and $\overline{\mathcal{P}_{ABCD}} = \bar{\mathcal{P}}^{ABCD}$.
The fermion transformation laws are:

$$\delta\psi_A = D\epsilon_A + a_3 T_{AB\,\mu\nu}\gamma^a\gamma^{\mu\nu}\epsilon^B V_a + \cdots \tag{5.245}$$

$$\delta\chi_{ABC} = a_4 \mathcal{P}_{ABCD\,i}\partial_\mu\phi^i\gamma^\mu\epsilon^D + a_5 T_{[AB\,\mu\nu}\gamma^{\mu\nu}\epsilon_{C]} + \cdots \tag{5.246}$$

$$\delta\chi_{(L)} = a_6 \bar{\mathcal{P}}^{ABCD}{}_{\bar{\imath}}\,\partial_\mu\bar{\phi}^{\bar{\imath}}\gamma^\mu\epsilon^E\epsilon_{ABCDE} + \cdots \tag{5.247}$$

where:

$$T_{AB} = \frac{1}{2}(h_{\Lambda\Sigma AB}F^{\Lambda\Sigma} - f^{\Lambda\Sigma}{}_{AB}G_{\Lambda\Sigma}) \tag{5.248}$$

$$\mathcal{N}_{\Lambda\Sigma\,\Delta\Pi} = \frac{1}{2}h_{\Lambda\Sigma\,AB}(f^{-1})^{AB}{}_{\Delta\Pi} . \tag{5.249}$$

With a by now familiar procedure one finds the following (complex) central charges:

$$Z_{AB} = i\overline{V(\phi_\infty)}^{-1} Q \tag{5.250}$$

where the charge vector is

$$Q = \begin{pmatrix} p^{\Lambda\Sigma} \\ q_{\Lambda\Sigma} \end{pmatrix} = \begin{pmatrix} \frac{1}{4\pi}\int_{S^2} F^{\Lambda\Sigma} \\ \frac{1}{4\pi}\int_{S^2} G_{\Lambda\Sigma} \end{pmatrix} \tag{5.251}$$

and ϕ_∞ is the constant value assumed by the scalar fields at spatial infinity. >From the equations (Maurer-Cartan equations)

$$\nabla^{(U(5))} f^{\Lambda\Sigma}{}_{AB} = \frac{1}{2}\bar{f}^{\Lambda\Sigma\,CD}\mathcal{P}_{ABCD} \tag{5.252}$$

and the analogous one for h we find:

$$\nabla^{(U(5))} Z_{AB} = \frac{1}{2}\bar{Z}^{CD}\mathcal{P}_{ABCD} . \tag{5.253}$$

Finally, the formula for the potential is, cf. (5.177),

$$\mathscr{V}_{BH} = \frac{1}{2}\bar{Z}^{AB}Z_{AB} = -\frac{1}{2}Q^t\mathcal{M}(\mathcal{N})Q \tag{5.254}$$

where the matrix $\mathcal{M}(\mathcal{N})$ has exactly the same form as in equation (5.179), and $\mathcal{N}$ is given in (5.249).

For $SU(5,1)$ there are only two $U(5)$ quartic invariants. In terms of the matrix $A_A^{\ B} = Z_{AC}\bar{Z}^{CB}$ they are:

$$\mathrm{Tr}A = Z_{AB}\bar{Z}^{BA} \quad , \qquad \mathrm{Tr}(A^2) = Z_{AB}\bar{Z}^{BC}Z_{CD}\bar{Z}^{DA} . \tag{5.255}$$

The $SU(5,1)$ invariant expression is

$$\mathscr{S} = \frac{1}{2}\sqrt{|4\mathrm{Tr}(A^2) - (\mathrm{Tr}A)^2|} . \tag{5.256}$$

The $N = 6$ theory

The scalar manifold of the $N = 6$ theory has the coset structure [65]:

$$G/H = \frac{SO^\star(12)}{U(6)} \tag{5.257}$$

We recall that $SO^\star(2n)$ is the real form of $O(2n,\mathbb{C})$ defined by the relation:

$$L^\dagger C L = C \ , \qquad C = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \tag{5.258}$$

The field content and transformation properties are given in Table 5, where $A, B, C = 1, \cdots, 6$

Table 5: Transformation properties of fields in $D = 4$, $N = 6$

	V^a	ψ_A	χ_{ABC}, χ_A	A^Λ	S_r^α	R_H
$SO^\star(12)$	1	1	1	-	32	-
$SU(6)$	1	6	$(20+6)$	1	$(15,1)+(\bar{15},\bar{1})$	$\bar{15}$
$U(1)$	0	$\frac{1}{2}$	$(\frac{3}{2},-\frac{5}{2})$	0	$(1,-3)+(-1,3)$	2

are $SU(6)$ indices in the fundamental representation and $\Lambda = 1, \cdots, 16$. The 32 spinor representation of $SO^\star(12)$ can be given in terms of a $Sp(32,\mathbb{R})$ matrix, which in the complex basis we denote by S_r^α ($\alpha, r = 1, \cdots, 32$). It is the double cover of $SO^\star(12)$ that embeds in $Sp(32,\mathbb{R})$ and therefore the duality group is this spin group. Employing the usual notation we may set:

$$S_r^\alpha = \frac{1}{\sqrt{2}} \begin{pmatrix} f^\Lambda_M + \mathrm{i}h_{\Lambda M} & \bar{f}^\Lambda_M + \mathrm{i}\bar{h}_{\Lambda M} \\ f^\Lambda_M - \mathrm{i}h_{\Lambda M} & \bar{f}^\Lambda_M - \mathrm{i}\bar{h}_{\Lambda M} \end{pmatrix} \tag{5.259}$$

where $\Lambda, M = 1, \cdots, 16$. With respect to $SU(6)$, the sixteen symplectic vectors $(f^\Lambda_M, h_{\Lambda M})$, $(M = 1, \cdots, 16)$ are reducible into the antisymmetric 15 dimensional representation plus a singlet of $SU(6)$:

$$(f^\Lambda_M, h_{\Lambda M}) \to (f^\Lambda{}_{AB}, h_{\Lambda AB}) + (\bar{f}^\Lambda, \bar{h}_\Lambda) \ . \tag{5.260}$$

It is precisely the existence of a $SU(6)$ singlet which allows for the Special Geometry structure of $\frac{SO^*(12)}{U(6)}$ (cf. (5.367), (5.368))*. Note that the element S_r^α has no definite $U(1)$ weight since the submatrices $f^\Lambda{}_{AB}, \bar{f}^\Lambda$ have the weights 1 and -3 respectively. The vielbein matrix is

$$\mathcal{P} = \begin{pmatrix} \mathcal{P}_{ABCD} & \mathcal{P}_{AB} \\ \mathcal{P}_{CD} & 0 \end{pmatrix} \ , \tag{5.261}$$

where

$$\mathcal{P}_{AB} = \frac{1}{4!}\epsilon_{ABCDEF}\mathcal{P}^{CDEF}; \quad \bar{\mathcal{P}}^{AB} = \overline{\mathcal{P}_{AB}} \ . \tag{5.262}$$

The supercovariant field strengths and the coset manifold vielbein have the following expression:

$$\hat{F}^\Lambda = dA^\Lambda + \big[f^\Lambda{}_{AB}(a_1\bar{\psi}^A\psi^B + a_2\bar{\psi}_C\gamma_a\chi^{ABC}V^a) + a_3 f^\Lambda\bar{\psi}_C\gamma_a\chi^C V^a + h.c.\big] \tag{5.263}$$

$$\hat{\mathcal{P}}_{ABCD} = \mathcal{P}_{ABCD} - \bar{\chi}_{[ABC}\psi_{D]} - \epsilon_{ABCDEF}\bar{\chi}^E\psi^F \tag{5.264}$$

*Due to its Special Geometry structure the coset space $\frac{SO^*(12)}{U(6)}$ is also the scalar manifold of an $N = 2$ supergravity. The two supergravity theories have the same bosonic fields however the fermion sector is different.

The fermion transformation laws are:

$$\delta\psi_A = D\epsilon_A + b_1 T_{AB\,\mu\nu}\gamma^a\gamma^{\mu\nu}\epsilon^B V_a + \cdots \tag{5.265}$$
$$\delta\chi_{ABC} = b_2 \mathcal{P}_{ABCD\,i}\partial_a z^i \gamma^a \epsilon^D + b_3 T_{[AB\,ab}\gamma^{ab}\epsilon_{C]} + \cdots \tag{5.266}$$
$$\delta\chi_A = b_4 \mathcal{P}^{BCDE}{}_i \partial_a z^i \gamma^a \epsilon^F \epsilon_{ABCDEF} + b_5 T_{ab}\gamma^{ab}\epsilon_A + \cdots \tag{5.267}$$

where according to the general definition (5.171):

$$\begin{aligned} T_{AB} &= h_{\Lambda\,AB}F^\Lambda - f^\Lambda{}_{AB}G_\Lambda \\ \bar{T} &= \bar{h}_\Lambda F^\Lambda - \bar{f}^\Lambda G_\Lambda \end{aligned} \tag{5.268}$$

With the usual procedure we have the following complex dyonic central charges:

$$Z_{AB} = h_{\Lambda AB}p^\Lambda - f^\Lambda_{AB}q_\Lambda \tag{5.269}$$
$$\bar{Z} = \bar{h}_\Lambda p^\Lambda - \bar{f}^\Lambda q_\Lambda \tag{5.270}$$

in the $\overline{15}$ (recall (5.169)) and singlet representation of $SU(6)$ respectively. Notice that although we have 16 graviphotons, only 15 central charges are present in the supersymmetry algebra. The singlet charge plays a role analogous to a "matter" charge (hence our notation $\bar{Z}$, $\bar{f}^\Lambda$, $\bar{h}_\Lambda$). The charges differential relations are

$$\nabla^{(U(6))}Z_{AB} = \frac{1}{2}\bar{Z}^{CD}\mathcal{P}_{ABCD} + \frac{1}{4!}Z\epsilon_{ABCDEF}\mathcal{P}^{CDEF} \tag{5.271}$$
$$\nabla^{(U(1))}\bar{Z} = \frac{1}{2!4!}\bar{Z}^{AB}\epsilon_{ABCDEF}\mathcal{P}^{CDEF} \tag{5.272}$$

and the formula for the potential reads, cf. (5.177),

$$\mathscr{V}_{BH} = \frac{1}{2}\bar{Z}^{AB}Z_{AB} + \bar{Z}Z = -\frac{1}{2}Q^t\mathcal{M}(\mathcal{N})Q\ . \tag{5.273}$$

The quartic $U(6)$ invariants are

$$I_1 = (TrA)^2 \tag{5.274}$$
$$I_2 = Tr(A^2) \tag{5.275}$$
$$I_3 = \frac{1}{2^3 3!}\mathrm{Re}\,(\epsilon^{ABCDEF}Z_{AB}Z_{CD}Z_{EF}Z) \tag{5.276}$$
$$I_4 = (TrA)Z\bar{Z} \tag{5.277}$$
$$I_5 = Z^2\bar{Z}^2 \tag{5.278}$$

where $A_A^B = Z_{AC}\bar{Z}^{CB}$. The unique $SO^*(12)$ invariant is

$$\mathscr{S} = \frac{1}{2}\sqrt{|4I_2 - I_1 + 32I_3 + 4I_4 + 4I_5|}\ . \tag{5.279}$$

The $N = 8$ theory

In the $N = 8$ case [5] the coset manifold is:

$$G/H = \frac{E_{7(7)}}{SU(8)/\mathbb{Z}_2}. \tag{5.280}$$

The field content and group assignments are given in Table 6.

Table 6: Field content and group assignments in $D=4$, $N=8$ supergravity

	V^a	ψ_A	$A^{\Lambda\Sigma}$	χ_{ABC}	S^α_r	R_H
$E_{7(7)}$	1	1	-	1	56	-
$SU(8)$	1	8	1	56	$28+\overline{28}$	70

The embedding in $Sp(56,\mathbb{R})$ is automatically realized because the 56 defining representation of $E_{7(7)}$ is a real symplectic representation. The components of the f and h matrices and their complex conjugates are

$$f^{\Lambda\Sigma}_{\ \ AB}\ ,\quad h_{\Lambda\Sigma AB}\ ,\quad \bar{f}_{\Lambda\Sigma}^{\ \ AB}\ ,\quad \bar{h}^{\Lambda\Sigma\, AB}\ , \tag{5.281}$$

here $\Lambda\Sigma, AB$ are couples of antisymmetric indices, with Λ,Σ,A,B running from 1 to 8. The 70 under which the vielbein of G/H transform is obtained from the four times antisymmetric of $SU(8)$ by imposing the self duality condition

$$\bar{t}^{\bar{A}\bar{B}\bar{C}\bar{D}} = \frac{1}{4!}\epsilon^{\bar{A}\bar{B}\bar{C}\bar{D}}{}_{A'B'C'D'}t^{A'B'C'D'} \tag{5.282}$$

The supercovariant field strengths and coset manifold vielbein are:

$$\hat{F}^{\Lambda\Sigma} = dA^{\Lambda\Sigma} + [f^{\Lambda\Sigma}_{\ \ AB}(a_1\bar{\psi}^A\psi^B + a_2\bar{\chi}^{ABC}\gamma_a\psi_C V^a) + h.c.] \tag{5.283}$$

$$\hat{\mathcal{P}}_{ABCD} = \mathcal{P}_{ABCD} - \bar{\chi}_{[ABC}\psi_{D]} + h.c. \tag{5.284}$$

where $\mathcal{P}_{ABCD} = \frac{1}{4!}\epsilon_{ABCDEFGH}\bar{\mathcal{P}}^{EFGH} \equiv (L^{-1}\nabla^{SU(8)}L)_{AB\,CD} = \mathcal{P}_{ABCD\,i}d\phi^i$ (ϕ^i coordinates of G/H). In the complex basis the vielbein $\mathcal{P}_{ABCD}$ of G/H are 28×28 matrices completely antisymmetric and self dual as in (5.282). The fermion transformation laws are given by:

$$\delta\psi_A = D\epsilon_A + a_3 T_{AB\,\mu\nu}\gamma^a\gamma^{\mu\nu}\epsilon^B V_a + \cdots \tag{5.285}$$

$$\delta\chi_{ABC} = a_4\mathcal{P}_{ABCD\,i}\partial_a\phi^i\gamma^a\epsilon^D + a_5 T_{[AB\,\mu\nu}\gamma^{\mu\nu}\epsilon_{C]} + \cdots \tag{5.286}$$

where:

$$T_{AB} = \frac{1}{2}(h_{\Lambda\Sigma AB}F^{\Lambda\Sigma} - f^{\Lambda\Sigma}_{\ \ AB}G_{\Lambda\Sigma}) \tag{5.287}$$

with:

$$\mathcal{N}_{\Lambda\Sigma\,\Gamma\Delta} = \frac{1}{2}h_{\Lambda\Sigma AB}(f^{-1})^{AB}_{\ \ \Gamma\Delta}\ . \tag{5.288}$$

With the usual manipulations we obtain the central charges:

$$Z_{AB} = \frac{1}{2}(h_{\Lambda\Sigma AB}p^{\Lambda\Sigma} - f^{\Lambda\Sigma}_{\ \ AB}q_{\Lambda\Sigma}), \tag{5.289}$$

the differential relations:

$$\nabla^{SU(8)}Z_{\ AB} = \frac{1}{2}\bar{Z}^{\ CD}\mathcal{P}_{ABCD} \tag{5.290}$$

and the formula for the potential, cf. (5.177),

$$\mathcal{V}_{BH} = \frac{1}{2}\bar{Z}^{AB}Z_{AB} = -\frac{1}{2}Q^t\mathcal{M}(\mathcal{N})Q \tag{5.291}$$

where the matrix $\mathcal{M}(\mathcal{N})$ is given in equation (5.179), and $\mathcal{N}$ in (5.288).

For $N=8$ the $SU(8)$ invariants are

$$I_1 = (TrA)^2 \tag{5.292}$$

$$I_2 = Tr(A^2) \tag{5.293}$$

$$I_3 = Pf\ Z = \frac{1}{2^4 4!}\epsilon^{ABCDEFGH}Z_{AB}Z_{CD}Z_{EF}Z_{GH} \tag{5.294}$$

where PfZ denotes the Pfaffian of the antisymmetric matrix $(Z_{AB})_{A,B=1,...8}$, and where $A_A^{\ B} = Z_{AC}\bar{Z}^{CB}$. One finds the following $E_{7(7)}$ invariant [44]:

$$\mathscr{S} = \frac{1}{2}\sqrt{|4\text{Tr}(A^2) - (TrA)^2 + 32\text{Re}\,(Pf\,Z)|} \tag{5.295}$$

For a very recent study of $E_{7(7)}$ duality rotations and of the corresponding conserved charges see [66].

Electric subgroups and the $D = 4$ and $N = 8$ theory.

A duality rotation is really a strong-weak duality if there is a rotation between electric and magnetic fields, more precisely if some of the rotated field strengths F'^{Λ} depend on the initial dual fields G^{Σ}, i.e. if the submatrix $B \neq 0$ in the symplectic matrix $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$. Only in this case the gauge kinetic term may transform nonlinearly, via a fractional transformation. On the other hand, under infinitesimal duality rotations $\left(\begin{smallmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{smallmatrix}\right) + \left(\begin{smallmatrix} a & 0 \\ c & d \end{smallmatrix}\right)$, with $b = 0$, the lagrangian changes by a total derivative so that (in the absence of instantons) these transformations are symmetries of the action, not just of the equation of motion. Furthermore if $c = 0$ the lagrangian itself is invariant.

We call electric any subgroup G_e of the duality group G with the property that it (locally) embeds in the symplectic group via matrices $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ with $B = 0$. The parameter space of true strong-weak duality rotations is G/G_e.

The electric subgroup of $Sp(2n, \mathbb{R})$ is the subgroup of all matrices of the kind

$$\begin{pmatrix} A & 0 \\ C & A^{t^{-1}} \end{pmatrix} \; ; \tag{5.296}$$

we denote it by $Sp_e(2n, \mathbb{R})$. It is *the* electric subgroup because any other electric subgroup is included in $Sp_e(2n, \mathbb{R})$. This subgroup is maximal in $Sp(2n, \mathbb{R})$ (see for example the appendices in[50, 68]). In particular if an action is invariant under infinitesimal $Sp_e(2n, \mathbb{R})$ transformations, and if the equations of motion admit also a $\pi/2$ duality rotation symmetry $F^{\Lambda} \to G^{\Lambda}$, $G^{\Lambda} \to -F^{\Lambda}$ for one or more indices Λ (no transformation on the other indices) then the theory has $Sp(2n, \mathbb{R})$ duality.

It is easy to generalize the results of Section 2.2 and prove that duality symmetry under these $\pi/2$ rotations is equivalent to the following invariance property of the lagrangian under the Legendre transformation associated to F^{Λ},

$$\mathcal{L}_D(F, \mathcal{N}') = \mathcal{L}(\mathcal{F}, \mathcal{N}) \; , \tag{5.297}$$

where $\mathcal{N}' = (\mathcal{C} + \mathcal{D}\mathcal{N})(\mathcal{A} + \mathcal{B}\mathcal{N})^{-\infty}$ are the transformed scalar fields, the matrix $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ implementing the $\pi/2$ rotation $F^{\Lambda} \to G^{\Lambda}$, $G^{\Lambda} \to -F^{\Lambda}$. We conclude that $Sp(2n, \mathbb{R})$ duality symmetry holds if there is $Sp_e(2n, \mathbb{R})$ symmetry and if the lagrangian satisfies (5.297).

When the duality group G is not $Sp(2n, \mathbb{R})$ then there may exist different maximal electric subgroups of G, say G_e and G'_e. Consider now a theory with G duality symmetry, the electric subgroup G_e hints at the existence of an action $S = \int \mathcal{L}$ invariant under the Lie algebra Lie(G_e) and under Legendre transformation that are $\pi/2$ duality rotation in G. Similarly G'_e leads to a different action $S' = \int \mathcal{L}'$ that is invariant under Lie(G'_e) and under Legendre transformations that are $\pi/2$ duality rotation in G. The equations of motion of both actions have G duality symmetry. They are equivalent if $\mathcal{L}$ and $\mathcal{L}'$ are related by a Legendre transformation. Since $\mathcal{L}'(F, \mathcal{N}') \neq \mathcal{L}(\mathcal{F}, \mathcal{N})$, this Legendre transformation cannot be a duality symmetry, it is a $\pi/2$ rotation $F^{\Lambda} \to G^{\Lambda}$, $G^{\Lambda} \to -F^{\Lambda}$ that is not in G, this is possible since $G \neq Sp(2n, \mathbb{R})$.

As an example consider the $G_e = SL(8, \mathbb{R})$ symmetry of the $N = 8$, $D = 4$ supergravity lagrangian whose duality group is $G = E_{7,(7)}$ this is the formulation of Cremmer-Julia. An alternative formulation, obtained from dimensional reduction of the $D = 5$ supergravity, exhibits an electric group $G'_e = [E_{6,(6)} \times SO(1,1)] \ltimes T_{27}$ where the nonsemisimple group G'_e is realized as a lower triangular subgroup of $E_{7,(7)}$ in its fundamental (symplectic) 56 dimensional representation. G_e and G'_e are both maximal subgroups of $E_{7(7)}$. The corrseponding lagrangians can be related only after a proper duality rotation of electric and magnetic fields which involves a suitable Legendre transformation.

A way to construct new supergravity theories is to promote a compact rigid electric subgroup symmetry to a local symmetry, thus constructing gauged supergravity models (see for a recent review [67], and references therein). Inequivalent choices of electric subgroups give different gauged supergravities. Consider again $D = 4$, $N = 8$ supergravity. The maximal compact subgroups of $G_e = SL(8, \mathbb{R})$ and of $G'_e = [E_{6,(6)} \times SO(1,1)] \ltimes T_{27}$ are $SO(8)$ and $Sp(8) = U(16) \cap Sp(16, \mathbb{C})$ respectively. The gauging of $SO(8)$ corresponds to the gauged $N = 8$ supergravity of De Witt and Nicolai [33]. As shown in [34] the gauging of the nonsemisimple group $U(1) \ltimes T_{27} \subset G'_e$ corresponds to the gauging of a flat group in the sense of Scherk and Schwarz dimensional reduction [35], and gives the massive deformation of the $N = 8$ supergravity as obtained by Cremmer, Scherk and Schwarz [36].

5.5 Special Geometry and $N = 2$ Supergravity

In the case of $N = 2$ supergravity the requirements imposed by supersymmetry on the scalar manifold M_{scalar} of the theory dictate that it should be the following direct product: $M_{scalar} = M \times M^Q$ where M is a special Kähler manifold of complex dimension n and M^Q a quaternionic manifold of real dimension $4n_H$, here n and n_H are respectively the number of vector multiplets and hypermultiplets contained in the theory. The direct product structure imposed by supersymmetry precisely reflects the fact that the quaternionic and special Kähler scalars belong to different supermultiplets. We do not discuss the hypermultiplets any further and refer to [77] for the full structure of N=2 supergravity. Since we are concerned with duality rotations we here concentrate our attention to an $N = 2$ supergravity where the graviton multiplet, containing besides the graviton $g_{\mu\nu}$ also a graviphoton A^0_μ, is coupled to n' vector multiplets. Such a theory has a bosonic action of type (5.151) where the number of (real) gauge fields is $n = 1 + n'$ and the number of (real) scalar fields is $2n'$. Compatibiliy of their couplings with local $N = 2$ supersymmetry lead to the formulation of special Kgeometry [75],[76].

The formalism we have developed so far for the $D = 4$, $N > 2$ theories is completely determined by the (local) embedding of the coset representative of the scalar manifold $M = G/H$ in $Sp(2n, \mathbb{R})$. It leads to a flat -actually a trivial- symplectic bundle with local symplectic sections $\mathcal{V}_\eta$, determined by the symplectic matrix V, or equivalently by the matrices f and h. We want now to show that these matrices, the differential relations among charges and their quadratic invariant $\mathscr{V}_{BH}$ (5.177) are also central for the description of $N = 2$ matter-coupled supergravity. This follows essentially from the fact that, though the scalar manifold M of the $N = 2$ theory is not in general a coset manifold, nevertheless, as for the $N > 2$ theories, we have a flat symplectic bundle associated to M, with symplectic sections $\mathcal{V}_\eta$. While the formalism is very similar there is a difference, the bundle is not a trivial bundle anymore, and it is in virtue of duality rotations that the theory can be globally defined on M.

In the next section we study the geometry of the scalar manifold M and in detail its associated flat symplectic bundle. Then in Section 5.2 we see how, in analogy with $N > 2$ supergravities, the flat symplectic bundle geometry of M enters the supersymmetry transformations laws of $N = 2$ supergravity and the differential relations among the matter and central charges.

5.5.1 Special Geometry

There are two kinds of special geometries: rigid and local. While rigid special Kmanifolds are the target space of the scalar fields present in the vector multiplets of $N = 2$ Yang Mills theories, the (local) special Kmanifolds, in the mathematical literature called projective special Kmanifolds, describe the target space of the scalar fields in the vector multiplets of $N = 2$ supergravity (that has local supersymmetry). In order to describe the structure of a (local or projective) special Kmanifold it is instructive to recall that of rigid Kmanifold.

Rigid Special Geometry

In short a rigid special Kmanifold is a Kmanifold M that has a flat connection on its tangent bundle. This connection must then be compatible with the symplectic and complex structure of M.

More precisely, following [49], see also [50], a **rigid special Kstructure** on a Kmanifold M with Kform K is a connection ∇ that is real, flat, torsionfree, compatible with the symplectic structure ω:

$$\nabla\omega = 0 \tag{5.298}$$

and compatible with the almost complex structure J of M:

$$d_\nabla J = 0 \tag{5.299}$$

where $d_\nabla : \Omega^1(TM) \to \Omega^2(TM)$ is the covariant exterior derivative on vector-valued forms. Explicitly, if $J = J^\xi \, \partial_\xi$ where J^ξ are 1-forms, and $\nabla\partial_\xi = A^\zeta{}_\xi \, \partial_\zeta$, with $A^\zeta{}_\xi$ 1-forms, then $d_\nabla J = dJ^\xi \, \partial_\xi - J^\xi \wedge A^\zeta{}_\xi \, \partial_\zeta = (dJ^\xi + A^\xi{}_\zeta \wedge J^\zeta) \, \partial_\xi$. Notice that the torsionfree condition can be similarly written $d_\nabla I = 0$, where I is the identity map in TM, locally $I = dx^\xi \otimes \partial_\xi$. The two conditions $d_\nabla J = 0$, $d_\nabla I = 0$ for the real connection ∇ can be written in the complexified tangent bundle simply as

$$d_\nabla \pi^{1,0} = 0 \ , \tag{5.300}$$

where $\pi^{1,0}$ is the projection onto the $(1,0)$ part of the complexified tangent bundle; locally $\pi^{1,0} = dz^i \otimes \frac{\partial}{\partial z^i}$.

The flatness condition is equivalent to require the existence of a covering of M with local frames $\{\mathsf{e}_\xi\}$ that are covariantly constant, $\nabla \mathsf{e}_\xi = 0$. The corresponding transition functions of the real tangent bundle TM are therefore constant invertible matrices; compatibility with the symplectic structure, equation (5.298), further implies that these matrices belong to the fundamental of $Sp(2n, \mathbb{R})$, where $2n$ is the real dimension of M (each frame $\{\mathsf{e}_\xi\}$ can be chosen to have mutually symplectic vectors e_ξ).

Flatness of ∇ (i.e., the vanishing of the curvature R_∇ or equivalently $d_\nabla^2 = 0$) implies that (5.300) is equivalent to the existence of a local complex vector field ξ that satisfies

$$\nabla\xi = \pi^{1,0} \tag{5.301}$$

[hint: in a flat reference frame $d_\nabla = d$, and Poincaré lemma for d implies that any d_∇-closed section is also d_∇-exact]. Studying the components of this vector field (with respect to a flat Darboux coordinate system) we obtain the existence of local holomorphic coordinates on M, called special coordinates, their transition functions are constant $Sp(2n, \mathbb{R})$ matrices, so that the holomorphic tangent bundle TM is a flat symplectic holomorphic one. Corresponding to these special coordinates we have a holomorhic function $\mathcal{F}$, the holomorphic prepotential. In terms of this data the Kpotential and the Kform read

$$\mathcal{K} = \frac{1}{2}\mathrm{Im}\big(\frac{\partial \mathcal{F}}{\partial z^i}\bar{z}^i\big)dz^i \wedge d\bar{z}^j \ , \tag{5.302}$$

$$K = i\partial\bar{\partial}\mathcal{K} = \frac{i}{2}\mathrm{Im}\big(\frac{\partial^2 \mathcal{F}}{\partial z^i \partial z^j}\big)dz^i \wedge d\bar{z}^j = \frac{i}{2}\mathrm{Im}(\tau_{ij})dz^i \wedge d\bar{z}^j \ , \tag{5.303}$$

where z^i are special coordinates, and $\tau_{ij} = \frac{\partial^2 \mathcal{F}}{\partial z^i \partial z^j}$.

An equivalent way of characterizing rigid special Kmanifolds is via a holomorphic symmetric 3-tensor C. This tensor measures the difference between the symplectic connection ∇ and the Levi-Civita connection D, whose connection coefficients we here denote γ^k_{ij} and $\bar{\gamma}^{\bar{k}}_{\bar{i}\bar{j}}$.
Define

$$\mathcal{P}_{\mathbb{R}} = \nabla - D \ .$$

The nonvanishing components of $\mathcal{P}_{\mathbb{R}}$ are

$$A^k_{ij} - \gamma^k_{ij} \ , \quad A^{\bar{k}}_{ij} \ , \ A^{\bar{k}}_{\bar{i}\bar{j}} - \gamma^{\bar{k}}_{\bar{i}\bar{j}} \ , \ A^k_{\bar{i}\bar{j}} \ , \tag{5.304}$$

this is so because the components A of the connection ∇ are constrained by condition (5.300). Since D and ∇ are real and torsionfree we further have that the lower indices in (5.304) are symmetric, and the reality conditions $\overline{A^k_{ij} - \gamma^k_{ij}} = A^{\bar{k}}_{\bar{i}\bar{j}} - \gamma^{\bar{k}}_{\bar{i}\bar{j}}$, $\overline{A^{\bar{k}}_{ij}} = A^k_{\bar{i}\bar{j}}$. Since both D and ∇ are symplectic we have that for any vector $u \in T_mM$, $(\mathcal{P}_{\mathbb{R}})_u : T_mM \to T_mM$ is a generator of a symlectic transformation,

$$\begin{aligned} u(K(v,w)) &= D_u(K(v,w)) = K(D_u v, w) + K(v, D_u w) \\ u(K(v,w)) &= \nabla_u(K(v,w)) = K(\nabla_u v, w) + K(v, \nabla_u w) \\ 0 &= K((\mathcal{P}_{\mathbb{R}})_u v, w) + K(v, (\mathcal{P}_{\mathbb{R}})_u w) \ . \end{aligned} \tag{5.305}$$

If we set $u = \partial_k$, $v = \partial_i$, $w = \bar{\partial}_{\bar{j}}$, and use that K is a $(1,1)$-form, we obtain

$$A^k_{ij} - \gamma^k_{ij} = 0 \ . \tag{5.306}$$

Then the components of

$$\mathcal{P}_{\mathbb{R}} = \mathcal{P} + \overline{\mathcal{P}}$$

are just $A^{\bar{k}}_{ij}$ and $A^k_{\bar{i}\bar{j}}$. This leads to define the tensor

$$C_{ijk} = -ig_{i\bar{\ell}}A^{\bar{\ell}}_{jk} \ . \tag{5.307}$$

Setting $u = \partial_k$, $v = \partial_i$, $w = \partial_j$ in (5.305) we obtain that C_{ijk} is totally symmetric in its indices. Since $D_j\pi^{(1,0)} = 0$ we easily compute, recalling (5.301), $C_{ijk} = -\langle \nabla_i\xi, \nabla_j\nabla_k\xi\rangle$, hence we obtain the coordinate independent expression for $C = C_{ijk}dz^i \otimes dz^j \otimes dz^k$,

$$C = -\langle \nabla\xi, \nabla\nabla\xi\rangle \ . \tag{5.308}$$

Flatness of $\nabla = D + \mathcal{P}_{\mathbb{R}}$, i.e. $d^2_{\nabla} = 0$, is equivalent to

$$R + d_D\mathcal{P} + d_D\overline{\mathcal{P}} + \mathcal{P} \wedge \overline{\mathcal{P}} + \overline{\mathcal{P}} \wedge \mathcal{P} = 0 \tag{5.309}$$

where $R = d^2_D$ is the Levi-Civita curvature and $d_{\nabla}\mathcal{P}$ is the exterior covariant derivative action on the 1-form $\mathcal{P}$ with values in $T_{\mathbb{C}}M \otimes T^*_{\mathbb{C}}M$ (where $T^*_{\mathbb{C}}M$ is the complexified cotangent bundle). Now in (5.309), the term $R + \mathcal{P} \wedge \overline{\mathcal{P}} + \overline{\mathcal{P}} \wedge \mathcal{P} \in \Omega^{(1,1)}(M, End(T_{\mathbb{C}}M, T_{\mathbb{C}}M))$, i.e., this term maps

$T^{(1,0)}M$ (or $T^{(0,1)}M$) vectors into (1,1)-forms valued in $T^{(1,0)}M$ (or $T^{(0,1)}M$). On the other hand $\mathcal{P} \in \Omega(End(T_{\mathbb{C}}M, \overline{T_{\mathbb{C}}M}))$, in particular it maps $T^{(1,0)}M$ vectors into forms valued in $T^{(0,1)}M$, and annihilates $T^{(0,1)}M$ vectors (hence $\mathcal{P} \wedge \mathcal{P} = 0$). Similar properties hold for the complex conjugate $\overline{\mathcal{P}}$, with $T^{(1,0)}M$ replaced by $T^{(0,1)}M$, and for $d_D\mathcal{P}$ and $d_D\overline{\mathcal{P}}$. It follows that equation (5.309) is equivalent to two independent equations,

$$R + \mathcal{P} \wedge \overline{\mathcal{P}} + \overline{\mathcal{P}} \wedge \mathcal{P} = 0 \tag{5.310}$$

$$d_D\mathcal{P} = 0 \ . \tag{5.311}$$

Since the covariant derivative of the metric vanishes, this last equation is equivalent to $d_DC = 0$. In local coordinates we have

$$dC_{\ell j} - \gamma^k{}_\ell \wedge C_{kj} - \gamma^k{}_j \wedge C_{\ell k} = 0 \ . \tag{5.312}$$

where $C_{ij} = C_{ikj}dz^k$. This equation splits in the condition

$$\bar{\partial} C = 0 \ , \tag{5.313}$$

so that C is holomorphic, and the condition $\partial_D C = 0$, that can be equivalently written

$$D_iC_j = D_jC_i \tag{5.314}$$

where C_i is the matrix $C_i = (C_{ki\ell})_{k,\ell=1,\ldots n}$, i.e., $C_i \in \Omega^0(M, T^{*(1,0)}M \otimes T^{*(1,0)}M)$, so that D_i is the covariant derivative on functions valued in $T^{*(1,0)}M \otimes T^{*(1,0)}M$.

The local coordinates expression of (5.310) is

$$R_{\bar{\imath}j\bar{k}\ell} = -\overline{C}_{\bar{\imath}\bar{k}\bar{s}}g^{\bar{s}p}C_{pj\ell} \ . \tag{5.315}$$

In conclusion a rigid special Kstructure on M implies the existence of a holomorphic symmetric 3-tensor (cubic form) C that satisfies (5.310) and (5.314).

Viceversa if a Kmanifold M admits a symmetric holomorphic 3-tensor C that satisfies (5.310) and (5.314), then M is a special Kmanifold. Indeed the contraction of C with the metric gives $\mathcal{P}$, so that we can define $\nabla = D - \mathcal{P}_{\mathbb{R}}$. The symmetry of C implies that $d_\nabla \pi^{1,0} = 0$ so that ∇ is torsionfree and compatible with the complex structure, $d_\nabla J = 0$. The symmetry of C also implies (5.305) so that ∇ is symplectic. Finally (5.310) and (5.314) imply that ∇ is flat.

In special coordinates the holomorphic 3-tensor C is simply given by $C_{ijk} = \frac{1}{4}\frac{\partial^3\mathcal{F}}{\partial z^i \partial z^j \partial z^k}$.

Local Special Geometry

We have recalled that to a rigid special Kmanifold of dimension n there is canonically associated a holomorphic n dimensional flat symplectic vector bundle. On the other hand, to a projective (or local) special Kmanifold M, of dimension n' there is canonically associated a holomorphic $n = n'+1$ dimensional flat symplectic vector bundle. The increase by one unit of the rank of the vector bundle with respect to the dimension of the manifold is due to the graviton multiplet. The mathematical description involves the $n = n' + 1$ dimensional manifold L, total space of a line bundle over M.

K-Hodge manifolds and their associated principal bundles $\tilde{M} \to M$

Consider a K-Hodge manifold, i.e. a triple (M, L, K), where M is Kwith integral Kform K, so that it defines a class $[K] \in H^2(M, \mathbb{Z})$, and

$$L \xrightarrow{\pi} M$$

is a holomorphic hermitian line bundle with first Chern class equal to $[K]$, and with curvature equal to $-2\pi i K$ (recall that on a hermitian holomorphic vector bundle there is a unique connection compatible with the hermitian holomorphic structure).

Consider the complex manifold $\tilde{M}$, that is L without the zero section of $L \xrightarrow{\pi} M$. The manifold $\tilde{M}$ is a principal bundle over M, with structure group $\mathbb{C}^\times$ (complex numbers minus the zero); the action of $\mathbb{C}^\times$ on $\tilde{M}$ is holomorphic. The hermitian connection canonically associated to $L \to M$ induces a connection on $\tilde{M}$ so that in $T\tilde{M}$ we have the subspaces of horizontal and vertical tangent vectors.

Another property of the manifold $\tilde{M}$ is that it has a canonical hermitian line bundle $\pi^*L \to \tilde{M}$; it is the pullback to $\tilde{M}$ of $L \to M$, so that the fiber on the point $\tilde{m} \in \tilde{M}$ is just the fiber of L on the point $m = \pi(\tilde{m}) \in M$,

$$\begin{array}{ccc} \pi^* L & \longrightarrow & L \\ \downarrow & & \downarrow{\scriptstyle \pi} \\ \tilde{M} & \xrightarrow{\ \pi\ } & M \end{array}$$

Explicitly $\pi^\star L = \{(\tilde{m}, \ell)\,;\ \pi(\ell) = \pi(\tilde{m})\}$. The line bundle $\pi^* L$ is trivial indeed we have the globally defined nonzero holomorphic section

$$\begin{aligned} \Omega\, :\ \tilde{M} \quad &\to \pi^* L \\ \tilde{m} \quad &\mapsto (\tilde{m}, \tilde{m}) \\ (m, \lambda) &\mapsto (m, \lambda, \lambda)\ . \end{aligned} \tag{5.316}$$

In the last line we used a local trivialization of $\tilde{M} \to M$ (and henceforth of $L \to M$) given by a local section s, say $\tilde{m} = \lambda s(m) \sim (m, \lambda)$. This induces a local trivialization $\tilde{s} = \pi^* s$ of the line bundle $\pi^* L \to \tilde{M}$. Explicitly $\tilde{s}$ associates to $\tilde{m}$ the point $s(m)$ of L, so that a generic element $\tilde{\ell} = \sigma \tilde{s}(\tilde{m}) \in \tilde{L}$ is described by the triple (m, λ, σ), and in particular

$$\Omega(\tilde{m}) = \Omega(\lambda s(m)) = \lambda \tilde{s}(\tilde{m}) \sim (m, \lambda, \lambda)\ . \tag{5.317}$$

It can be shown that $\tilde{M}$ is a pseudo-Kmanifold (i.e. a Kmanifold where the metric has pseudo-Riemannian signature). The Kform is

$$\tilde{K} = \frac{i}{2\pi} \bar{\partial} \partial |\Omega|^2\ , \tag{5.318}$$

where $|\Omega|^2$ is the evaluation on Ω of the hermitian structure of $\pi^\star(L)$ (this latter is trivially inherited from the hermitian structure of L). With respect to the corresponding Kmetric, horizontal and vertical vectors are orthogonal, moreover the Kmetric is negative definite along vertical vectors, and positive definite along horizontal vectors, where $\tilde{K}|_{hor} = |\Omega|^2 \pi^* K$.* Thus $(\tilde{M}, \tilde{K})$ has Lorentzian signature.

Concerning the pullback $\pi^* K$ on $\tilde{M}$ of the Kform K on M; while K is in general only closed, $\pi^* K$ is exact,

$$\pi^* K = \frac{i}{2\pi} \bar{\partial} \partial \log |\Omega|^2\ . \tag{5.319}$$

*Hint: in the coordinates (z^i, λ), associated to the local trivialization $\tilde{m} = \lambda s(m) \sim (m, \lambda)$ induced by a section s of L, we have $|\Omega|^2 = \lambda \bar{\lambda} |s|^2$. Moreover horizontal vectors read $u = u^i \partial_i - u^i a_i \lambda \frac{\partial}{\partial \lambda}$ where the local connection 1-form on M is $a = a_i dz^i = |s|^{-2} \partial |s|^2$. The pseudo-Kform reads $-2\pi i \tilde{K} = \lambda \bar{\lambda} \partial_i \partial_{\bar{\jmath}} |s|^2 dz^i \wedge d\bar{z}^{\bar{\jmath}} + |s|^2 d\lambda \wedge d\bar{\lambda} + \lambda \partial_i |s|^2 dz^i \wedge d\bar{\lambda} + \bar{\lambda} \partial_{\bar{\jmath}} |s|^2 d\lambda \wedge d\bar{z}^{\bar{\jmath}}$.

This last formula easily follows by pulling back the usual local curvature formula for the hermitian connection $K = \frac{i}{2\pi}\bar{\partial}\partial \log |s|^2$ and by observing that $\pi^*\log|s|^2 = \log|\tilde{s}|^2 = \log|\Omega|^2 - \log\lambda - \log\bar{\lambda}$.

In conclusion, one can canonically associate to a K-Hodge manifold (M, L, K) a pseudo-Kmanifold $(\tilde{M}, \tilde{K})$ that carries a free and holomorphic $\mathbb{C}^\times$ action, and a line bundle $\pi^*L \to \tilde{M}$ that has a canonical global holomorphic section Ω.

The bundle $\tilde{L}$ can be naturally identified as the holomorphic subbundle of $T\tilde{M}$ given by the vertical vectors of $\tilde{M}$ with respect to the holomorphic $\mathbb{C}^\times$ action. The global holomorphic section Ω corresponds to the vertical vector field that gives the infinitesimal $\mathbb{C}^\times$ action. Under this identification we have

$$\tilde{K}(\Omega, \Omega) = -\frac{i}{2\pi}|\Omega|^2 \ . \tag{5.320}$$

This equation shows that under the identification $T\tilde{M}|_{vert} \simeq L$ the corresponding hermitian structures are mapped one into minus the other.

Special Kmanifolds
Following [49], (M, L, K) is special Kif $(\tilde{M}, \tilde{K})$ is rigid special Kand if Ω is compatible with the symplectic connection $\tilde{\nabla}$.

A (projective or local) **special Kmanifold** is a K-Hodge manifold (M, L, K) such that the associated pseudo-Kmanifold $(\tilde{M}, \tilde{K})$ has a rigid special pseudo-Kstructure $\tilde{\nabla}$ which satisfies

$$\tilde{\nabla}\Omega = \pi^{(1,0)} \ . \tag{5.321}$$

Notice that (5.321) is equivalent to the condition $\tilde{\nabla}_u\Omega = u$ for any $u \in T^{(1,0)}\tilde{M}$. As shown in [50], since $\tilde{\nabla}$ is torsionfree and flat, then condition (5.321) implies the $\mathbb{C}^\times$ invariance of $\tilde{\nabla}$, i.e. $dR_b(\tilde{\nabla}_u v) = \tilde{\nabla}_{dR_b u} dR_b v$ where R_b denotes the action of $b \in \mathbb{C}^\times$. Notice also that equation (5.321) is the global version of eq. (5.301).

For ease of notation in the following we denote the flat torsionfree symplectic connection $\tilde{\nabla}$ on $\tilde{M}$ simply by ∇.

We now construct a flat symplectic $2n = 2n' + 2$ dimensional bundle $\boldsymbol{\mathcal{H}}$ on M that is frequently used in the literature in order to characterize projective special Kmanifolds. We introduce a new $\mathbb{C}^\times$ action on $T\tilde{M}$. On $\tilde{M}$ it is the usual one $R_b\tilde{m} = \tilde{m}b = b\tilde{m}$, where $b \in \mathbb{C}^\times$, while on vectors we have

$$v_{\tilde{m}} \mapsto b^{-1} dR_b \, v_{\tilde{m}} \ . \tag{5.322}$$

>From now on by $\mathbb{C}^\times$ action we understand the new above defined one. Thus for example since $b^{-1}dR_b\Omega_{\tilde{m}} = b^{-1}\Omega_{\tilde{m}b}$, then Ω is not invariant under (5.322). On the other hand the local section (vertical vector field) $\tilde{s}$, obtained from a local section s of L, satisfies $b^{-1}dR_b\tilde{s}_{\tilde{m}} = \tilde{s}_{b\tilde{m}}$ (or $b^{-1}R_{b*}\tilde{s} = \tilde{s}$) and is therefore $\mathbb{C}^\times$ invariant. A $\mathbb{C}^\times$ invariant frame associated with local coordinates z^i of M and with the local section s of L is $(\lambda^{-1}\frac{\partial}{\partial z^i}, \frac{\partial}{\partial \lambda})$; it is given by the coordinates $(X^i, X^0) = (\lambda z^i, \lambda)$, they are $\mathbb{C}^\times$ invariant $(b^{-1}{R_b}^*X = X)$ and therefore are homogeneous (projective) coordinates of M.

We define the $2n = 2n' + 2$ dimensional real vector bundle on M $(\dim_{\mathbb{R}} M = 2n')$,

$$\boldsymbol{\mathcal{H}} \to M \tag{5.323}$$

by identifying its local sections with the $\mathbb{C}^\times$ invariant sections of $T\tilde{M}$. In other words $\boldsymbol{\mathcal{H}}$ is the quotient of $T\tilde{M}$ via the $\mathbb{C}^\times$ action (5.322). A point $(m, h) \in \boldsymbol{\mathcal{H}}$ is the equivalence class

$[(\tilde{m}, v_{\tilde{m}})]$ where $(\tilde{m}, v_{\tilde{m}}) \sim (\tilde{m}', u_{\tilde{m}'})$ if $m' = mb$ and $b^{-1} dR_b v_{\tilde{m}} = u_{\tilde{m}'}$. Under this quotient $\pi^\star L \subset T\tilde{M}$ becomes L, while the subbundle $T\tilde{M}|_{hor}$ of horizontal vectors becomes $L \otimes TM$.* Therefore we have two natural inclusions

$$L \subset \mathcal{H} \quad \text{and} \quad L \otimes TM \subset \mathcal{H} \ . \tag{5.324}$$

Since the $\mathbb{C}^\times$ action is holomorphic, then $\mathcal{H}$ is a holomorphic vector bundle on M of rank $n' + 1$. Since $\tilde{K}$ is a $\mathbb{C}^\times$ invariant 2-form the symplectic structure of $T\tilde{M}$ goes to the quotient $\mathcal{H}$: indeed $\tilde{K}(u, v)$ is a homogeneous function on M if u and v are $\mathbb{C}^\times$ invariant vector fields of $T\tilde{M}$. Similarly also the flat symplectic connection ∇ induces a flat symplectic connection on $\mathcal{H}$ (see for example [50]). The inclusion $L \subset \mathcal{H}$ implies that

$$L^{-1} \otimes \mathcal{H} \to M \tag{5.325}$$

has a nonvanishing global holomorphic section.

In the following we work in $T\tilde{M}$, but we choose $\mathbb{C}^\times$ invariant tensors and therefore our results immediately apply to the bundle $\mathcal{H}$. Let's consider a $\mathbb{C}^\times$ invariant flat local symplectic framing of $T\tilde{M}$, that we denote by $\{\mathsf{e}_\xi\} = \{\mathsf{e}_\Lambda, \mathsf{f}^\Lambda\}$, $\xi = 1, \ldots 2n$, $\Lambda = 1, \ldots n$. The framing is flat because $\nabla \mathsf{e}_\Lambda = 0, \nabla \mathsf{f}^\Lambda = 0$, and it is symplectic because in this basis the symplectic matrix is in canonical form: the components $\tilde{K}(\mathsf{e}_\Lambda, \mathsf{e}_\Sigma)$, $\tilde{K}(\mathsf{e}_\Lambda, \mathsf{f}^\Sigma)$, $\tilde{K}(\mathsf{f}^\Lambda, \mathsf{e}_\Sigma)$, $\tilde{K}(\mathsf{f}^\Lambda, \mathsf{f}^\Sigma)$ read

$$\begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \tag{5.326}$$

With respect to the $\{\mathsf{e}_\Lambda, \mathsf{f}^\Lambda\}$ frame, the global section Ω has local components $\Omega = \Omega^\xi \mathsf{e}_\xi = X^\Lambda \mathsf{e}_\Lambda + F_\Lambda \mathsf{f}^\Lambda$. We also denote by Ω this column vector of coefficients,

$$\Omega = (\Omega^\xi) = \begin{pmatrix} X^\Lambda \\ F_\Lambda \end{pmatrix} \ . \tag{5.327}$$

The local functions X^Λ, F_Λ on $\tilde{M}$ are holomorphic, indeed (5.321) implies that $\nabla\Omega$ is a $(1,0)$-form valued in $T\tilde{M}$, since $\nabla(\Omega^\xi \mathsf{e}_\xi) = d\Omega^\xi \, \mathsf{e}_\xi = \partial\Omega^\xi \, \mathsf{e}_\xi + \bar{\partial}\Omega^\xi \, \mathsf{e}_\xi$, we obtain $\bar{\partial}\Omega^\xi = 0$. In conclusion (X^Λ, F_Λ) are local components of the global symplectic section Ω of the tangent bundle $T\tilde{M}$.

Each entry X^Λ, F_Λ is also a local holomorphic section of the line bundle $L^{-1} \to M$. Indeed from the transformation properties of Ω under the $\mathbb{C}^\times$ action $\tilde{m} \mapsto R_{e^{-f(m)}}(\tilde{m}) = e^{-f(m)}\tilde{m}$ (or under a change of local trivialization $s'(m) = e^{f(m)} s(m)$) we have

$$\begin{pmatrix} X^\Lambda \\ F_\Lambda \end{pmatrix}' = e^{-f(m)} \begin{pmatrix} X^\Lambda \\ F_\Lambda \end{pmatrix} \ , \tag{5.328}$$

therefore for each invertible Ω^ξ we have that ${\Omega^\xi}^{-1}(s)s$ is a section of $L \to M$ or equivalently each X^Λ and each F_Λ are the coefficients of sections of $L^{-1} \to M$.

In conclusion (X^Λ, F_Λ) are local components of the global symplectic section Ω of the tangent bundle $T\tilde{M}$. Each entry is also a local holomorphic section of the line bundle $L^{-1} \to M$. Under change of local trivialization of $T\tilde{M}$ we have

$$\begin{pmatrix} X^\Lambda \\ F_\Lambda \end{pmatrix}' = S \begin{pmatrix} X^\Lambda \\ F_\Lambda \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} X^\Lambda \\ F_\Lambda \end{pmatrix} \ , \tag{5.329}$$

*Hint: denote by $\widehat{v_m}|_{\tilde{m}}$ the horizontal lift in $T_{\tilde{m}}\tilde{M}$ of the vector $v_m \in T_m M$. Then the map $L \otimes TM \to (T\tilde{M}|_{hor})/_{\mathbb{C}^\times \text{action}}$ defined by $(\ell_m \otimes v_m) \mapsto [(\ell_m, \widehat{v_m}|_{\ell_m})]$ if $\ell_m \neq 0$, and by $0 \mapsto 0$ is well defined, linear and injective.

where $S = \binom{A\,C}{B\,D}$ is a constant symplectic matrix. We can also consider a change of coordinates on M, say $z \to z'$. Provided we keep fixed the frame of $T\tilde{M}$ and the trivalization of L we then have that X^Λ and F_Λ behave like local functions on M, $X^\Lambda(z) = X'^\Lambda(z')$, $F_\Lambda(z) = F'_\Lambda(z')$ (here $X^A(z) = X^A(s(z))$ etc.).

It can be shown [50] that from the set of $2n$ elements $\{X^\Lambda, F_\Lambda\}$ one can always choose a subset of n elements that form a local coordinate system on $\tilde{M}$. Contrary to the Kcase (where the metric is Riemanninan) in this pseudo-Kcase in general neither $\{X^\Lambda\}$ nor $\{F_\Lambda\}$ are coordinates systems on $\tilde{M}$. The frame $\{\mathsf{e}_\Lambda, \mathsf{f}^\Lambda\}$ is determined up to a symplectic transformation, if using this freedom we have that the $\{X^\Lambda\}$ are coordinates functions then the $\{X^\Lambda\}$ are named special coordinates. The sections F_Λ can then be seen as functions of the X^Λ and are obtained via a prepotential $\mathcal{F}$,

$$F_\Lambda = \frac{\partial \mathcal{F}}{\partial X^\Lambda} \ . \tag{5.330}$$

Recalling (5.319) and (5.320) we have

$$\pi^* K = \frac{i}{2\pi} \bar{\partial}\partial \log i\langle \Omega, \overline{\Omega} \rangle \tag{5.331}$$

and for the corresponding "K" potential $\mathcal{K}$ we have*

$$\mathcal{K} = -\log i\langle \Omega, \overline{\Omega} \rangle \ ; \tag{5.332}$$

in these formulae we used the standard notation

$$\langle \Omega, \overline{\Omega} \rangle = \tilde{K}(\Omega, \overline{\Omega}) \ .$$

Using the components (X^Λ, F_Λ) expression (5.332) reads

$$\mathcal{K} = -\log\big[i(X,F)\begin{pmatrix} 0 & -1\!\!1 \\ 1\!\!1 & 0 \end{pmatrix}\begin{pmatrix} \bar{X} \\ \bar{F} \end{pmatrix}\big] = -\log[i(F_\Lambda \bar{X}^\Lambda - X^\Lambda \bar{F}_\Lambda)] \ . \tag{5.333}$$

By considering local sections of the bundle $\tilde{M} \to M$, we can then pull back the potential $\mathcal{K}$ to local Kpotentials on M.

Under the action of $e^{-f(m)} \in \mathbb{C}^\times$ on $\tilde{M}$ (or equivalently under change of trivialization of $\tilde{M} \to M$) we have

$$\mathcal{K}' = \mathcal{K} + f + \bar{f} \tag{5.334}$$

thus showing that $e^{-\mathcal{K}}$ defines a global nonvanishing section of the bundle $L \otimes \overline{L} \to M$, in particular this bundle is trivial. Explicitly this global section is $e^{\mathcal{K}(s)}[s, \bar{s}]$ where s is any local section of $\tilde{M} \to M$ and $[s, \bar{s}] = \{(s\lambda, \lambda^{-1}\bar{s}), \lambda \in \mathbb{C}^\times\}$ is the corresponding local section of $L \otimes \bar{L}$.

Symplectic Sections and Matrices from local coordinates frames on M

Let's examine few more properties of special Kmanifolds and introduce those symplectic vectors that we have seen characterizing the geometry of the supergravity scalar fields. Consider a vector $u \in T_m^{(1,0)}M$, this can be lifted to a horizontal vector $\hat{u} \in T_{\tilde{m}}^{(1,0)}\tilde{M}$. Because of (5.321) the covariant derivative $\nabla_{\hat{u}}\Omega$ is again a vector in $T_{\tilde{m}}^{(1,0)}\tilde{M}$, then

$$\langle \Omega, \nabla_{\hat{u}}\Omega \rangle = 0 \quad , \quad \langle \bar{\Omega}, \nabla_{\hat{u}}\Omega \rangle = 0 \quad ; \tag{5.335}$$

the first relation holds because $\tilde{K} = \langle \ , \ \rangle$ is a $(1,1)$-form, the second relation holds because horizontal and vertical vectors are orthogonal under $\tilde{K}$ (recall paragraph after (5.318)).

*As usual when K is integral $K = \frac{i}{2\pi} g_{i\bar{j}} dz^i \wedge d\bar{z}^{\bar{j}} = \frac{i}{2\pi}\partial_i \partial_{\bar{j}} \mathcal{K} dz^i \wedge d\bar{z}^{\bar{j}} = \frac{i}{2\pi}\partial\bar{\partial}\mathcal{K}$.

Subordinate to a holomorphic coordinate system $\{z^i\}$ of M, and a local section s of $L \to M$ we have the local coordinates (z^i, λ) on $\tilde{M}$. The corresponding vector fields are $(\partial_i, \frac{\partial}{\partial\lambda})$. A more natural frame on $\tilde{M}$ is given by considering the vertical vector field associated to the action of $\mathbb{C}^\times$ on $\tilde{M}$,

$$\hat{\partial}_0 \equiv \Omega = \lambda \frac{\partial}{\partial\lambda} \ , \tag{5.336}$$

and the horizontal lift $\hat{\partial}_i$ of the vector fields ∂_i on M

$$\hat{\partial}_i = \partial_i - |s|^{-2}\partial_i|s|^2\lambda\frac{\partial}{\partial\lambda} = \partial_i + \partial_i\mathcal{K}\lambda\frac{\partial}{\partial\lambda} .\partial_i + \partial_i\mathcal{K}\hat{\partial}_0 \tag{5.337}$$

In (5.337), $|s|^2 = h(s,s)$ is the hermitian form of $L \to M$. All these vector fields have degree 1 and are independent from the section s of $L \to M$.

We define

$$\nabla_i = \nabla_{\hat{\partial}_i} \ . \tag{5.338}$$

The new sections $\nabla_i\Omega$ are exactly the horizontal vector fields $\hat{\partial}_i$, indeed from (5.321) we obtain

$$\nabla_i\Omega = \hat{\partial}_i \quad , \quad \nabla_0\Omega = \hat{\partial}_0 = \Omega \quad . \tag{5.339}$$

Similarly

$$\bar{\nabla}_{\bar{\imath}}\bar{\Omega} = 0 \quad , \quad \bar{\nabla}_{\bar{0}}\Omega = 0 \quad . \tag{5.340}$$

Recalling (5.335) we obtain

$$\langle \Omega, \nabla_i\Omega \rangle = 0 \tag{5.341}$$

$$\langle \nabla_i\Omega, \nabla_j\Omega \rangle = 0 \tag{5.342}$$

$$\langle \bar{\Omega}, \nabla_i\Omega \rangle = 0 \ . \tag{5.343}$$

Notice also that $\langle \Omega, \bar{\Omega} \rangle$ is invariant under horizontal vector fields,

$$\hat{\partial}_i\langle \Omega, \bar{\Omega} \rangle = \nabla_i\langle \Omega, \bar{\Omega} \rangle = \langle \nabla_i\Omega, \bar{\Omega} \rangle + \langle \Omega, \nabla_i\bar{\Omega} \rangle = 0 \tag{5.344}$$

where in the last passage we used (5.335) and (5.340). Similarly $\bar{\nabla}_{\bar{\imath}}\langle \Omega, \bar{\Omega} \rangle = 0$.

The metric associated to the Kform (5.318) on $\tilde{M}$ is block diagonal in the $\hat{\partial}_0, \hat{\partial}_i$ basis, (see paragraph following (5.318)),

$$\begin{pmatrix} \tilde{g}_{0\bar{0}} & 0 \\ 0 & \tilde{g}_{i\bar{\jmath}} \end{pmatrix} = \begin{pmatrix} -\lambda\bar{\lambda}|s|^2 & 0 \\ 0 & \lambda\bar{\lambda}|s|^2 g_{i\bar{\jmath}}\circ\pi \end{pmatrix} = \begin{pmatrix} -|\Omega|^2 & 0 \\ 0 & |\Omega|^2\, g_{i\bar{\jmath}}\circ\pi \end{pmatrix} \ . \tag{5.345}$$

Because of (5.344) the associated Levi-Civita connection coefficients of $\tilde{M}$ in the $\hat{\partial}_i$ basis of horizontal vectors coincide with those of M in the $\frac{\partial}{\partial z^i}$ basis,

$$\tilde{\Gamma}^\ell_{ij} = \tilde{g}^{\bar{k}\ell}\hat{\partial}_i\tilde{g}_{j\bar{k}} = g^{\bar{k}\ell}\partial_i g_{j\bar{k}} = \Gamma^\ell_{ij} \ . \tag{5.346}$$

In terms of the symplectic frame $\{\mathsf{e}_\xi\} = \{\mathsf{e}_\Lambda, \mathsf{f}^\Lambda\}$, that is flat, we have $\nabla\Omega = \nabla(\Omega^\xi\mathsf{e}_\xi) = d(\Omega^\xi)\mathsf{e}_\xi$, and $\nabla_i\Omega = \hat{\partial}_i(\Omega^\xi)\mathsf{e}_\xi = \partial_i\Omega^\xi + \partial_i\mathcal{K}\Omega^\xi$, i.e.,

$$\nabla_i\begin{pmatrix} X^\Lambda \\ F_\Lambda \end{pmatrix} = \partial_i\begin{pmatrix} X^\Lambda \\ F_\Lambda \end{pmatrix} + \partial_i\mathcal{K}\begin{pmatrix} X^\Lambda \\ F_\Lambda \end{pmatrix} \ . \tag{5.347}$$

Recalling the interpretation of X^Λ or F_Λ as coefficients of local sections of $L^{-1} \to M$, we read in equation (5.347) the covariant derivative of $L^{-1} \to M$.

It is also convenient to normalize Ω and thus consider the (non holomorphic) nonvanishing global vector field on $\tilde{M}$ given by

$$\mathcal{V} = e^{\mathcal{K}/2}\Omega \ . \tag{5.348}$$

>From (5.344) the covariant derivatives of $\mathcal{V}$ are

$$\nabla_i \mathcal{V} = e^{\mathcal{K}/2}\nabla_i\Omega \quad , \quad \bar{\nabla}_{\bar{\imath}} \mathcal{V} = e^{\mathcal{K}/2}\bar{\nabla}_{\bar{\imath}}\Omega = 0 \ ,$$
$$\bar{\nabla}_{\bar{\imath}} \bar{\mathcal{V}} = e^{\mathcal{K}/2}\bar{\nabla}_{\bar{\imath}}\bar{\Omega} \quad , \quad \nabla_i \bar{\mathcal{V}} = e^{\mathcal{K}/2}\nabla_i\bar{\Omega} = 0 \ .$$

Explicitly we have*

$$\nabla_i \mathcal{V} = (\partial_i V^\xi + \frac{1}{2}\partial_i \mathcal{K} V^\xi)\, \mathsf{e}_\xi \quad , \quad \bar{\nabla}_{\bar{\imath}} \mathcal{V} = (\bar{\partial}_{\bar{\imath}} V^\xi - \frac{1}{2}\bar{\partial}_{\bar{\imath}} \mathcal{K}\, V^\xi)\, \mathsf{e}_\xi = 0 \tag{5.349}$$

$$\bar{\nabla}_{\bar{\imath}} \bar{\mathcal{V}} = (\bar{\partial}_{\bar{\imath}} \bar{V}^\xi + \frac{1}{2}\bar{\partial}_{\bar{\imath}} \mathcal{K}\, \bar{V}^\xi)\, \mathsf{e}_\xi \quad , \quad \nabla_i \bar{\mathcal{V}} = (\partial_i \bar{V}^\xi - \frac{1}{2}\partial_i \mathcal{K}\, \bar{V}^\xi)\, \mathsf{e}_\xi = 0 \ . \tag{5.350}$$

Each coefficient V^ξ of $\mathcal{V}$ with respect to the $\mathbb{C}^\times$ invariant basis e_ξ is also a coefficient of a local section of the bundle $L^{-1/2} \otimes \bar{L}^{1/2} \to M$. This bundle has connection $\frac{1}{2}\partial_i \mathcal{K} - \frac{1}{2}\bar{\partial}_{\bar{\imath}} \mathcal{K}$. Equation (5.349) can be interpreted as the covariant derivative of these line bundle local sections.

From (5.332), and (5.341)-(5.343) we have

$$\langle \mathcal{V}, \bar{\mathcal{V}} \rangle = -i \quad , \tag{5.351}$$
$$\langle \mathcal{V}, \nabla_i \mathcal{V} \rangle = 0 \quad , \tag{5.352}$$
$$\langle \nabla_i \mathcal{V}, \nabla_j \mathcal{V} \rangle = 0 \quad , \tag{5.353}$$
$$\langle \mathcal{V}, \bar{\nabla}_{\bar{\imath}} \bar{\mathcal{V}} \rangle = 0 \quad . \tag{5.354}$$

>From (5.345), or also from $[\nabla_j, \bar{\nabla}_{\bar{\imath}}] = -\partial_j \bar{\partial}_{\bar{\imath}} \mathcal{K} = -g_{j\bar{\imath}}$ and $\langle \bar{\nabla}_{\bar{\imath}} \nabla_j \mathcal{V}, \bar{\mathcal{V}} \rangle + \langle \nabla_j \mathcal{V}, \bar{\nabla}_{\bar{\imath}} \bar{\mathcal{V}} \rangle = 0$, we have

$$\langle \nabla_j \mathcal{V}, \bar{\nabla}_{\bar{\imath}} \bar{\mathcal{V}} \rangle = i g_{j\bar{\imath}} \ , \tag{5.355}$$

(where $g_{j\bar{\imath}} = \partial_j \bar{\partial}_{\bar{\imath}} \mathcal{K} = -2\pi i\, \pi^* K_{j\bar{\imath}}$ is actually $g_{j\bar{\imath}} \circ \pi$, the pull back via π of the positive definite metric on M). If we consider an orthonormal frame $\{e_I\}$, $(I = 1, \ldots n')$ on M,

$$e_I = e_I^j \partial_j \quad , \quad \partial_j = e_j^I e_I \ , \ g_{j\bar{\imath}} = e_j^I\, \bar{e}_{\bar{\imath}}^{\bar{J}}\, \delta_{I\bar{J}} \ , \tag{5.356}$$

we lift this frame to a frame of horizontal vectors of $T^{(1,0)}\tilde{M}$, and if we set

$$\mathcal{V}_M = (\mathcal{V}, \bar{\nabla}_{\bar{I}} \bar{\mathcal{V}}) \quad , \quad M = 0, 1, \ldots n' \ , \tag{5.357}$$

(where $\bar{\nabla}_{\bar{I}} = \bar{e}_{\bar{I}}^{\bar{\imath}}\, \bar{\nabla}_{\bar{\imath}}$), then relations (5.352), (5.353), (5.351), (5.355) read

$$\langle \mathcal{V}_M, \mathcal{V}_N \rangle = 0 \quad , \quad \langle \bar{\mathcal{V}}_M, \mathcal{V}_N \rangle = i\delta_{MN} \ . \tag{5.358}$$

The index M mixes holomorphic and antiholomorphic indices in order to compensate for the Lorentian signature of the metric $\begin{pmatrix} -1 & 0 \\ 0 & g_{j\bar{\imath}} \end{pmatrix}$ in (5.351), (5.355).

*we find also instructive to obtain the covariant derivative of the section $\mathcal{V}$ via this straighforward calculation that uses $\lambda \frac{\partial}{\partial \lambda} \mathcal{K} = -1$,

$$\nabla_i \mathcal{V} = \nabla_i (e^{\mathcal{K}/2} \Omega^\xi \mathsf{e}_\xi) = \hat{\partial}_i (e^{\mathcal{K}/2} \Omega^\xi) \mathsf{e}_\xi = \partial_i (e^{\mathcal{K}/2} \Omega^\xi) \mathsf{e}_\xi + \partial_i \mathcal{K} \lambda \frac{\partial}{\partial \lambda} (e^{\mathcal{K}/2} \Omega^\xi) \mathsf{e}_\xi = (\partial_i V^\xi + \frac{1}{2}\partial_i \mathcal{K} V^\xi)\, \mathsf{e}_\xi \ .$$

Explicitly the column vectors of the components of the sections $\mathcal{V}_M = V^\xi{}_M \mathsf{e}_\xi$ are

$$(V^\xi) = \begin{pmatrix} L^\Lambda \\ M_\Lambda \end{pmatrix} = e^{\mathcal{K}/2} \begin{pmatrix} X^\Lambda \\ F_\Lambda \end{pmatrix} \quad , \quad (\bar{\nabla}_{\bar{I}} \bar{V}^\xi) = \begin{pmatrix} \bar{\nabla}_{\bar{I}} L^\Lambda \\ \bar{\nabla}_{\bar{I}} M^\Lambda \end{pmatrix} \quad , \tag{5.359}$$

and they can be organized in a $2n \times n$ matrix

$$(V^\xi{}_M) = (V, \bar{\nabla}_{\bar{I}} \bar{V}^\xi) = \begin{pmatrix} L^\Lambda & \bar{\nabla}_{\bar{I}} L^\Lambda \\ M_\Lambda & \bar{\nabla}_{\bar{I}} M^\Lambda \end{pmatrix} = \begin{pmatrix} f^\Lambda{}_M \\ h_{\Lambda M} \end{pmatrix} = \begin{pmatrix} f \\ h \end{pmatrix} \ . \tag{5.360}$$

In the last passage we have denoted by f (respectively h) the $n \times n$ matrix of entries $f^\Lambda{}_M$ (respectively $h_{\Lambda M}$).

The $N = 2$ special geometry relations (5.358) are equivalent to

$$(f^\dagger, h^\dagger) \begin{pmatrix} 0 & -1\!\!1 \\ 1\!\!1 & 0 \end{pmatrix} \begin{pmatrix} f \\ h \end{pmatrix} = i 1\!\!1 \qquad i.e. \quad -f^\dagger h + h^\dagger f = i 1\!\!1 \tag{5.361}$$

and

$$(f^t, h^t) \begin{pmatrix} 0 & -1\!\!1 \\ 1\!\!1 & 0 \end{pmatrix} \begin{pmatrix} f \\ h \end{pmatrix} = 0 \qquad i.e. \quad -f^t h + h^t f = 0 \tag{5.362}$$

These two relations are equivalent to require the real matrix

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \sqrt{2} \begin{pmatrix} \mathrm{Re} f & -\mathrm{Im} f \\ \mathrm{Re} h & -\mathrm{Im} h \end{pmatrix} \tag{5.363}$$

to be symplectic. Vice versa any symplectic matrix $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ leads to relations (5.361), (5.362) by defining $\left(\begin{smallmatrix} f \\ h \end{smallmatrix}\right) = \frac{1}{\sqrt{2}} \left(\begin{smallmatrix} A - iB \\ C - iD \end{smallmatrix}\right)$. The matrix

$$V = \begin{pmatrix} f & \bar{f} \\ h & \bar{h} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \mathcal{A} \ , \tag{5.364}$$

where $\mathcal{A} = \frac{1}{\sqrt{2}} \left(\begin{smallmatrix} 1\!\!1 & 1\!\!1 \\ -i 1\!\!1 & i 1\!\!1 \end{smallmatrix}\right)$, rotates the flat real symplectic frame $\{\mathsf{e}_\xi\} = \{\mathsf{e}^\Lambda, \mathsf{f}_\Lambda\}$ in the frame $\{\mathcal{V}_M, \bar{\mathcal{V}}_{\bar{M}}\}$ that up to a rotation by $\mathcal{A}^{-1} = \mathcal{A}^\dagger$ is also real and symplectic (but not flat). This $\{\mathcal{V}_M, \bar{\mathcal{V}}_{\bar{M}}\}$ frame comes from a local coordinate frame on M, indeed $\bar{\mathcal{V}}_{\bar{M}} = (e^{\mathcal{K}/2} \bar{\Omega}, e^{\mathcal{K}/2} e^j_I \hat{\partial}_j)$. The symplectic connection 1-form in this frame is simply $\Gamma = V^{-1} dV$, indeed $\nabla \mathsf{e}_\xi = 0$ is equivalent to

$$dV = V \Gamma \ . \tag{5.365}$$

We can write $\Gamma = \begin{pmatrix} \omega & \bar{\mathcal{P}} \\ \mathcal{P} & \bar{\omega} \end{pmatrix}$, and see this equation as a condition on the Levi-Civita connection ω and the tensor $\mathcal{P}$ of $\tilde{M}$. The block decomposition $\left(\begin{smallmatrix} \omega & \bar{\mathcal{P}} \\ \mathcal{P} & \bar{\omega} \end{smallmatrix}\right)$ follows by recalling that $\tilde{M}$ is in particular a rigid special Kmanifold. The difference $\mathcal{P}_{\mathbb{R}} = \nabla - D$ between the flat symplectic connection and the Levi-Civita connection is given by the holomorphic symmetric three form C (c.f. (5.308))

$$C = -\langle \nabla \Omega, \nabla \nabla \Omega \rangle \ . \tag{5.366}$$

The properties of C previously discussed in the rigid case apply also to this projective special geometry case.

5.5.2 The $N = 2$ theory

>From the previous section we see that the $N = 2$ supergravity theories and the higher N theories have a similar flat symplectic structure. The formalism is the same, indeed since the antisymmetric of the $U(2)$ authomorphism group of the $N = 2$ supersymmetry algebra is a singlet we have

$$f^{\Lambda}{}_{AB} = f^{\Lambda}{}_{0}\epsilon_{AB} \quad , h_{\Lambda AB} = h_{\Lambda 0}\epsilon_{AB} \tag{5.367}$$

where $f^{\Lambda}{}_{0}, h_{\Lambda 0}$ are the components of the global section $\mathcal{V}$, therefore from (5.360) we have as in (5.172),

$$\begin{aligned} f &= (f^{\Lambda}{}_{M}) = (f^{\Lambda}{}_{AB}, \bar{f}^{\Lambda}{}_{\bar{I}}) , \\ h &= (h_{\Lambda M}) = (h_{\Lambda AB}, \bar{h}_{\Lambda \bar{I}}) , \end{aligned} \tag{5.368}$$

as it should be, the sections $\binom{\bar{f}^{\Lambda}{}_{\bar{I}}}{\bar{h}_{\Lambda\bar{I}}}$ have Kweight opposite to the $\binom{f^{\Lambda}{}_{AB}}{h_{\Lambda I}}$ sections.

The difference between the $N = 2$ cases and the $N > 2$ cases is that the scalar manifold M of the $N = 2$ case is not in general a coset manifold. The flat symplectic bundle is therefore not in general a trivial bundle. The gauge kinetic term $\mathcal{N} = \langle_{\mathcal{M}}\{^{-\infty\mathcal{M}}$ depends on the choice of the flat symplectic frame $\{\mathsf{e}_{\xi}\} = \{\mathsf{e}_{\Lambda}, \mathsf{f}^{\Lambda}\}$. This latter can be defined only locally on $\tilde{M}$ (and therefore on M). In another region we have a different frame $\{\mathsf{e}'_{\xi}\} = \{\mathsf{e}'_{\Lambda}, \mathsf{f}'^{\Lambda}\}$ and therefore a different gauge kinetic term $\mathcal{N}'$. In the common overlapping region the two formulations should give the same theory, this is indeed the case because the corresponding equations of motion are related by a duality rotation. As a consequence the notion of electric or magnetic charge depends on the flat frame chosen. In this sense the notion of electric and magnetic charge is not a fundamental one. The symplectic group is a gauge group (where just constant gauge transformations are allowed) and only gauge invariant quantities are physical.

A related aspect of the comparison between the $N = 2$ and the $N > 2$ theories is that the special Kstructure determines the presence of a new geometric quantity, the holomorphic cubic form C, which physically corresponds to the anomalous magnetic moments of the $N = 2$ theory. When the special Kmanifold M is itself a coset manifold [78], then the anomalous magnetic moments C_{ijk} are expressible in terms of the vielbein of G/H, this is for example the case of the $N = 2$ theories with scalar manifold $G/H = \frac{SU(1,1)}{U(1)} \times \frac{O(6,2)}{O(6)\times O(2)}$ and $G/H = \frac{SO^{\star}(12)}{U(6)}$ [78].

To complete the analogy between the $N = 2$ theory with n' vector multiplets and the higher N theories in $D = 4$, we also give the supersymmetry transformation laws, the central and matter charges, the differential relations among them and the formula for the potential $\mathcal{V}_{BH}$.

The supercovariant electric field strength $\hat{F}^{\Lambda}$ is

$$\hat{F}^{\Lambda} = F^{\Lambda} + f^{\Lambda}\bar{\psi}^{A}\psi^{B}\epsilon_{AB} - \mathrm{i}\bar{f}^{\Lambda}_{\bar{\imath}}\bar{\lambda}^{\bar{\imath}}_{A}\gamma_{a}\psi_{B}\epsilon^{AB}V^{a} + h.c. \tag{5.369}$$

The transformation laws for the chiral gravitino ψ_A and gaugino λ^{iA} fields are:

$$\delta\psi_{A\mu} = \nabla_{\mu}\,\epsilon_{A} + \epsilon_{AB}T_{\mu\nu}\gamma^{\nu}\epsilon^{B} + \cdots , \tag{5.370}$$

$$\delta\lambda^{iA} = \mathrm{i}\partial_{\mu}z^{i}\gamma^{\mu}\epsilon^{A} + \frac{\mathrm{i}}{2}\bar{T}_{\bar{\jmath}\mu\nu}\gamma^{\mu\nu}g^{i\bar{\jmath}}\epsilon^{AB}\epsilon_{B} + \cdots , \tag{5.371}$$

where:

$$T = h_{\Lambda}F^{\Lambda} - f^{\Lambda}G_{\Lambda} , \tag{5.372}$$

$$\bar{T}_{\bar{\imath}} = \bar{T}_{\bar{I}}\bar{e}^{\bar{I}}{}_{\bar{\imath}} , \text{ with } \bar{T}_{\bar{I}} = \bar{h}_{\Lambda\bar{I}}F^{\Lambda} - \bar{f}^{\Lambda}_{\bar{I}}G_{\Lambda} , \tag{5.373}$$

are respectively the graviphoton and the matter vectors. In (5.370), (5.371) the position of the $SU(2)$ automorphism index A $(A, B = 1, 2)$ is related to chirality, namely (ψ_A, λ^{iA}) are chiral, $(\psi^A, \lambda^{\bar{\imath}}_A)$ antichiral.

In order to define the symplectic invariant charges let us recall the definition of the magnetic and electric charges (the moduli independent charges) in (5.160). The central charges and the matter charges are then defined as the integrals over a sphere at spatial infinity of the dressed graviphoton and matter vectors (5.171), they are given in (5.173), (5.174):

$$(Z_M) = \left(Z, \bar{Z}_{\bar{I}}\right) = i\overline{V(\phi_\infty)}^{-1} Q \tag{5.374}$$

where ϕ_∞ is the value of the scalar fields at spatial infinity. Because of (5.357) we get immediately:

$$\nabla_I Z = Z_I \, . \tag{5.375}$$

This relation can also be written $\nabla_I Z_{AB} = Z_I \epsilon_{AB}$, and considering the vielbein 1-form $\mathcal{P}^I$ dual to the frame e_I introduced in (5.356) and setting $\nabla \equiv \mathcal{P}^I \nabla_I$ we obtain $\nabla Z_{AB} = Z_I \mathcal{P}^I \epsilon_{AB}$.

The positive definite quadratic invariant $\mathscr{V}_{BH}$ in terms of the charges Z and Z_I reads

$$\mathscr{V}_{BH} = \frac{1}{2} Z\bar{Z} + Z_I \bar{Z}^I = -\frac{1}{2} Q^t \mathcal{M}(\mathcal{N}) Q \, . \tag{5.376}$$

Equation (5.376) is obtained by using exactly the same procedure as in (5.177). Invariance of $\mathscr{V}_{BH}$ implies that it is a well defined positive function on M.

5.6 Duality rotations in Noncommutative Spacetime

Field theories on noncommutative spaces have received renewed interest since their relevance in describing Dp-branes effective actions (see [79] and references therein). Noncommutativity in this context is due to a nonvanishing NS background two form on the Dp-brane. First space-like (magnetic) backgrounds $(B^{ij} \neq 0)$ were considered, then NCYM theories also with time noncommutativity $(B^{0i} \neq 0)$ have been studied [82]. The NCYM theories that can be obtained from open strings in the decoupling limit $\alpha' \to 0$ are those with B space-like or light-like (e.g. $B_{0i} = -B_{1i}$), these were also considered the only theories without unitatrity problems [83], however by applying a proper perturbative setup it was shown that also time-space noncommutative field theories can be unitary [84].

Following [79], gauge theory on a Dp-brane with constant two-form B can be described via a commutative Lagrangian and field strength $\mathcal{L}(F + B)$ or via a noncommutative one $\widehat{\mathcal{L}}(\widehat{F})$, where $\widehat{F}_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - i[A_\mu \overset{\star}{,} A_\nu]$. Here $\star$ is the star product, on coordinates $[x^\mu \overset{\star}{,} x^\nu] = x^\mu \star x^\nu - x^\nu \star x^\nu = i\Theta^{\mu\nu}$, where Θ depends on B and the metric on the Dp-brane. The commutative and the noncommutative descriptions are complementary and are related by Seiberg-Witten map (SW map) [79], [80, 81]. In the $\alpha' \to 0$ limit [79] the exact effective electromagnetic theory on a Dp-brane is noncommutative electromagnetism (NCEM), this is equivalent, via SW map, to a nonlinear commutative $U(1)$ gauge theory.

In this section we consider a D3-brane action in the slowly varying field approximation, we give an explicit expression of this nonlinear $U(1)$ theory and we show that it is self-dual when B (or Θ) is light-like. Via SW map solutions of $U(1)$ nonlinear electromagnetism are mapped into solutions of NCEM, so that duality rotations are also a symmetry of NCEM, i.e., NCEM is self-dual [85], [52]. When Θ is space-like we do not have self-duality and the S-dual of space-like NCYM is a noncommutative open string theory decoupled from closed strings [87]. Related work appeared in [88, 89, 90]. We mention that self-duality of NCEM was initially studied in [86] to first order in Θ. On one hand it is per se interesting to provide new examples of self-dual nonlinear electromagnetism, as the one we give with the lagrangian (15.355). On the other hand this lagrangian is via Seiberg-Witten map, and for slowly varying fields, just NCEM. Formally NCEM resembles $U(N)$ YM on commutative space, and on tori with rational Θ the two theories are T-dual [91]. Self-duality of NCEM then hints to a possible duality symmetry property of the equations of motion of $U(N)$ YM.

Self-Duality of the $D3$-brane action

Consider the $D3$-brane effective action in a IIB supergravity background with constant axion, dilaton NS and RR two-forms. The background two-forms can be gauged away in the bulk and we are left with the field strength $\mathcal{F} = F + B$ on the D3-brane. Here B is defined as the constant part of $\mathcal{F}$, or $B = \mathcal{F}|_{\mathrm{spatial}\,\infty}$ since F vanish at spatial infinity. For slowly varying fields the Lagrangian, in Einstein frame is essentially the Born-Infeld action with axion and dilaton. We set for simplicity $\mathcal{N} = -\rangle\infty\infty$ and $g_s = 1$, where g_s is the string coupling constant. The lagrangian is then $\mathcal{L} = \frac{-1}{\alpha'^2}\sqrt{-\det(g + \alpha'\mathcal{F})}$. The explicit expression of $\mathcal{G}$, is obtained from the definition $\mathcal{G} := \frac{\partial\mathcal{L}}{\partial F}$ and is (cf. (5.38))

$$\mathcal{G}_{\mu\nu} = \frac{\mathcal{F}^*_{\mu\nu} + \frac{\alpha'^2}{4}\mathcal{F}\mathcal{F}^*\mathcal{F}_{\mu\nu}}{\sqrt{1 + \frac{\alpha'^2}{2}\mathcal{F}^2 - \frac{\alpha'^4}{16}(\mathcal{F}\mathcal{F}^*)^2}} \ . \tag{5.377}$$

Here $\mathcal{F}^*_{\mu\nu} = \sqrt{g}\epsilon_{\mu\nu\rho\sigma}\mathcal{F}^{\rho\sigma}$, cf. footnote 2, Section 2.1. One can then consider a duality rotation by an angle γ and extract how B (the constant part of $\mathcal{F}$) transforms

$$B'_{\mu\nu} = \cos\gamma\, B_{\mu\nu} - \sin\gamma \frac{B^*_{\mu\nu} + \frac{\alpha'^2}{4}BB^*B_{\mu\nu}}{\sqrt{1 + \frac{\alpha'^2}{2}B^2 - \frac{\alpha'^4}{16}(BB^*)^2}} \ . \tag{5.378}$$

Open/closed strings and light-like noncommutativity

The open and closed string parameters are related by (see [79], the expressions for G and Θ first appeared in [92])

$$\frac{1}{g + \alpha' B} = \mathsf{G}^{-1} + \frac{\Theta}{\alpha'}$$

$$g^{-1} = (\mathsf{G}^{-1} - \Theta/\alpha')\,\mathsf{G}\,(\mathsf{G}^{-1} + \Theta/\alpha') = \mathsf{G}^{-1} - \alpha'^{-2}\Theta\,\mathsf{G}\,\Theta$$

$$\alpha' B = -(\mathsf{G}^{-1} - \Theta/\alpha')\,\Theta/\alpha'\,(\mathsf{G}^{-1} + \Theta/\alpha')$$

$$\mathsf{G}_s = g_s\sqrt{\frac{\det\mathsf{G}}{\det(g + \alpha' B)}} = g_s\sqrt{\det\mathsf{G}\ \det(\mathsf{G}^{-1} + \Theta/\alpha')} = g_s\sqrt{\det g^{-1}\ \det(g + \alpha' B)}$$

The decoupling limit $\alpha' \to 0$ with $\mathsf{G}_s, \mathsf{G}, \Theta$ nonzero and finite [79] leads to a well defined field theory only if B is space-like or light-like. Looking at the closed and open string coupling constants it is easy to see why one needs this space-like or light-like condition on B in performing this limit. Consider the coupling constants ratio G_s/g_s, that expanding the 4x4 determinant reads (here $B^2 = B_{\mu\nu}B_{\rho\sigma}g^{\mu\rho}g^{\nu\sigma}$, $\Theta^2 = \Theta^{\mu\nu}\Theta^{\rho\sigma}\mathsf{G}_{\mu\rho}\mathsf{G}_{\nu\sigma}$ and so on)

$$\frac{\mathsf{G}_s}{g_s} = \sqrt{1 + \frac{\alpha'^{-2}}{2}\Theta^2 - \frac{\alpha'^{-4}}{16}(\Theta\Theta^*)^2} = \sqrt{1 + \frac{\alpha'^2}{2}B^2 - \frac{\alpha'^4}{16}(BB^*)^2} \ . \tag{5.379}$$

Both G_s and g_s must be positive; since G and Θ are by definition finite for $\alpha' \to 0$ this implies $\Theta\Theta^* = 0$ and $\Theta^2 \geq 0$. Now $\Theta\Theta^* = 0 \Leftrightarrow \det\Theta = 0 \Leftrightarrow \det B = 0 \Leftrightarrow BB^* = 0$. In this case from (5.379) we also have $\Theta^2 = \alpha'^4 B^2$. In conclusion in order for the $\alpha' \to 0$ limit defined by keeping $\mathsf{G}_s, \mathsf{G}, \Theta$ nonzero and finite [79], to be well defined we need

$$B^2 \geq 0\ , \quad BB^* = 0 \qquad \text{i.e.} \qquad \Theta^2 \geq 0\ , \quad \Theta\Theta^* = 0 \tag{5.380}$$

This is the condition for B (and Θ) to be space-like or light-like. Indeed with Minkowski metric and in three vector notation (5.380) reads $\boldsymbol{B}^2 - \boldsymbol{E}^2 \geq 0$ and $\boldsymbol{E} \perp \boldsymbol{B}$.

If we now require the $\alpha' \to 0$ limit to be compatible with duality rotations, we immediately see that we have to consider only the light-like case $B^2 = BB^* = 0$. Indeed under $U(1)$ rotations the electric and magnetic fields mix up, in particular under a $\pi/2$ rotation (5.378) a space-like B becomes time-like.

In the light-like case $\det(g + \alpha' B) = \det(g)$, relations (5.379) simplify considerably. The open and closed string coupling constants coincide, since we set $g_s = 1$ we have $\mathsf{G}_s = g_s = 1$, this also implies $\det(\mathsf{G}) = \det(g)$ so that the hodge dual field $F^\star$ with the g metric equals the one with the G metric. Use of the relations

$$\Omega^*_{\mu\rho}\Omega^{*\rho\nu} - \Omega_{\mu\rho}\Omega^{\rho\nu} = \frac{1}{2}\Omega^2\,\delta_\mu^{\ \nu}\ \ , \quad \Omega_{\mu\rho}\Omega^{*\rho\nu} = \Omega^*{}_{\mu\rho}\Omega^{\rho\nu} = \frac{-1}{4}\Omega\Omega^*\,\delta_\mu^{\ \nu} \tag{5.381}$$

valid for any antisymmetric tensor Ω, shows that any two-tensor at least cubic in Θ (or B) vanishes. It follows that $g^{-1}\mathsf{G}\,\Theta = \Theta$ and that the raising or lowering of the Θ and B indices is independent from the metric used. We also have

$$B_{\mu\nu} = -\alpha'^{-2}\Theta_{\mu\nu}\ \ . \tag{5.382}$$

Self-duality of NCBI and NCEM

We now study duality rotations for noncommutative Born-Infeld (NCBI) theory and its zero slope limit that is NCEM. The relation between the NCBI and the BI Lagrangians is [79]

$$\widehat{\mathcal{L}}_{BI}(\widehat{F}, \mathsf{G}, \Theta, \mathsf{G}_s) = \mathcal{L}_{BI}(F + B, g) + O(\partial F) + \text{tot.der.} \tag{5.383}$$

where $O(\partial F)$ stands for higher order derivative corrections, $\widehat{F}$ is the noncommutative $U(1)$ field strength and we have set $g_s = 1$. The NCBI Lagrangian is

$$\widehat{\mathcal{L}}_{BI}(\widehat{F}, \mathsf{G}, \Theta, \mathsf{G}_s) = \frac{-1}{\alpha'^2 \mathsf{G}_s}\sqrt{-\det(\mathsf{G} + \alpha'\widehat{F})} + O(\partial\widehat{F})\ . \tag{5.384}$$

In the slowly varying field approximation the action of duality rotations on $\widehat{\mathcal{L}}_{BI}$ is derived from self-duality of $\mathcal{L}_{BI}$. If $\widehat{F}$ is a solution of the $\widehat{\mathcal{L}}_{BI}^{\mathsf{G}_s,\mathsf{G},\Theta}$ EOM then $\widehat{F}'$ obtained via

$$\widehat{F} \overset{\text{SW map}}{\longleftrightarrow} \mathcal{F} \overset{\text{duality rot.}}{\longleftrightarrow} \mathcal{F}' \overset{\text{SW map}}{\longleftrightarrow} \widehat{F}'$$

is a solution of the $\widehat{\mathcal{L}}_{BI}^{\mathsf{G}'_s,\mathsf{G}',\Theta'}$ EOM where $\mathsf{G}'_s, \mathsf{G}', \Theta'$ are obtained using (5.379) from g', B' and $g'_s = g_s = 1$.

In the light-like case we have $\mathsf{G}_s = g_s = 1$, the B rotation (5.378) simplifies to

$$B'_{\mu\nu} = \cos\gamma\, B_{\mu\nu} - \sin\gamma\, B^*_{\mu\nu}\ . \tag{5.385}$$

Using (5.385) the $U(1)$ duality action on the open string variables is

$$\mathsf{G}' = \mathsf{G}\quad , \quad \Theta'^{\mu\nu} = \cos\gamma\,\Theta^{\mu\nu} - \sin\gamma\,\Theta^{*\mu\nu}\ . \tag{5.386}$$

For Θ light-like, solutions $\widehat{F}$ of $\widehat{\mathcal{L}}^{\mathsf{G},\Theta}$ are mapped into solutions $\widehat{F}'$ of $\widehat{\mathcal{L}}^{\mathsf{G},\Theta'}$. Thus we can map solutions of $\widehat{\mathcal{L}}^{\mathsf{G},\Theta}$ into solutions of $\widehat{\mathcal{L}}^{\mathsf{G},\Theta}$, therefore the theory described by $\widehat{\mathcal{L}}^{\mathsf{G},\Theta}$ has $U(1)$ duality rotation symmetry.

In order to show self-duality of NCEM we consider the zero slope limit of (5.383) and verify that the resulting lagrangian on the r.h.s. of (5.383) is self-dual. We rewrite $\mathcal{L}_{BI}$ in terms of the open string parameters G, Θ

$$\mathcal{L}_{BI} = \frac{-1}{\alpha'^2}\sqrt{-\det(g+\alpha'\mathcal{F})} = \frac{-\sqrt{\mathsf{G}}}{\alpha'^2}\sqrt{\frac{\det(g+\alpha' B+\alpha' F)}{\det(g+\alpha' B)}}$$

$$= \frac{-1}{\alpha'^2}\sqrt{-\det(\mathsf{G}+\alpha' F+\mathsf{G}\Theta F)}\ . \tag{5.387}$$

The determinant in the last line can be evaluated as sum of products of traces (Newton-Leverrier formula). Each trace can then be rewritten in terms of the six basic Lorentz invariants $F^2,\ FF^*,\ F\Theta,\ F\Theta^*,\ \Theta^2=\Theta\Theta^*=0$, explicitly

$$\det\mathsf{G}^{-1}\det(\mathsf{G}+\alpha' F+\mathsf{G}\Theta F) = (1-\tfrac{1}{2}\Theta F)^2+\alpha'^2[\tfrac{1}{2}F^2+\tfrac{1}{4}\Theta F^*\ FF^*]-\alpha'^4(\tfrac{1}{4}FF^*)^2$$

Finally we take the $\alpha'\to 0$ limit of (5.387), by dropping the infinite constant and total derivatives the resulting Lagrangian is $\sqrt{\mathsf{G}}$ times

$$\frac{-\frac{1}{4}F^2-\frac{1}{8}\Theta F^*\ FF^*}{1-\frac{1}{2}\Theta F}\ . \tag{5.388}$$

We thus have an expression for NCEM in terms of F, Θ and G (of course $\mathsf{G}_{\mu\nu}$ can be taken $\eta_{\mu\nu}$), $\widehat{\mathcal{L}}_{\rm EM} = \sqrt{\mathsf{G}}\,\widehat{L}_{\rm EM}$,

$$\widehat{L}_{\rm EM} \equiv -\frac{1}{4}\widehat{F}\widehat{F} = \frac{-\frac{1}{4}F^2-\frac{1}{8}\Theta F^*\ FF^*}{1-\frac{1}{2}\Theta F} + O(\partial F) + \text{tot. der.} \tag{5.389}$$

The Lagrangian (15.355) satisfies the self-duality condition (5.60) with $\varphi=\Theta$, $\kappa=0$, $a=d=0$, $c=-b$ and therefore NCEM is self-dual under the $U(1)$ duality rotations (5.386) and $F'=\cos\gamma\, F-\sin\gamma\, G$. The change in $\Theta\to\Theta'$, that is not a dynamical field, can be cancelled by a rotation in space so that therefore we can map solution of the EOM of (5.389) into solutions of the EOM of (5.389) with the same value of Θ.

This duality can be enhanced to $Sp(2,\mathbb{R})$ by considering also axion and dilaton fields; also Higgs fields can be coupled, the coupling is minimal in the noncommutative theory. Using this duality one can relate space-noncommutative magnetic monopoles with a string (D1-string D3-brane configuration) to space-noncommutative electric monopoles (possibly an F-string ending on a D3-brane) [52, 53].

Acknowledgments
We thank Laura Andrianopoli, Leonardo Castellani, Anna Ceresole, Riccardo D'auria, Pietro Fré and Mario Trigiante for valuable discussions.

The work of P.A. was supported in part by the European Community's Human Potential Program under contract MRTN-CT-2004-005104 "*Constituents, Fundamental Forces and Symmetries of the Universe*" in association with University of Alessandria, and in part by the Italian MIUR under contract PRIN-2005023102.

The work of S.F. was supported in part by the European Community Human Potential Programme under contract MRTN-CT-2004-503369 "*Constituents, Fundamental Forces and Symmetries of the Universe*" in association with LNF-INFN and in part by DOE grant DE-FG03-91ER40662, Task C.

The work of B.Z. was supported in part by the Director, Office of Science, Office of High Energy and Nuclear Physics, Division of High Energy Physics of the U.S. Department of Energy under Contract DE-AC03-76SF00098, in part by the National Science Foundation under grant PHY-0457315.

5.7 Appendix: Symplectic group and transformations

5.7.1 Symplectic group (A, B, C, D and f, h and V matrices)

The symplectic group $Sp(2n, \mathbb{R})$ is the group of real $2n \times 2n$ matrices that satisfy

$$S^t \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} S = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \tag{5.390}$$

Setting $S = \left(\begin{smallmatrix} A & C \\ B & D \end{smallmatrix}\right)$ we explicitly have

$$A^t C - C^t A = 0 \ , \quad B^t D - D^t B = 0 \ , \quad A^t D - C^t B = 1 \ . \tag{5.391}$$

Since the transpose of a symplectic matrix is again symplectic we equivalently have

$$AB^t - BA^t = 0 \ , \quad CD^t - DC^t = 0 \ , \quad AD^t - BC^t = 1 \ . \tag{5.392}$$

In particular $A^tC, B^tD, CA^{-1}, BD^{-1}, A^{-1}B, D^{-1}C, AB^t, DC^t$ are symmetric matrices (in case they exist).

If D is invertible we have the factorization

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} \mathbb{1} & BD^{-1} \\ 0 & \mathbb{1} \end{pmatrix} \begin{pmatrix} D^{t^{-1}} & 0 \\ 0 & D \end{pmatrix} \begin{pmatrix} \mathbb{1} & 0 \\ D^{-1}C & \mathbb{1} \end{pmatrix} \tag{5.393}$$

where $A = D^{t^{-1}} + BD^{-1}C$ follows from $BD^{-1} = D^{t^{-1}}B^t$.

The complex basis

It is often convenient to consider the complex basis $\frac{1}{\sqrt{2}}\left(\begin{smallmatrix} F+iG \\ F-iG \end{smallmatrix}\right)$ rather than $\left(\begin{smallmatrix} F \\ G \end{smallmatrix}\right)$. The transition from the real to the complex basis is given by the symplectic and unitary matrix $\bar{A}^{-1}$, where

$$\mathcal{A} = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1} & \mathbb{1} \\ -i\mathbb{1} & i\mathbb{1} \end{pmatrix} \quad , \quad \mathcal{A}^{-1} = \mathcal{A}^\dagger \ . \tag{5.394}$$

A symplectic matrix S, belonging to the fundamental representation of $Sp(2n, \mathbb{R})$, in the complex basis reads

$$U = \mathcal{A}^{-1} S \mathcal{A} \ . \tag{5.395}$$

There is a 1-1 correspondence between matrices U as in (5.395) and complex $2n \times 2n$ matrices belonging to $U(n, n) \cap Sp(2n, \mathbb{C})$,

$$U^\dagger \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} U = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \quad , \quad U^t \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} U = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \ . \tag{5.396}$$

Equations (5.396) define a representation of $Sp(2n, \mathbb{R})$ on the complex vector space $\mathbb{C}^{2n}$. It is the direct sum of the representations $\left(\begin{smallmatrix} \psi \\ \bar{\psi} \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} \psi \\ -\bar{\psi} \end{smallmatrix}\right)$, these are real representations of real dimension $2n$. (The representation $\left(\begin{smallmatrix} \psi \\ \bar{\psi} \end{smallmatrix}\right)$ is the vector space of all linear combinations, with coefficients in $\mathbb{R}$, of vectors of the kind $\left(\begin{smallmatrix} \psi \\ \bar{\psi} \end{smallmatrix}\right)$).

The maximal compact subgroup of $U(n,n)$ is $U(n)\times U(n)$; because of the second relation in (5.396) the maximal compact subgroup of $Sp(2n,\mathbb{R})$ is $U(n)$. The usual embedding of $U(n)$ into the complex and the fundamental representations of $Sp(2n,\mathbb{R})$ are respectively

$$\begin{pmatrix} u & 0 \\ 0 & \bar{u} \end{pmatrix} \quad , \quad \begin{pmatrix} \mathrm{Re}\,u & -\mathrm{Im}\,u \\ \mathrm{Im}\,u & \mathrm{Re}\,u \end{pmatrix} \quad , \tag{5.397}$$

where u belongs to the fundamental of $U(n)$.

The f and h matrices

The f and h matrices are $n\times n$ complex matrices that satisfy the two conditions

$$(f^\dagger, h^\dagger)\begin{pmatrix} 0 & -1\!\!1 \\ 1\!\!1 & 0 \end{pmatrix}\begin{pmatrix} f \\ h \end{pmatrix} = i1\!\!1 \qquad i.e. \quad -f^\dagger h + h^\dagger f = i1\!\!1 \tag{5.398}$$

and

$$(f^t, h^t)\begin{pmatrix} 0 & -1\!\!1 \\ 1\!\!1 & 0 \end{pmatrix}\begin{pmatrix} f \\ h \end{pmatrix} = 0 \qquad i.e. \quad -f^t h + h^t f = 0 \tag{5.399}$$

These two relations are equivalent to require the real matrix

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \sqrt{2}\begin{pmatrix} \mathrm{Re}f & -\mathrm{Im}f \\ \mathrm{Re}h & -\mathrm{Im}h \end{pmatrix} \tag{5.400}$$

to be in the fundamental representation of $Sp(2n,\mathbb{R})$. Vice versa any symplectic matrix $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ leads to relations (5.398), (5.399) by defining

$$\begin{pmatrix} f \\ h \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} A - iB \\ C - iD \end{pmatrix} \ . \tag{5.401}$$

In terms of the f and h matrices we have

$$U = \mathcal{A}^{-1}\begin{pmatrix} A & B \\ C & D \end{pmatrix}\mathcal{A} = \frac{1}{\sqrt{2}}\begin{pmatrix} f + ih & \bar{f} + i\bar{h} \\ f - ih & \bar{f} - i\bar{h} \end{pmatrix} \ . \tag{5.402}$$

The V matrix and its symplectic vectors

The matrix

$$V = \begin{pmatrix} A & B \\ C & D \end{pmatrix}\mathcal{A} = \begin{pmatrix} f & \bar{f} \\ h & \bar{h} \end{pmatrix} \tag{5.403}$$

transforms from the left via the fundamental representation of $Sp(2n,\mathbb{R})$ and from the right via the complex representation of $Sp(2n,\mathbb{R})$. Since $\bar{A}$ is a symplectic matrix we have that V is a symplectic matrix, $V^t\left(\begin{smallmatrix} 0 & -1\!\!1 \\ 1\!\!1 & 0 \end{smallmatrix}\right)V = \left(\begin{smallmatrix} 0 & -1\!\!1 \\ 1\!\!1 & 0 \end{smallmatrix}\right)$, hence also its transpose V^t, $V\left(\begin{smallmatrix} 0 & -1\!\!1 \\ 1\!\!1 & 0 \end{smallmatrix}\right)V^t = \left(\begin{smallmatrix} 0 & -1\!\!1 \\ 1\!\!1 & 0 \end{smallmatrix}\right)$. The columns of the V matrix are therefore mutually symplectic vectors; also the rows are mutually symplectic vectors. Explicitly if V^ξ is the vector with components given by the ξ-th row of V, then $V^\xi{}_\rho \Omega^{\rho\sigma} V^\zeta{}_\sigma = \Omega^{\xi\zeta}$, where $\Omega = \left(\begin{smallmatrix} 0 & -1\!\!1 \\ 1\!\!1 & 0 \end{smallmatrix}\right)$.

5.7.2 The coset space $Sp(2n,\mathbb{R})/U(n)$ (M^* and $\mathcal{N}$ matrices)

All positive definite symmetric and symplectic matrices $\mathcal{S}$ are of the form

$$\mathcal{S} = gg^t \quad , \quad g \in Sp(2n,\mathbb{R}) \ . \tag{5.404}$$

Indeed consider the factorization (5.393) (since $\mathcal{S}$ is positive definite also its restriction to an n dimensional subspace is positive definite, therefore D is invertible). The factorization (5.404) is obtained for example by considering the symplectic matrix

$$g = \begin{pmatrix} 1\!\!1 & BD^{-1} \\ 0 & 1\!\!1 \end{pmatrix} \begin{pmatrix} \sqrt{D^{-1}} & 0 \\ 0 & \sqrt{D} \end{pmatrix} , \tag{5.405}$$

where the matrix $\sqrt{D}$ is the unique positive definite square root of the symmetric and positive definite matrix D. (Notice that the same proof shows that any symmetric and symplectic matrix $\left(\begin{smallmatrix} A & B \\ B^t & D \end{smallmatrix}\right)$ with invertible and positive definite matrix D is of the form gg^t and therefore is positive definite).

We can now show that the coset space $Sp(2n,\mathbb{R})/U(n)$ is the space of all positive definite symmetric and symplectic matrices. The maximal compact subgroup of $Sp(2n,\mathbb{R})$ is $H := \{g \in Sp(2n,\mathbb{R}); gg^t = 1\!\!1\}$, and we have seen in (5.397) that it is $U(n)$.

We then denote by gH the elements of $Sp(2n,\mathbb{R})/U(n)$, where $H = U(n)$, and consider the map

$$\begin{aligned} \sigma \; : \; \frac{Sp(2n,\mathbb{R})}{U(n)} \; &\to \; \{\mathcal{S} \in Sp(2n,\mathbb{R}); \mathcal{S} = \mathcal{S}^t \text{ and } \mathcal{S} \text{ positive definite}\} \\ gH \; &\mapsto \; gg^t \end{aligned} \tag{5.406}$$

This map is well defined because it does not depend on the representative $g \in Sp(2n,\mathbb{R})$ of the equivalence class gH. Formula (5.404) shows that this map is surjective. Injectivity is also easily proven: if $gg^t = g'g'^t$ then $g'^{-1}g(g'^{-1}g)^t = 1$, so that $u = g'^{-1}g$ is an element of $Sp(2n,\mathbb{R})$ that satisfies $uu^t = 1$. Therefore $u = g'^{-1}g$ belongs to the maximal compact subgroup $H = U(n)$, hence g and g' belong to the same coset.

The M^* and $\mathcal{N}$ matrices

Notice that the $n\times n$ matrices $f = (f^\Lambda{}_a)_{\Lambda,a=1,\dots n}$, are invertible. Indeed if the columns of f were linearly dependent, say $f^\Lambda{}_a\psi^a = 0$, i.e. $f\psi = 0$, with a nonzero vector ψ, then sandwiching (5.398) between $\psi^\dagger$ and ψ we would obtain

$$-(f\psi)^\dagger h\,\psi + \psi^\dagger h^\dagger f\psi = i\psi^\dagger\psi \neq 0 \tag{5.407}$$

that is absurd. Similarly also the matrix $h = (h_{\Lambda a})$ is invertible. We can then define the invertible $n \times n$ matrix

$$\mathcal{N} = hf^{-1} \tag{5.408}$$

that is symmetric (cf. (5.399)) and that has negative definite imaginary part (cf. (5.398))

$$\mathcal{N} = \mathcal{N}^{\sqcup} \quad , \quad \operatorname{Im}\mathcal{N} = -\frac{\rangle}{\in}(\mathcal{N} - \mathcal{N}^\dagger) = -\frac{\infty}{\in}(\{\{^\dagger)^{-\infty} , \tag{5.409}$$

(while $\mathcal{N}^{-\infty}$ has positive definite imaginary part $\mathcal{N}^{-\infty} - \mathcal{N}^{-\dagger} = \rangle(\langle\langle^\dagger)^{-\infty}$). Any symmetric matrix with negative definite imaginary part is of the form (5.408) for some (f,h) satisfying (5.398) and (5.399) (just consider any f that satisfyes (5.409)). There is also a 1-1- correspondence between symmetric complex matrices $\mathcal{N}$ with negative definite imaginary part and symmetric negative definite matrices M^* of $Sp(2n,\mathbb{R})$. Given $\mathcal{N}$ we consider

$$\begin{aligned}
M^*(\mathcal{N}) &= \begin{pmatrix} \mathbb{1} & -\mathrm{Re}\,\mathcal{N} \\ 0 & \mathbb{1} \end{pmatrix} \begin{pmatrix} \mathrm{Im}\,\mathcal{N} & 0 \\ 0 & \mathrm{Im}\,\mathcal{N}^{-\infty} \end{pmatrix} \begin{pmatrix} \mathbb{1} & 0 \\ -\mathrm{Re}\,\mathcal{N} & \mathbb{1} \end{pmatrix} \\
&= \begin{pmatrix} \mathrm{Im}\,\mathcal{N} + \mathrm{Re}\,\mathcal{N}\,\mathrm{Im}\,\mathcal{N}^{-\infty}\,\mathrm{Re}\,\mathcal{N} & -\mathrm{Re}\,\mathcal{N}\,\mathrm{Im}\,\mathcal{N}^{-\infty} \\ -\mathrm{Im}\,\mathcal{N}^{-\infty}\,\mathrm{Re}\,\mathcal{N} & \mathrm{Im}\,\mathcal{N}^{-\infty} \end{pmatrix} \\
&= -i \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} + \begin{pmatrix} \mathcal{N}\,\mathrm{Im}\,\mathcal{N}^{-\infty}\,\mathcal{N}^\dagger & -\mathcal{N}\,\mathrm{Im}\,\mathcal{N}^{-\infty} \\ -\mathrm{Im}\,\mathcal{N}^{-\infty}\,\mathcal{N}^\dagger & \mathrm{Im}\,\mathcal{N}^{-\infty} \end{pmatrix} \\
&= -i \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} - 2 \begin{pmatrix} hh^\dagger & -hf^\dagger \\ -fh^\dagger & ff^\dagger \end{pmatrix} \\
&= -i \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} - 2 \begin{pmatrix} -h \\ f \end{pmatrix} (-h^\dagger \;\; f^\dagger) \\
&= -2\,\mathrm{Re}\left[\begin{pmatrix} -h \\ f \end{pmatrix} (-h^\dagger \;\; f^\dagger) \right]
\end{aligned} \tag{5.410}$$

Since symmetric negative definite matrices M^* of $Sp(2n,\mathbb{R})$ parametrize the coset space $Sp(2n,\mathbb{R})/U(n)$, the matrices $\mathcal{N}$ too parametrize this coset space.

Under symplectic rotations (5.329) we have

$$\begin{pmatrix} f \\ h \end{pmatrix} \to \begin{pmatrix} f \\ h \end{pmatrix}' = S \begin{pmatrix} f \\ h \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} f \\ h \end{pmatrix} \tag{5.411}$$

and

$$\mathcal{N} \to \mathcal{N}' = (C + D\mathcal{N})(A + B\mathcal{N})^{-1} \;. \tag{5.412}$$

The transformation of the imaginary part of $\mathcal{N}$ is (recall (5.409))

$$\mathrm{Im}\,\mathcal{N} \to \mathrm{Im}\,\mathcal{N}' = (\mathcal{A} + \mathcal{B}\mathcal{N})^{-\dagger}\mathrm{Im}\,\mathcal{N}(\mathcal{A} + \mathcal{B}\mathcal{N})^{-\infty} \tag{5.413}$$

The transformation of the corresponding matrix $M^*(\mathcal{N})$ is

$$M^*(\mathcal{N}) \to \mathcal{M}^*(\mathcal{N}') = \mathcal{S}^{\sqcup-\infty}\,\mathcal{M}^*(\mathcal{N})\,\mathcal{S}^{-\infty} \;, \tag{5.414}$$

this last relation easily follows from (5.410) and from $\binom{-h}{f} = \left(\begin{smallmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{smallmatrix}\right)\binom{f}{h}$.

The relation between the negative definite symmetric matrix M^* defined in (5.410) and $\mathcal{S}$ defined in (5.404) can be obtained from their transformation properties under $Sp(2n,\mathbb{R})$,

$$M^* = -\mathcal{S}^{-1} = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \mathcal{S} \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \;. \tag{5.415}$$

We also have $\mathcal{M} = -V^{-\dagger}V^{-1}$.

5.7.3 Lie algebra of $Sp(2n,\mathbb{R})$ and $U(n)$ $(a,b,c,d$ matrices)

If we write $\left(\begin{smallmatrix} A & C \\ B & D \end{smallmatrix}\right) = \left(\begin{smallmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{smallmatrix}\right) + \epsilon \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with ϵ infinitesimal we obtain that the $2n \times 2n$ matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{5.416}$$

belongs to the Lie algebra of $Sp(2n,\mathbb{R})$ if a, b, c, d are real $n \times n$ matrices that satisfy the relations

$$a^t = -d \;, \quad b^t = b \;, \quad c^t = c \;. \tag{5.417}$$

The Lie algebra of $U(n)$ in this fundamental representation of $Sp(2n, \mathbb{R})$ is given by the matrices

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

with $b = b^t$, $a = -a^t$.

In the complex basis (5.395) the Lie algebra of $Sp(2n, \mathbb{R})$ is given by the $2n \times 2n$ matrices

$$\begin{pmatrix} \mathsf{a} & \mathsf{b} \\ \bar{\mathsf{b}} & \bar{\mathsf{a}} \end{pmatrix} \tag{5.418}$$

where a and b are complex $n \times n$ matrices that satisfy the relations

$$\mathsf{a}^\dagger = -\mathsf{a} \ , \quad \mathsf{b}^t = \mathsf{b} \ . \tag{5.419}$$

The Lie algebra of $U(n)$ in this complex basis is given by the matrices $\begin{pmatrix} \mathsf{a} & 0 \\ 0 & \bar{\mathsf{a}} \end{pmatrix}$ with $\mathsf{a}^\dagger = -\mathsf{a}$.

5.8 Appendix: Unilateral Matrix Equations

The remarkable symmetry property of the trace of the solution of the matrix equation (5.142) holds for more general matrix equations. This trace property and the structure of the solution itself are studied in [18], and with a different method in [70]; see also [71] for a unified approach based on the generalized Bezout theorem, and [69] for convergence of perturbative solutions of matrix equations and a new form of the noncommutative Lagrange inversion formula.

In this appendix we prove the symmetry property of the trace of certain solutions (and their powers) of unilateral matrix equations. These are N^{th} order matrix equations for the variable X with matrix coefficients A_i which are all on one side, e.g. on the left

$$X = A_0 + A_1 X + A_2 X^2 + \ldots + A_N X^N. \tag{5.420}$$

The matrices are all square and of arbitrary degree. We may equally consider the A_i's as generators of an associative algebra, and X an element of this algebra which satisfies the above equation. We consider the formal solution of (5.420) obtained as the limit of the sequence $X_0 = 0$, $X_{k+1} = A_0 + A_1 X_k + A_2 X_k^2 + \ldots + A_N X_k^N$. It is convenient to assign to every matrix a dimension d such that $d(X) = -1$. Using (5.420), the dimension of the matrix A_i is given by $d(A_i) = i - 1$.

First note that we can rewrite equation (5.420) as

$$1 - \sum_{i=0}^{N} A_i \ = \ 1 - X - \sum_{k=1}^{N} A_k (1 - X^k)$$

The right hand side factorizes

$$1 - \sum_{i=0}^{N} A_i \ = \ (1 - \sum_{k=1}^{N} \sum_{m=0}^{k-1} A_k X^m)(1 - X) \ .$$

Under the trace we can use the fundamental property of the logarithm, even for noncommutative objects, and obtain

$$\text{Tr} \log(1 - \sum_{i=0}^{N} A_i) \ = \ \text{Tr} \log(1 - \sum_{k=1}^{N} \sum_{m=0}^{k-1} A_k X^m) + \text{Tr} \log(1 - X) \ .$$

Using $d(A_k) = k-1$ and $d(X) = -1$ we have $d(A_k X^m) = k-m-1$ and we see that all the words in the argument of the first logarithm on the right hand side have semi-positive dimension. Since all the words in the expansion of the second term have negative dimension we obtain

$$\mathrm{Tr}\,\log(1-X) = \mathrm{Tr}\,\log(1-\sum_{i=0}^{N} A_i)\Big|_{d<0} \,. \tag{5.421}$$

On the right hand side of (5.421) one must expand the logarithm and restrict the sum to words of negative dimension. Since $d(X^r) = -r$ by extracting the dimension $d = -r$ terms from the right hand side of (5.421) we obtain

$$\mathrm{Tr}\,\phi^r \;=\; r \sum_{\substack{\{a_i\} \\ \sum(i-1)a_i = -r}} \frac{\left(\sum_{i=0}^N a_i - 1\right)!}{a_0!a_1!\dots a_N!}\,\mathrm{Tr}\,\mathcal{S}(A_0^{a_0}A_1^{a_1}\dots A_N^{a_N})\,. \tag{5.422}$$

The relevant point is that all the terms in the expansion of $\mathrm{Tr}\,\log(1-\sum_{i=0}^N A_i)$ are automatically symmetrized, this explains the symmetrization operator $\mathcal{S}$ in the $A_0, A_1, ...A_N$ matrix coefficients.

If the coefficient A_N is unity, we have the following identity for the symmetrization operators of $N+1$ and of N coefficients (words)

$$\mathcal{S}(A_0^{a_0}A_1^{a_1}\dots A_N^{a_N})|_{A_N=1} = \mathcal{S}(A_0^{a_0}A_1^{a_1}\dots A_{N-1}^{a_{N-1}})\,.$$

This is obviously true up to normalization; the normalization can be checked in the commutative case.

The trace of the solution of (5.142) can now be obtained from (5.422) by considering $r=1$ and $N=2$ and by setting A_2 to unity.

6 Classical Electromagnetic Duality for Children by JM Figueroa-O'Farrill

In this chapter we treat classical electromagnetic duality, and its manifestation (Montonen-Olive duality) in some spontaneously broken gauge theories. We start by reviewing the Dirac monopole, and then quickly move on to the 't Hooft-Polyakov monopole solution in the model described by the bosonic part of the $SO(3)$ Georgi-Glashow model. We focus on the monopole solution in the Prasad-Sommerfield limit and derive the Bogomol'nyi bound for the mass of the monopole. We show that the classical spectrum of the model is invariant under electromagnetic duality. This leads to the conjecture of Montonen and Olive. We then discuss "the Witten effect" and show that the $\mathbb{Z}_2$ electromagnetic duality extends to an $SL(2,\mathbb{Z})$ duality.

6.1 1.1 The Dirac Monopole

In this section we discuss the Dirac monopole and the Dirac(-ZwanzigerSchwinger) quantisation condition in the light of classical electromagnetic duality.

6.1.1 1.1.1 And in the beginning there was Maxwell...

Maxwell's equations in vacuo, given by

$$\vec{\partial} \cdot \vec{E} = 0 \qquad \vec{\partial} \cdot \vec{B} = 0$$
$$\vec{\partial} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \qquad \vec{\partial} \times \vec{B} = \frac{\partial \vec{E}}{\partial t}$$

are highly symmetric. In fact, they are invariant under both Lorentz transformations (in fact, conformal) and under electromagnetic duality:

$$(\vec{E}, \vec{B}) \mapsto (\vec{B}, -\vec{E}) \tag{1.2}$$

Lorentz invariance can be made manifest by introducing the field-strength $F_{\mu\nu}$, defined by

$$F^{0i} = -F^{i0} = -E^i \quad F^{ij} = -\epsilon_{ijk} B^k$$

In terms of $F_{\mu\nu}$, Maxwell's equations (1.1) become

$$\partial_\nu F^{\mu\nu} = 0 \quad \partial_\nu {}^\star F^{\mu\nu} = 0 \tag{1.3}$$

where

$${}^\star F^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\lambda\rho} F_{\lambda\rho}$$

with $\epsilon^{0123} = +1$. This formulation has the added virtue that the duality transformation (1.2) is simply

$$F^{\mu\nu} \mapsto {}^\star F^{\mu\nu} \quad {}^\star F^{\mu\nu} \mapsto -F^{\mu\nu}, \tag{1.4}$$

where the sign in the second equation is due to the fact that in Minkowski space $\star^2 = -1$.

In Minkowski space $\partial_\nu {}^\star F^{\mu\nu} = 0$ implies that $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, for some electromagnetic potential A_μ. Similarly, $\partial_\nu F^{\mu\nu} = 0$ implies that ${}^\star F_{\mu\nu} = \partial_\mu \tilde{A}_\nu - \partial_\nu \tilde{A}_\mu$, for some dual electromagnetic field $\tilde{A}_\mu$. Notice however that the duality transformation relating A_μ and $\tilde{A}_\mu$ is nonlocal. It may be easier to visualise the following two-dimensional analogue, where the duality transformation relates functions ϕ and $\tilde{\phi}$ which satisfy $\epsilon_{\alpha\beta}\partial^\beta \phi = \partial_\alpha \tilde{\phi}$, where α, β take the values 0 and 1 now.

In the presence of sources, duality is preserved provided that we include both electric and magnetic sources:

$$\partial_\nu F^{\mu\nu} = j^\nu \quad \partial_\nu {}^\star F^{\mu\nu} = k^\mu$$

and that we supplement the duality transformations (1.4) by a similar transformation of the sources:

$$j^\mu \mapsto k^\mu \quad k^\mu \mapsto -j^\mu$$

A charged point-particle of in the presence of an electromagnetic field behaves according to the Lorentz force law. If the particle is also magnetically charge, the Lorentz law is then given by

$$m \frac{d^2 x^\mu}{d\tau^2} = \left(q F^{\mu\nu} + g {}^\star F^{\mu\nu} \right) \frac{dx_\nu}{d\tau}$$

*[1] In these lectures, we shall pretend to live in Minkowski space with signature $(+---)$. We will set $c = 1$ but will often keep λ explicit.

where τ is the proper time, and m, q and g are the mass, the electric and magnetic charges, respectively. This formula is also invariant under duality provided we interchange the electric and magnetic charges of the particle: $(q, g) \mapsto (g, -q)$.

Problem: Derive the above force law from a particle action.

Notice that in the presence of magnetic sources, $\partial_\nu {}^\star F^{\mu\nu} \neq 0$ whence there is no electromagnetic potential A_μ. Nevertheless if at any given moment in time, the magnetic sources are localised in space, one may define A_μ in those regions where $k^\mu = 0$. The topology of such regions is generically nontrivial and therefore a nonsingular A_μ need not exist throughout. Instead one solves for A_μ locally, any two solutions being related, in their common domain of definition, by a gauge transformation. We will see this explicitly for the magnetic monopole.

6.1.2 1.1.2 The Dirac quantisation condition

Whereas a particle interacting classically with an electromagnetic field does so solely via the field-strength $F^{\mu\nu}$, quantum mechanically the electromagnetic potential enters explicitly in the expression for the hamiltonian. Therefore the non-existence of the potential could spell trouble for the quantisation of, say, a charged particle interacting with the magnetic field of a monopole. In his celebrated paper of 1931, Dirac [Dir31] studied the problem of the quantum mechanics of a particle in the presence of a magnetic monopole and found that a consistent quantisation forced a relation between the electric charge of the particle and the magnetic charge of the monopole: the so-called Dirac quantisation condition. We will now derive this relation.

A magnetic monopole is a point-like source of magnetic field. If we place the source at the origin in $\mathbb{R}^3$, then the magnetic field is given by

$$\vec{\boldsymbol{B}}(\vec{\boldsymbol{r}}) = \frac{g}{4\pi}\frac{\vec{\boldsymbol{r}}}{r^3} \tag{1.5}$$

where g is the magnetic charge. In these conventions, the magnetic charge is also the magnetic flux. Indeed, if Σ denotes the unit sphere in $\mathbb{R}^3$, then

$$g = \int_\Sigma \vec{\boldsymbol{B}} \cdot d\vec{\boldsymbol{S}}$$

In the complement of the origin in $\mathbb{R}^3$, $\vec{\boldsymbol{\partial}} \times \vec{\boldsymbol{B}} = 0$, whence one can try to solve for a vector potential $\vec{\boldsymbol{A}}$ obeying $\vec{\boldsymbol{B}} = \vec{\boldsymbol{\partial}} \times \vec{\boldsymbol{A}}$. For example, we can consider

$$\vec{\boldsymbol{A}}_+(\vec{\boldsymbol{r}}) = \frac{g}{4\pi r}\frac{1-\cos\theta}{\sin\theta}\hat{\boldsymbol{e}}_\phi$$

where (r, θ, ϕ) are spherical coordinates. $\vec{\boldsymbol{\partial}} \times \vec{\boldsymbol{A}}_+ = \vec{\boldsymbol{B}}$ everywhere but on the negative z-axis where $\theta = \pi$ and hence $\vec{\boldsymbol{A}}_+$is singular. Similarly,

$$\vec{\boldsymbol{A}}_-(\vec{\boldsymbol{r}}) = -\frac{g}{4\pi r}\frac{1+\cos\theta}{\sin\theta}\hat{\boldsymbol{e}}_\phi$$

also obeys $\vec{\boldsymbol{\partial}} \times \vec{\boldsymbol{A}}_- = \vec{\boldsymbol{B}}$ everywhere but on the positive z-axis where $\theta = 0$ and $\vec{\boldsymbol{A}}_-$is singular. It isn't that we haven't been clever enough, but that any $\vec{\boldsymbol{A}}$ which obeys $\vec{\boldsymbol{\partial}} \times \vec{\boldsymbol{A}} = \vec{\boldsymbol{B}}$ over some region will always be singular on some string-like region: the celebrated Dirac string.

Over their common domain of definition (the complement of the z-axis in $\mathbb{R}^3$) $\vec{\partial} \times \left(\vec{\boldsymbol{A}}_+ - \vec{\boldsymbol{A}}_-\right) = 0$, whence one would expect that there exists a function χ so that $\vec{A}_+ - \vec{A}_- = \vec{\partial}\chi$. However the complement of the z-axis is not simply-connected, and χ need only be defined locally. For example, restricting ourselves to $\theta = \frac{\pi}{2}$, we find that

$$\vec{\boldsymbol{A}}_+ - \vec{\boldsymbol{A}}_- = \frac{g}{2\pi r}\hat{\boldsymbol{e}}_\phi = \vec{\boldsymbol{\partial}}\left(\frac{g}{2\pi}\phi\right)$$

but notice that since ϕ is an angle, the function χ is not continuous. It couldn't possibly be continuous, for if it were there would be no flux. Indeed, if Σ again denotes the unit sphere in $\mathbb{R}^3, \Sigma^\pm$the upper and lower hemispheres respectively, and E the equator, the flux can be computed in terms of χ as follows:

$$\begin{aligned} g &= \int_\Sigma \vec{\boldsymbol{B}} \cdot d\vec{\boldsymbol{S}} \\ &= \int_{\Sigma^+} \left(\vec{\boldsymbol{\partial}} \times \vec{\boldsymbol{A}}_+\right) \cdot d\vec{\boldsymbol{S}} + \int_{\Sigma^-} \left(\vec{\boldsymbol{\partial}} \times \vec{\boldsymbol{A}}_-\right) \cdot d\vec{\boldsymbol{S}} \\ &= \int_E \vec{\boldsymbol{A}}_+ \cdot d\vec{\boldsymbol{\ell}} - \int_E \vec{\boldsymbol{A}}_- \cdot d\vec{\boldsymbol{\ell}} \\ &= \int_E \vec{\boldsymbol{\partial}}\chi \cdot d\vec{\boldsymbol{\ell}} \\ &= \chi(2\pi) - \chi(0) \end{aligned}$$

Suppose now that we are quantising a particle of mass m and charge q in the field of a magnetic monopole. The Schrödinger equation satisfied by the wave-function is

$$-\frac{\lambda^2}{2m}\nabla^2\psi = i\lambda\frac{\partial\psi}{\partial t}$$

where $\vec{\boldsymbol{\nabla}} = \vec{\boldsymbol{\partial}} + ie\vec{\boldsymbol{A}}$, for $e = q/\lambda$. The Schrödinger equation is invariant under the gauge-transformations:

$$\vec{\boldsymbol{A}} \mapsto \vec{\boldsymbol{A}} + \vec{\boldsymbol{\partial}}\chi \quad \text{and} \quad \psi \mapsto \exp(-ie\chi)\psi$$

This gauge invariance guarantees that solutions of the Schrödinger equation obtained locally with a particular $\vec{\boldsymbol{A}}$ will patch up nicely, provided that the wave-function be single-valued. This condition means that

$$\exp(-ie\chi) = \exp(-ieg\phi/2\pi)$$

must be a single-valued function, which is equivalent to the Dirac quantisation condition:

$$eg = 2\pi n \quad \text{for some} \quad n \in \mathbb{Z} \tag{1.6}$$

The Dirac quantisation condition has the following physical interpretation. Classically, there is not much of a distinction between a a magnetic monopole and a very long and very thin solenoid. The field inside the solenoid is of course, different, but in the limit in which the solenoid becomes infinitely long (on one end only) and infinitesimally thin, so that the inside of the solenoid lies beyond the probe of a classical experiment, the field at the end of the solenoid is indistinguishable from that of a magnetic monopole. Quantum mechanically, however, one can in principle detect the solenoid through the quantum interference pattern predicted by the Bohm-Aharanov effect. The condition for the absence of the interference is precisely the Dirac quantisation condition.

6.1.3 1.1.3 Dyons and the Zwanziger-Schwinger quantisation condition

A quicker, more heuristic derivation of the Dirac quantisation condition (1.6) follows by invoking the quantisation of angular momentum. The orbital angular momentum $\vec{\boldsymbol{L}} = \vec{\boldsymbol{r}} \times m\dot{\vec{\boldsymbol{r}}}$ of a particle of mass m and charge q in the presence of a magnetic monopole (1.5) is not conserved. Indeed, using the Lorentz force law,

$$\begin{aligned}\frac{d\vec{\boldsymbol{L}}}{dt} &= \vec{\boldsymbol{r}} \times m\ddot{\vec{\boldsymbol{r}}} \\ &= \vec{\boldsymbol{r}} \times (q\dot{\vec{\boldsymbol{r}}} \times \vec{\boldsymbol{B}}) \\ &= \frac{qg}{4\pi r^3}\vec{\boldsymbol{r}} \times (\dot{\vec{\boldsymbol{r}}} \times \vec{\boldsymbol{r}}) \\ &= \frac{d}{dt}\left(\frac{qg}{4\pi}\frac{\vec{\boldsymbol{r}}}{r}\right)\end{aligned}$$

whence the conserved quantity is instead

$$\vec{\boldsymbol{J}} \equiv \vec{\boldsymbol{L}} - \frac{qg}{4\pi}\frac{\vec{\boldsymbol{r}}}{r}$$

a result dating to 1896 and due to Poincaré.

Exercise

1.1 (Angular momentum due to the electromagnetic field)

Show that the correction term is in fact nothing else but the angular momentum of the electromagnetic field itself:

$$\vec{\boldsymbol{J}}_{\text{em}} = \int_{\mathbb{R}^3} d^3r\, \vec{\boldsymbol{r}} \times (\vec{\boldsymbol{E}} \times \vec{\boldsymbol{B}})$$

where the $\vec{\boldsymbol{E}}$-field is the one due to the charged particle.

If we now assume that the electromagnetic angular momentum is separately quantised, so that

$$\left|\vec{\boldsymbol{J}}_{\text{em}}\right| = \frac{1}{2}n\lambda \quad \text{for some} \quad n \in \mathbb{Z}$$

we recover (1.6) again. The virtue of this derivation is that it provides a quick proof of the Zwanziger-Schwinger quantisation condition for dyons, as the following exercise asks you to show.

Exercise 1.2 (The Zwanziger-Schwinger quantisation condition)

A dyon is a particle which possesses both electric and magnetic charge. Consider two dyons of charges $(q = e\lambda, g)$ and $(q' = e'\lambda, g')$. Show that imposing the quantisation of the angular momentum of the resulting electromagnetic field yields the following condition:

$$eg' - e'g = 2\pi n \quad \text{for some } n \in \mathbb{Z} \tag{1.7}$$

Notice that the existence of the "electron" (that is, a particle with charges $(e, 0)$) does not tell us anything about the electric charge of a monopole (q, g); although it does tell us something about the difference between the electric charges of two such monopoles: (q, g) and (q', g). Indeed, (1.7) tells us immediately that $g\,(q - q') = 2\pi n$ for some integer n. If g has the minimum magnetic charge $g = 2\pi/e$, then the difference between the electric charges of the dyons (q, g) and (q', g) is an integer multiple of the electric charge of the electron: $q - q' = ne$ for some integer n. But we cannot say anything further about the absolute magnitude of either q or q'.

Exercise 1.3 (Dyonic spectrum in CP non-violating theories)

Prove that if CP is not violated, then in fact there are only two (mutually exclusive) possibilities: either $q =$ ne or $q = ne + \frac{1}{2}e$.

(Hint: use that under CP: $(q, g) \mapsto (-q, g)$. Why?)

We will see later when we discuss the so-called "Witten effect" that this gets modified in the presence of a CP-violating term, and the electric charge of the dyon will depend explicitly on the θ angle measuring the extent of the CP violation.

6.2 1.2 The 't Hooft-Polyakov Monopole

In 1974, 't Hooft [tH74] and Polyakov [Pol74] independently discovered that the bosonic part of the Georgi-Glashow model admits finite energy solutions that from far away look like Dirac monopoles. In contrast with the Dirac monopole, these solutions are everywhere regular and do not necessitate the introduction of a source of magnetic charge - this being due to the "twists" in (the vacuum expectation value of) the Higgs field.

6.2.1 1.2.1 The bosonic part of the Georgi-Glashow model

The Georgi-Glashow model was an early proposal to describe the electroweak interactions. We will be concerned here only with the bosonic part of the model which consists of an $SO(3)$ Yang-Mills field theory coupled to a Higgs field in the adjoint representation. The lagrangian density is given by

$$\mathcal{L} = -\frac{1}{4}\overrightarrow{\mathrm{G}}^{\mu\nu} \cdot \overrightarrow{\mathrm{G}}_{\mu\nu} + \frac{1}{2}D^{\mu}\vec{\phi} \cdot D_{\mu}\overrightarrow{\boldsymbol{\phi}} - V(\phi) \tag{1.8}$$

where

the gauge field-strength $\overrightarrow{\mathrm{G}}_{\mu\nu}$ is defined by

$$\overrightarrow{\mathrm{G}}_{\mu\nu} = \partial_{\mu}\overrightarrow{\mathrm{W}}_{\nu} - \partial_{\nu}\overrightarrow{\mathrm{W}}_{\mu} - e\overrightarrow{\mathrm{W}}_{\mu} \times \overrightarrow{\mathrm{W}}_{\nu}$$

where $\overrightarrow{\mathrm{W}}_{\mu}$ are gauge potentials taking values in the Lie algebra of $SO(3)$, which we identify with $\mathbb{R}^3$ with the cross product for Lie bracket;

the Higgs field $\overrightarrow{\boldsymbol{\phi}}$ is a vector in the (three-dimensional) adjoint representation of $SO(3)$, with components $\phi_a = (\phi_1, \phi_2, \phi_3)$ which is minimally coupled to the gauge field via the gauge-covariant derivative:

$$D_{\mu}\overrightarrow{\boldsymbol{\phi}} = \partial_{\mu}\overrightarrow{\boldsymbol{\phi}} - e\overrightarrow{\mathrm{W}_{\mu} \times \vec{\phi}}$$

the Higgs potential $V(\phi)$ is given by

$$V(\phi) = \frac{1}{4}\lambda\left(\phi^2 - a^2\right)^2$$

where $\phi^2 = \vec{\phi} \cdot \vec{\phi}$ and λ is assumed non-negative.

The lagrangian density $\mathcal{L}$ is invariant under the following $SO(3)$ gauge transformations:

$$\begin{aligned} &\overrightarrow{\boldsymbol{\phi}} \mapsto \overrightarrow{\boldsymbol{\phi}}' = g(x)\overrightarrow{\boldsymbol{\phi}} \\ &\overrightarrow{\mathrm{W}}_{\mu} \mapsto \overrightarrow{\mathrm{W}}'_{\mu} = g(x)\overrightarrow{\mathrm{W}}_{\mu}g(x)^{-1} + \frac{1}{e}\partial_{\mu}g(x)g(x)^{-1} \end{aligned} \tag{1.9}$$

where $g(x)$ is a possibly x-dependent 3×3 orthogonal matrix with unit determinant.

The classical dynamics of the fields $\vec{W}_{\mu}$ and $\vec{\phi}$ are determined from the equations of motion

$$D_\nu \overrightarrow{\mathrm{G}}^{\mu\nu} = -e\overrightarrow{\boldsymbol{\phi}} \times D^\mu \overrightarrow{\boldsymbol{\phi}} \quad D^\mu D_\mu \vec{\phi} = -\lambda \left(\phi^2 - a^2\right) \vec{\phi} \tag{1.10}$$

and by the Bianchi identity

$$D_\mu {}^\star \vec{G}^{\mu\nu} = 0 \tag{1.11}$$

where ${}^*\vec{G}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\lambda\rho}\overrightarrow{\mathrm{G}}_{\lambda\rho}$.

The canonically conjugate momenta to the gauge field $\overrightarrow{\mathrm{W}}_\mu$ and the Higgs $\vec{\phi}$ are given by

$$\overrightarrow{\mathrm{E}}^i = -\overrightarrow{\mathrm{G}}^{0i} \quad \vec{\Pi} = D_0\vec{\phi} \tag{1.12}$$

Defining $\overrightarrow{\mathrm{B}}_i$ by

$$\overrightarrow{\mathrm{G}}_{ij} = -\epsilon_{ijk}\overrightarrow{\mathrm{B}}^k = +\epsilon_{ijk}\overrightarrow{\mathrm{B}}_k$$

we can write the energy density as

$$\mathcal{H} = \frac{1}{2}\overrightarrow{\mathrm{E}}_i \cdot \overrightarrow{\mathrm{E}}_i + \frac{1}{2}\overrightarrow{\boldsymbol{\Pi}} \cdot \overrightarrow{\boldsymbol{\Pi}} + \frac{1}{2}\overrightarrow{\mathrm{B}}_i \cdot \overrightarrow{\mathrm{B}}_i + \frac{1}{2}D_i\overrightarrow{\boldsymbol{\phi}} \cdot D_i\overrightarrow{\boldsymbol{\phi}} + V(\phi) \tag{1.13}$$

which is manifestly positive-semidefinite and also gauge-invariant.

We define a vacuum configuration to be one for which the energy density vanishes. This means that

$$\overrightarrow{\mathrm{G}}_{\mu\nu} = 0 \quad D^\mu\overrightarrow{\boldsymbol{\phi}} = 0 \quad V(\phi) = 0$$

For example, $\overrightarrow{\boldsymbol{\phi}} = a\hat{\boldsymbol{e}}_3$ and $\overrightarrow{\mathrm{W}}_\mu = 0$ is such a configuration, where $(\hat{\boldsymbol{e}}_a)$ is an orthonormal basis for the three-dimensional representation space where the Higgs field takes values. We also define the Higgs vacuum as those configurations of the Higgs field which satisfy the latter two equations above. Notice that in the Higgs vacuum, the Higgs field obeys $\phi^2 = a^2$. Any such vacuum configuration is not invariant under the whole $SO(3)$, but only under an $SO(2) \cong U(1)$ subgroup, therefore this model exhibits spontaneous symmetry breaking.

Exercise 1.4 (The spectrum of the model)

Let $\overrightarrow{\boldsymbol{\phi}} = \overrightarrow{\boldsymbol{a}} + \overrightarrow{\boldsymbol{\varphi}}$ where $\overrightarrow{\boldsymbol{a}}$ is a constant vector obeying $\overrightarrow{\boldsymbol{a}} \cdot \overrightarrow{\boldsymbol{a}} = a^2$. Expanding the lagrangian density in terms of $\overrightarrow{\boldsymbol{\varphi}}$, show that the model consists of a massless vector boson $A_\mu = \frac{1}{a}\overrightarrow{\boldsymbol{a}} \cdot \overrightarrow{\mathrm{W}}_\mu$ which we will identify with the photon, a massive scalar field $\varphi = \frac{1}{a}\overrightarrow{\boldsymbol{a}} \cdot \overrightarrow{\boldsymbol{\phi}}$ and two massive vector bosons $W_\mu^\pm$with the charge assignments given
in Table 1.1.

(Hint: The masses are read off from the quadratic terms of the lagrangian density:

$$\mathcal{L} = \cdots + \frac{1}{2}\left(\frac{M_H}{\lambda}\right)^2 \varphi^2 + \frac{1}{2}\left(\frac{M_W}{\lambda}\right)^2 W_\mu^+ W^{\mu-} + \cdots$$

whereas the charges are read off from the coupling to the photon. The photon couples minimally via the covariant derivative $\nabla_\mu = \partial_\mu + iQ/\lambda A_\mu$. By examining how this covariant derivative embeds in the $SO(3)$ covariant derivative one can read off what Q are for the fields in the spectrum.)

Field	Mass	Charge
A_μ	0	0
φ	$M_H = a\sqrt{2\lambda}\lambda$	0
$W_\mu^\pm$	$M_W = ae\lambda$	$\pm e\lambda$

Table 1.1: The perturbative spectrum after higgsing.

6.2.2 1.2.2 Finite-energy solutions: the 't Hooft-Polyakov Ansatz

We now investigate the properties of finite-energy non-dissipative solutions to the equations of motion (1.10). But first let us remark a few properties of arbitrary finite-energy field configurations. The energy of a given field configuration is the spatial integral $E = \int d^3x\mathcal{H}$ of the energy density $\mathcal{H}$ given by equation (1.13). Finite energy means that the integral exists, hence the fields must approach a vacuum configuration asymptotically. In particular the Higgs field approaches the Higgs vacuum at spatial infinity. If we think of the Higgs potential V as a function $V : \mathbb{R}^3 \to \mathbb{R}$, let us define $\mathcal{M}_0 \subset \mathbb{R}^3$ as those points $\overrightarrow{\boldsymbol{x}} \in \mathbb{R}^3$ for which $V(\overrightarrow{\boldsymbol{x}}) = 0$. In the model at hand, $\mathcal{M}_0$ is the sphere of radius a, hence in any finite-energy configuration the Higgs field defines a function from the sphere at spatial infinity to $\mathcal{M}_0$:

$$\vec{\phi}_\infty(\hat{\boldsymbol{r}}) \equiv \lim_{r\to\infty} \vec{\phi}(\overrightarrow{\boldsymbol{r}}) \in \mathcal{M}_0$$

We will assume that the resulting function $\vec{\phi}_\infty$ is actually continuous. This would follow from some uniformity property of the limit and such a property has been proven by Taubes JT80.

It is well-known that the space of continuous functions from a sphere to a sphere is disconnected: it has an infinite number of connected components
indexed by an integer called the degree of the map. A constant map has degree zero, whereas the identity map has degree 1. Heuristically, the degree is the number of times one sphere wraps around the other. It is the direct two-dimensional generalisation of the winding number for maps from a circle to a circle.

Taking these remarks into account it is not difficult to construct maps of arbitrary degree. Consider the map $f_n : \mathbb{R}^3\backslash\{0\} \to \mathbb{R}^3$ defined by

$$f_n(\overrightarrow{\boldsymbol{r}}) = (\sin\theta\cos n\varphi, \sin\theta\sin n\varphi, \cos\theta) \tag{1.14}$$

where (r, θ, φ) are spherical coordinates. The map f_n restricts to a map from the unit sphere in $\mathbb{R}^3$ to itself which has degree n.

The topological number of a finite-energy configuration is defined to be the degree of the map $\vec{\phi}_\infty$. The zero energy vacuum configuration $\vec{W}_\mu = 0$, $\vec{\phi} = a\hat{\boldsymbol{e}}_3$ has zero degree, since $\vec{\phi}_\infty$ is constant. The topological number of a field configuration-being an integer-is invariant under any continuous deformation. In particular it is invariant under time evolution, and under gauge transformations, since the gauge group is connected. Hence if we set up a finite-energy field configuration at some moment in time whose topological number is different from zero, it will never dissipate; that is, it will never evolve in time towards a trivial solution. In other words, in a sense it will be stable.

We now investigate whether such stable solutions actually exist. We will narrow our search to spherically symmetric static solutions-a solution is defined to be static if it is time-independent and in addition the timecomponent of the gauge field $\overrightarrow{\mathrm{W}}_0$ vanishes.

One may be tempted to think that the this latter condition is simply a choice of gauge. Indeed it is easy to show that $\vec{W}_0 = 0$ up to a gauge transformation, but the gauge transformation is actually time-dependent which is not allowed, since we are looking for time-independent solutions. Coming soon: More details on the explicit time-dependent gauge transformation.

It follows from (1.12) that for static field configurations both $\vec{E}$ and $\vec{\Pi}$ vanish, and hence the energy agrees up to a sign with the lagrangian. This means that a field configuration will be a solution to the classical equations of motion if and only if it extremises the energy.

The 't Hooft-Polyakov Ansatz for the monopole is given by

$$\begin{aligned}\vec{\boldsymbol{\phi}}(\vec{\boldsymbol{r}}) &= \frac{\vec{\boldsymbol{r}}}{er^2}H(aer)\\ \mathrm{W}_a^i &= -\epsilon_{aij}\frac{r^j}{er^2}(1-K(aer))\\ \mathrm{W}_a^0 &= 0\end{aligned} \tag{1.15}$$

for some arbitrary functions H and K.

Exercise

1.5 (Boundary conditions on H and K)

Plugging the Ansatz into the expression for the energy density derive the following formula for the energy:

$$E = \frac{4\pi a}{e}\int_0^\infty \frac{d\xi}{\xi^2}\times$$
$$\left(\xi^2\frac{dH}{d\xi} + \frac{1}{2}\left(\xi\frac{dH}{d\xi} - H\right)^2 + \frac{1}{2}\left(K^2-1\right)^2 + K^2H^2 + \frac{\lambda}{4e^2}\left(H^2-\xi^2\right)^2\right) \tag{1.16}$$

Deduce that the integral exists provided that the following boundary conditions hold:

$$\begin{aligned}&K\to 0 \quad \text{and} \quad H/\xi \to 1 \quad \text{sufficiently fast as } \xi\to\infty\\ &K-1\le O(\xi) \quad \text{and} \quad H\le O(\xi) \quad \text{as } \xi\to 0\end{aligned} \tag{1.17}$$

This last equation means that H and K approach 0 and 1 respectively at least linearly in ξ as $\xi \to 0$.

Notice that with the above boundary conditions,

$$\vec{\boldsymbol{\phi}}_\infty(\hat{\boldsymbol{r}}) \equiv \lim_{r\to\infty}\frac{\vec{\boldsymbol{r}}}{er^2}H(aer) = a\hat{\boldsymbol{r}}$$

which is (homotopic to) the identity map, and hence has degree 1. In other words, the topological number of such a field configuration is 1. If such a solution exists it is therefore stable and non-dissipative.

Exercise 1.6 (The equations of motion for H and K)

Work out the equations of motion for the functions H and K in either of two ways: either plug the Ansatz into the equations of motion (1.10) or else extremise the energy subject to the above boundary conditions. In either case you should get the following coupled nonlinear system of ordinary differential equations:

$$\begin{aligned}\xi^2\frac{d^2K}{d\xi^2} &= KH^2 + K\left(K^2-1\right)\\ \xi^2\frac{d^2H}{d\xi^2} &= 2K^2H + \frac{\lambda}{e^2}H\left(H^2-\xi^2\right)\end{aligned} \tag{1.18}$$

Initial numerical studies of the above differential equations for H and K together with the boundary conditions (1.17) suggested the existence of a solution. This was later proven rigourously by Taubes JT80. Notice that the asymptotic limit of the equations (1.18) in the limit $\xi \to \infty$ yields:

$$\frac{d^2K}{d\xi^2} = K$$
$$\frac{d^2h}{d\xi^2} = 2\frac{\lambda}{e^2}h$$

where $h \equiv H - \xi$. The above equations can be solved for at once and one finds that the solutions compatible with the boundary conditions are

$$K \sim \exp(-\xi) = \exp(-M_W r/\lambda)$$
$$h \sim \exp(-M_H r/\lambda)$$

where M_W and M_H were obtained in Exercise 1.4. This means that the solution describes an object of finite size given by the largest of the Compton wavelengths λ/M_H or λ/M_W.

In order to identify the solution provided by the 't Hooft-Polyakov Ansatz we investigate the asymptotic electromagnetic field. Recall that the electromagnetic potential is identified with $A_\mu = \frac{1}{a}\vec{\phi} \cdot \overrightarrow{\mathrm{W}}_\mu$, corresponding to the $U(1) \subset SO(3)$ defined as the stabiliser of $\overrightarrow{\boldsymbol{\phi}}$. The electromagnetic field can therefore be identified with $F_{\mu\nu} = \frac{1}{a}\vec{\phi} \cdot \overrightarrow{\mathrm{G}}_{\mu,\nu}$. Because the 't Hooft-Polyakov Ansatz corresponds to a static solution, there is no electric field: $F_{0i} = 0$. However, as the next exercise shows, there is a magnetic field.

Exercise 1.7 (Asymptotic form of the electromagnetic field)

Show that the asymptotic form of $F_{ij} = \frac{1}{a}\vec{\phi} \cdot \vec{G}_{ij}$ is given by

$$F_{ij} = \epsilon_{ijk}\frac{r^k}{er^3} \tag{1.19}$$

The form (1.19) of the electromagnetic field shows that the asymptotic magnetic field is that of a magnetic monopole:

$$\overrightarrow{\boldsymbol{B}} = -\frac{1}{e}\frac{\overrightarrow{\boldsymbol{r}}}{r^3}$$

A quick comparison with equation (1.5) reveals that the magnetic charge of a 't Hooft-Polyakov monopole is (up to a sign) twice the minimum magnetic charge consistent with the electric charge e and the Dirac quantisation condition; that is, twice the Dirac charge corresponding to e. This follows from the fact that the electromagnetic $U(1)$ is embedded in $SO(3)$ in such a way that the electric charge is the eigenvalue of the T_3 isospin generator, which here is in the adjoint representation, which has integral isospin. The minimum electric charge is therefore $e_{\min} = \frac{1}{2}e$, relative to which the charge of the 't Hooft-Polyakov monopole is indeed one Dirac charge, again up to a sign.

In fact, there is another solution with the opposite magnetic charge. It is obtained from the 't Hooft-Polyakov monopole by performing a parity transformation on the Ansatz.

One might wonder whether there also exist dyonic solutions. These solutions would not be static in that $\vec{W}_0$ would be different from zero, but time-independent dyonic solutions have been found by Julia and Zee JZ75 shortly after the results of 't Hooft and Polyakov.

In summary, we see that the 't Hooft-Polyakov solution describes an object of finite size which from far away cannot be distinguished from a Dirac monopole of charge $-4\pi/e$. In contrast with the Dirac monopole, the 't Hooft-Polyakov monopole is everywhere smooth- this being due to the massive fields which become relevant as we approach the "core" of the monopole.

6.2.3 1.2.3 The topological origin of the magnetic charge

Although the 't Hooft-Polyakov monopole is indistinguishable from far away from a Dirac monopole, as we approach its core the massive fields become relevant and the difference becomes evident. In contrast to the Dirac monopole, which necessitates a singular point-like magnetic source at the origin, the 't Hooft-Polyakov monopole is everywhere smooth and its magnetic charge is purely topological and, as we will see in this section, due completely to the behaviour of the Higgs field far away from the core.

Because of the exponential decay of the massive fields away from the core of the monopole, we notice that the Higgs field approaches the Higgs vacuum. In other words, a large distance away from the core of the monopole, the Higgs field satisfies

$$D_\mu \overrightarrow{\phi} = 0 \tag{1.20}$$

$$\vec{\phi} \cdot \vec{\phi} = a^2 \tag{1.21}$$

up to terms of order $O(\exp(-r/R))$ where R is the effective size of the monopole, which is governed by the mass of the heavy particles.

Notice that equation (1.20) already implies that $\vec{\phi} \cdot \vec{\phi}$ is a constant. Indeed,

$$\begin{aligned}
\partial_\mu(\vec{\phi} \cdot \vec{\phi}) &= 2\vec{\phi} \cdot \partial_\mu \vec{\phi} \\
&= 2e\vec{\phi} \cdot \left(\vec{W}_\mu \times \vec{\phi}\right) \qquad \text{by 1.20} \\
&= 0
\end{aligned}$$

What 1.21 tells us is that this constant is such that the potential attains its minimum.

It is therefore reasonable to assume that any finite-energy solution (not necessarily static or time-independent) of the Yang-Mills-Higgs system (1.10) satisfies equations (1.20) and (1.21) except in a finite number of well-separated compact localised regions in space, which we shall call monopoles. In other words, we are considering a "dilute gas of monopoles" surrounded by a Higgs vacuum.

Notice that in the Higgs vacuum, $\vec{\phi} \times \overrightarrow{\mathrm{W}}_\mu = -\frac{1}{e}\partial_\mu \vec{\phi}$, whence $\overrightarrow{\mathrm{W}}_\mu$ is fully determined except for the component in the $\vec{\phi}$-direction, which we denote by
A_μ. Computing the components perpendicular to $\vec{\phi}$ we find that

$$\overrightarrow{\mathrm{W}}_\mu = \frac{1}{a^2 e} \overrightarrow{\phi} \times \partial_\mu \overrightarrow{\phi} + \frac{1}{a} \overrightarrow{\phi} A_\mu$$

Exercise 1.8 (Gauge field-strength in the Higgs vacuum)

Show that the field-strength in the Higgs vacuum points in the $\overrightarrow{\phi}$-direction and is given by $\overrightarrow{\mathrm{G}}_{\mu\nu} = \frac{1}{a}\vec{\phi} F_{\mu\nu}$ where

$$F_{\mu\nu} = \frac{1}{a^3 e} \vec{\phi} \cdot \left(\partial_\mu \vec{\phi} \times \partial_\nu \vec{\phi}\right) + \partial_\mu A_\nu - \partial_\nu A_\mu$$

Using the equations of motion (1.10) and the Bianchi identity (1.11) prove that $F_{\mu\nu}$ satisfies Maxwell's equations (1.3).

Now let Σ be a surface in the Higgs vacuum enclosing some monopoles in the volume it bounds. The magnetic flux through Σ measures the magnetic charge. Notice that A_μ doesn't contribute, and that we get:

$$\begin{aligned} g_\Sigma &\equiv \int_\Sigma \overrightarrow{\boldsymbol{B}} \cdot d\overrightarrow{\boldsymbol{S}} \\ &= -\frac{1}{2ea^2} \int_\Sigma \epsilon_{ijk} \overrightarrow{\boldsymbol{\phi}} \cdot \left(\partial_j \overrightarrow{\boldsymbol{\phi}} \times \partial_k \overrightarrow{\boldsymbol{\phi}} \right) dS_i \end{aligned}$$

Notice that only the components of $\partial_i \vec{\phi}$ tangential to Σ contribute to the integral and therefore the magnetic charge only depends on the behaviour of $\vec{\phi}$ on Σ. Furthermore it only depends on the homotopy class of $\vec{\phi}$ as map $\Sigma \to \boldsymbol{\mathcal{M}}_0$; in other words, the above integral is invariant under deformations $\delta\vec{\phi}$ of $\vec{\phi}$ which preserve the Higgs vacuum:

$$D_\mu \delta\vec{\phi} = 0 \quad \text{and} \quad \vec{\phi} \cdot \delta\vec{\phi} = 0$$

To see this, let's compute the variation of g_Σ under such a deformation of $\overrightarrow{\boldsymbol{\phi}}$. Notice first that

$$\begin{aligned} &\delta \left(\epsilon_{ijk} \overrightarrow{\boldsymbol{\phi}} \cdot \left(\partial_j \overrightarrow{\boldsymbol{\phi}} \times \partial_k \overrightarrow{\boldsymbol{\phi}} \right) \right) = \\ &3\epsilon_{ijk} \delta \overrightarrow{\boldsymbol{\phi}} \cdot \left(\partial_j \overrightarrow{\boldsymbol{\phi}} \times \partial_k \overrightarrow{\boldsymbol{\phi}} \right) + 2\epsilon_{ijk} \partial_j \left(\overrightarrow{\boldsymbol{\phi}} \cdot \left(\delta \overrightarrow{\boldsymbol{\phi}} \times \partial_k \overrightarrow{\boldsymbol{\phi}} \right) \right) \end{aligned}$$

By Stokes' theorem, the second term in the right-hand-side integrates to zero. Now, because $\vec{\phi} \cdot \partial_j \vec{\phi} = 0, \vec{\phi} \cdot \left(\partial_j \vec{\phi} \times \partial_k \vec{\phi} \right) = 0$, whence $\partial_j \vec{\phi} \times \partial_k \vec{\phi}$ is parallel to $\vec{\phi}$. Hence, $\delta\vec{\phi} \cdot \left(\partial_j \vec{\phi} \times \partial_k \vec{\phi} \right) = 0$. In other words, $\delta g_\Sigma = 0$. This means that g_Σ is invariant under arbitrary deformations of $\vec{\phi}$ and hence under any deformation which can be achieved by iterating infinitesimal deformations: homotopies. Examples of homotopies are:

time-evolution of $\vec{\phi}$; continuous gauge transformations on $\vec{\phi}$; continuous changes of Σ within the Higgs vacuum.

Exercise 1.9 (Additivity of the magnetic charge g_Σ)

Use the invariance of the magnetic charge under the last of the above homotopies, to argue that the magnetic charge is additive.

(Hint: Use a "contour" deformation argument.)

Notice that the magnetic charge can be written as $g_\Sigma = -\frac{4\pi}{e} N_\Sigma$, where

$$N_\Sigma = \frac{1}{8\pi a^3} \int_\Sigma dS_i \epsilon_{ijk} \overrightarrow{\boldsymbol{\phi}} \times \left(\partial_j \overrightarrow{\boldsymbol{\phi}} \times \partial_k \overrightarrow{\boldsymbol{\phi}} \right) \tag{1.22}$$

which as the next exercise asks you to show, is the degree of the map $\vec{\phi}$: $\Sigma \to \boldsymbol{\mathcal{M}}_0$.

Exercise 1.10 (Dirac quantisation condition revisited)

Show that N_Σ is the integral of the jacobian of the map $\vec{\phi} : \Sigma \to \boldsymbol{\mathcal{M}}_0$, which is the classical definition of the degree of the map. This means that N_Σ is an integer; a fact of which you may convince yourself by showing that if f_n is the map defined by (1.14), then the value of N_Σ when Σ is, say, the unit sphere in $\mathbb{R}^3$, is equal to n. Taking this into account we recover again the Dirac quantisation condition:

$$eg_\Sigma = -4\pi N_\Sigma \tag{1.23}$$

with the same caveat as before about the fact that the minimum magnetic charge is twice the Dirac charge.

6.3 1.3 BPS-monopoles

Since the source for a Dirac monopole has to be put in by hand, its mass is a free parameter: it cannot be calculated. On the other hand, for the 't HooftPolyakov monopole there is no source, and the mass of the monopole is an intrinsic property of the Yang-Mills-Higgs system and as such it should be calculable. In the next section we derive a lower bound for its mass. A natural question to ask is whether there are solutions which saturate this bound, and in the section after that such a solution is found: the BPS-monopole.

6.3.1 1.3.1 Estimating the mass of a monopole: the Bogomol'nyi bound

In the centre of mass frame, all the energy of the monopole is concentrated in its mass. Therefore, taking equation (1.13) into account,

$$\begin{aligned} M &= \int_{\mathbb{R}^3} \left(\frac{1}{2} \overrightarrow{\mathrm{E}}_i \cdot \overrightarrow{\mathrm{E}}_i + \frac{1}{2} \overrightarrow{\mathrm{B}}_i \cdot \overrightarrow{\mathrm{B}}_i + \frac{1}{2} \overrightarrow{\mathbf{\Pi}} \cdot \overrightarrow{\mathbf{\Pi}} + \frac{1}{2} D_i \vec{\phi} \cdot D_i \vec{\phi} + V(\phi) \right) \\ &\geq \frac{1}{2} \int_{\mathbb{R}^3} \left(\overrightarrow{\mathrm{E}}_i \cdot \overrightarrow{\mathrm{E}}_i + \overrightarrow{\mathrm{B}}_i \cdot \overrightarrow{\mathrm{B}}_i + D_i \vec{\phi} \cdot D_i \vec{\phi} \right) \end{aligned}$$

where we have dropped some non-negative terms. We now redistribute the last term as follows: we introduce an angular parameter θ and we add and subtract $\overrightarrow{\mathrm{E}}_i \cdot D_i \vec{\phi} \sin\theta$ and $\overrightarrow{\mathrm{B}}_i \cdot D_i \overrightarrow{\boldsymbol{\phi}} \cos\theta$ to the integrand. This yields

$$\begin{aligned} M &\geq \frac{1}{2} \int_{\mathbb{R}^3} \left(\left\| \overrightarrow{\mathrm{E}}_i - D_i \vec{\phi} \sin\theta \right\|^2 + \left\| \overrightarrow{\mathrm{B}}_i - D_i \vec{\phi} \cos\theta \right\|^2 \right) \\ &\quad + \sin\theta \int_{\mathbb{R}^3} D_i \vec{\phi} \cdot \overrightarrow{\mathrm{E}}_i + \cos\theta \int_{\mathbb{R}^3} D_i \overrightarrow{\boldsymbol{\phi}} \cdot \overrightarrow{\mathrm{B}}_i \\ &\geq \sin\theta \int_{\mathbb{R}^3} D_i \overrightarrow{\boldsymbol{\phi}} \cdot \overrightarrow{\mathrm{E}}_i + \cos\theta \int_{\mathbb{R}^3} D_i \overrightarrow{\boldsymbol{\phi}} \cdot \overrightarrow{\mathrm{B}}_i \end{aligned}$$

where we have introduced the obvious shorthand $\|\mathrm{V}_i\|^2 = \mathrm{V}_i \cdot \mathrm{V}_i$. But now notice that

$$\begin{aligned} \int_{\mathbb{R}^3} D_i \overrightarrow{\boldsymbol{\phi}} \cdot \overrightarrow{\mathrm{B}}_i &= \int_{\mathbb{R}^3} \partial_i \left(\overrightarrow{\boldsymbol{\phi}} \cdot \overrightarrow{\mathrm{B}}_i \right) \quad \text{by the Bianchi identity (1.11)} \\ &= \int_{\Sigma_\infty} \overrightarrow{\boldsymbol{\phi}} \cdot \overrightarrow{\mathrm{B}}_i dS_i \quad \text{by Stokes} \\ &= a \int_{\Sigma_\infty} \overrightarrow{\boldsymbol{B}} \cdot d\overrightarrow{\boldsymbol{S}} \equiv ag \end{aligned} \tag{1.24}$$

where Σ_∞ is the sphere are spatial infinity and g is the magnetic charge of the solution. Notice that we have used the results of Exercise 1.8, which are valid since finite-energy demands that the sphere at spatial infinity be in the Higgs vacuum. Similarly, using the equations of motion this time instead of the Bianchi identity, one finds out that

$$\int_{\mathbb{R}^3} D_i \overrightarrow{\boldsymbol{\phi}} \cdot \overrightarrow{\mathrm{E}}_i = a \int_{\Sigma_\infty} \overrightarrow{\boldsymbol{E}} \cdot d\overrightarrow{\boldsymbol{S}} \equiv aq \tag{1.25}$$

where q is the electric charge of the solution. Therefore for all angles θ we have the following bound on the mass:

$$M \geq ag\cos\theta + aq\sin\theta \tag{1.26}$$

The sharpest bound occurs when the right hand side is a maximum, which happens for $q\cos\theta = g\sin\theta$. In other words, $\tan\theta = q/g$. Plugging this back into (1.26), we find the celebrated Bogomol'nyi bound for the mass of a monopole-like solution in terms of the electric and magnetic charges:

$$M \geq a\sqrt{q^2+g^2} \tag{1.27}$$

derived for the first time in [Bog76 (see also [CPNS76]).

For the 't Hooft-Polyakov monopole, which is electrically neutral, the Bogomol'nyi bound yields $M \geq a|g| = 4\pi a/e$. But $a/e = M_W/e^2\lambda$, whence $M \gtrsim M_W/\alpha$, where $\alpha \simeq 1/137$ is the fine structure constant. If $M_W \simeq 90$GeV, say, then $M_W \gtrsim 12$TeV-beyond the present experimental range. This concludes the phenomenological part of these lectures!

6.3.2 1.3.2 Saturating the bound: the BPS-monopole

Having derived the Bogomol'nyi bound, it is natural to ask whether there exist solutions which saturate the bound. We will follow custom and call such states BPS-states. We incurred in the inequalities for the mass by discarding certain terms from the mass formula. To saturate the bound, these terms would have to be equal to zero. Since they are all integrals of non-negative quantities, we must impose that these quantities vanish throughout space and not just asymptotically as the weaker requirement of finite-energy would demand.

Let us concentrate on static solutions which saturate the bound. Static solutions satisfy $\overrightarrow{\mathrm{E}}_i = 0$ and $D_0\overrightarrow{\boldsymbol{\phi}} = 0$. In particular they have no electric charge, so that $\sin\theta = 0$. This means that $\cos\theta = \pm 1$ correlated to the sign of the magnetic charge. A quick inspection at the way we derived the bound reveals that for saturation we must also require that $V(\phi)$ should vanish and that in addition the Bogomol'nyi equation should hold:

$$\overrightarrow{\mathrm{B}}_i = \pm D_i\overrightarrow{\boldsymbol{\phi}} \tag{1.28}$$

Now the only way to satisfy $V(\phi) = 0$ and yet obtain a solution with nonzero magnetic charge, is for λ to vanish. Why? Because for $\lambda \neq 0, \phi^2 = a^2$ throughout space, and in particular, $\vec{\phi}\cdot D_i\vec{\phi} = \vec{\phi}\cdot\partial_i\vec{\phi} = 0$. But using the Bogomol'nyi equation (1.28), this means that $\vec{\phi}\cdot\vec{B}_i = 0$, whence the solution carries no magnetic field. One way to understand the condition $\lambda = 0$ is as a limiting value. We let $\lambda \downarrow 0$, while at the same time retaining the boundary condition that at spatial infinity $\vec{\phi}$ satisfies (1.21). This is known as the Prasad-Sommerfield limit [PS75].

Exercise 1.11 (The Bogomol'nyi equation implies (1.10))

Show that the Bogomol'nyi equation together with the Bianchi identity (1.11) implies the equations of motion (1.10) for the Yang-Mills-Higgs system with $\lambda = 0$.

Of course, the advantage of the Bogomol'nyi equation lies in its simplicity. In fact, it is not hard to find an explicit solution to the Bogomol'nyi equation in the 't Hooft-Polyakov Ansatz, as the next exercise asks you to do.

Exercise 1.12 (The BPS-monopole)

Show that the Bogomol'nyi equation in the 't Hooft-Polyakov Ansatz yields the following systems of equations for the functions H and K:

$$\xi \frac{dK}{d\xi} = -KH$$
$$\xi \frac{dH}{d\xi} = H + 1 - K^2$$

Show that the following is a solution with the right asymptotic boundary conditions:

$$H(\xi) = \xi \coth \xi - 1$$
$$K(\xi) = \frac{\xi}{\sinh \xi}$$

Notice that the solution for the BPS-monopole is such that

$$H(\xi) - \xi = 1 + O(\exp(-\xi))$$

which does not contradict 1.17 because for $\lambda = 0$ the Higgs field is massless. Its interactions are long range and hence the BPS-monopole can be distinguished from a Dirac monopole from afar.

One consequence of the Bogomol'nyi equation is that both the photon (through $\overrightarrow{\mathrm{B}}_i$) and the Higgs (via $D_i\vec{\phi}$) contribute equally to the mass density. One can show that the longrange force exerted by the Higgs is always attractive and for static monopoles, it is equal in magnitude to the $1/r^2$ magnetic force. Therefore the forces add for oppositely charged monopoles, yet they cancel for equally charged monopoles. This is as it should be if static multi-monopole solutions saturating the Bogomol'nyi bound are to exist. To see this, notice that the mass of a two-monopole system with charges g and g' (of the same sign) is precisely equal to the sum of the masses of each of the BPS-monopoles. Hence there can be no net force between them.

Exercise

1.13 (The mass density at the origin is finite)

Show that the mass density at the origin for a BPS-monopole is not merely integrable, but actually finite!

(Hint: Notice that the mass density is given by $\left\| D_i\vec{\phi} \right\|^2$. Compute this for the BPS-monopole and expand as $\xi \sim 0$.)

6.4 1.4 Duality conjectures

In this section we discuss the observed duality symmetries between perturbative and nonperturbative states in the Georgi-Glashow model and the conjectures that this observation suggests. We start with the Montonen-Olive conjecture and then, after introducing a CP-violating term in the theory, the Witten effect will suggest an improved $SL(2, \mathbb{Z})$ duality conjecture.

6.4.1 1.4.1 The Montonen-Olive conjecture

At $\lambda = 0$, the (bosonic) spectrum of the Georgi-Glashow model (including the BPS-monopoles) is the following:

Particle	Mass	Electric Charge 0	Magnetic Charge 0	Spin/ Helicity ± 1
Photon	0	0	0	0
Higgs	0	0	0	1
$W_\pm$boson	aq	$\pm q$	$\pm g$	0

where $q = e\lambda$. Two features are immediately striking:

all particles satisfy the Bogomol'nyi bound; and the spectrum is invariant under electromagnetic $\mathbb{Z}_2$ duality: $(q, g) \mapsto (g, -q)$ provided that we also interchange the BPS-monopoles and the massive vector bosons.

The invariance of the spectrum under electromagnetic duality is a consequence of the fact that the formula for the Bogomol'nyi bound is invariant under electromagnetic duality and the fact that the spectrum saturates the bound. This observation prompted Montonen and Olive MO77 to conjecture that there should be a dual ("magnetic") description of this gauge theory where the elementary gauge particles are the BPS-monopoles and where the massive vector bosons appear as "electric monopoles". This conjecture is reinforced by the fact that two very different calculations for the inter-particle force between the massive vector bosons (done by computing tree diagrams in the quantum field theory) and between the BPS-monopoles (a calculation due to Manton) yield identical answers. Notice, however, that because of the Dirac quantisation condition, if the coupling constant e of the original theory is small, the coupling constant g of the magnetic theory must be large, and
viceversa. Hence the duality conjecture would imply that the strong coupling behaviour of a gauge theory could be determined by the weak coupling behaviour of its dual theory - a very attractive possibility.

The Montonen-Olive conjecture suffers from several drawbacks:

there is no reason to believe that the duality symmetry of the spectrum is not broken by radiative corrections through a renormalisation of the Bogomol'nyi bound; in order to understand the BPS-monopoles as gauge particles, we would expect that their spin be equal to one - yet it would seem naively that due to their rotational symmetry, they have spin zero; and the conjecture is untestable unless we get a better handle at strongly coupled theories - of course, this also means that it cannot be disproved!

We will see in the next chapter that supersymmetry solves the first two problems. The third problem is of course very difficult, but we will now see that by introducing a CP violating term in the action, the duality conjecture will imply a richer dyonic spectrum which can be tested in principle.

6.4.2 1.4.2 The Witten effect

Exercise 1.3 asked you to compute the dyonic spectrum consistent with the quantisation condition (1.7) in a CP non-violating theory. You should have found that the electric charge q of a dyon with minimal magnetic charge g could take one of two sets of mutually exclusive values: either $q =$ ne or $q = ne + \frac{1}{2}e$, where n is some integer. We will see that indeed it is the former case which holds.

Let N denote the operator which generates gauge transformations about the direction $\vec{\phi}$:

$$\begin{aligned} \delta \overrightarrow{\boldsymbol{v}} &= \frac{1}{a} \overrightarrow{\boldsymbol{\phi}} \times \overrightarrow{\boldsymbol{v}} \\ \delta \overrightarrow{\mathrm{W}}_\mu &= -\frac{1}{ea} D_\mu \overrightarrow{\boldsymbol{\phi}} \end{aligned} \tag{1.29}$$

where $\overrightarrow{\boldsymbol{v}}$ is any isovector. And consider the operator $\exp 2\pi i N$. In the background of a finite energy solution, the Higgs field is in the Higgs vacuum at spatial infinity, whence

*The formula for the Bogomol'nyi bound is actually invariant under rotations in the (g, q) plane. But the quantisation of the electric and magnetic charges, actually breaks this symmetry down to the $\mathbb{Z}_2$ duality symmetry. In their paper, Montonen and Olive speculate that the massless Higgs could play the role of a Goldstone boson associated to the breaking of this $SO(2)$ symmetry down to $\mathbb{Z}_2$. I am not aware of any further progress in this direction.

$\exp 2\pi i N$ generates the identity transformation. On isovectors it generates a rotation about $\vec{\phi}$ of magnitude $2\pi|\vec{\phi}|/a = 2\pi$, and on the gauge fields we notice that $D_\mu\vec{\phi} = 0$ in the Higgs vacuum. Since $\exp 2\pi i N = \mathbb{1}$, the eigenvalues of N are integral. To see what this means, we compute N.

We can compute N since it is the charge of the Noether current associated with the transformations (1.29). Indeed,

$$N = \int_{\mathbb{R}^3} \left(\frac{\partial \mathcal{L}}{\partial \partial_0 \overrightarrow{\mathrm{W}}_\mu} \cdot \delta \overrightarrow{\mathrm{W}}_\mu + \frac{\partial \mathcal{L}}{\partial \partial_0 \vec{\phi}} \cdot \delta \overrightarrow{\boldsymbol{\phi}} \right)$$

Using equation (1.29), and hence that $\delta \overrightarrow{\boldsymbol{\phi}} = 0$, we can rewrite N as

$$N = -\frac{1}{ae} \int_{\mathbb{R}^3} \frac{\partial \mathcal{L}}{\partial \partial_0 \overrightarrow{\mathrm{W}}_i} \cdot D_i \overrightarrow{\boldsymbol{\phi}}$$

Since the conjugate momentum to $\overrightarrow{\mathrm{W}}_i$ is $-\overrightarrow{\mathrm{G}}^{0i} = \overrightarrow{\mathrm{E}}^i = -\overrightarrow{\mathrm{E}}_i$, we find that

$$N = \frac{1}{ae} \int_{\mathbb{R}^3} \overrightarrow{\mathrm{E}}_i \cdot D_i \vec{\phi} = \frac{q}{e} \tag{1.30}$$

where we have used the expression (1.25) for the electric charge q of the configuration. The quantisation of N then implies that $q = ne$ for some integer n.

Let us now introduce a θ-term in the action:

$$\mathcal{L}_\theta = \frac{1}{2} \frac{e^2\theta}{32\pi^2} \epsilon^{\alpha\beta\mu\nu} \overrightarrow{\mathrm{G}}_{\alpha\beta} \cdot \overrightarrow{\mathrm{G}}_{\mu\nu} = -\frac{e^2\theta}{32\pi^2} \times \overrightarrow{\mathrm{G}}^{\mu\nu} \cdot \overrightarrow{\mathrm{G}}_{\mu\nu}$$

This term is locally a total derivative and hence does not contribute to the equations of motion. Its integral in a given configuration is an integral multiple (called the instanton number) of the parameter θ. θ is therefore an angular variable and parametrises inequivalent vacua. The Noether charge N gets modified in the presence of this term as follows:

$$N \mapsto N - \frac{1}{ae} \int_{\mathbb{R}^3} \frac{\partial \mathcal{L}_\theta}{\partial \partial_0 \vec{W}_i} \cdot D_i \vec{\phi}$$

Computing this we find

$$\begin{aligned} \Delta N &= -\frac{e\theta}{16\pi^2 a} \int_{\mathbb{R}^3} \epsilon^{0i\alpha\beta} \overrightarrow{\mathrm{G}}_{\alpha\beta} \cdot D_i \vec{\phi} \\ &= -\frac{e\theta}{16\pi^2 a} \int_{\mathbb{R}^3} \epsilon^{ijk} \overrightarrow{\mathrm{G}}_{jk} \cdot D_i \vec{\phi} \\ &= \frac{e\theta}{8\pi^2 a} \int_{\mathbb{R}^3} \overrightarrow{\mathrm{B}}_i \cdot D_i \vec{\phi} \\ &= \frac{e\theta}{8\pi^2} g \end{aligned}$$

where g, given by equation (1.24), is the magnetic charge of the configuration. In other words,

$$N = \frac{q}{e} + \frac{e\theta}{8\pi^2} g$$

For the 't Hooft-Polyakov monopole, eg $= -4\pi$, hence the integrality of N means that

$$q = ne + \frac{e\theta}{2\pi} \quad \text{for some} \quad n \in \mathbb{Z} \tag{1.31}$$

This result, which was first obtained by Witten in [Wit79], is of course consistent with the quantisation condition (1.7) since for a fixed θ the difference between any charges is an integral multiple of e.

6.4.3 1.4.3 $SL(2,\mathbb{Z})$ duality

The action defined by $\mathcal{L}+\mathcal{L}_\theta$ depends on four parameters: e,θ,λ and a. The dependence on the first two can be unified into a complex parameter τ. To see this, let us first rescale the gauge fields $\vec{\mathrm{W}}_\mu \mapsto e\vec{\mathrm{W}}_\mu$. This has the effect of bringing out into the open all the dependence on e. The lagrangian is now

$$\mathcal{L}+\mathcal{L}_\theta = -\frac{1}{4e^2}\vec{\mathrm{G}}_{\mu\nu}\cdot\vec{\mathrm{G}}^{\mu\nu} + \frac{\theta}{32\pi^2}\vec{\mathrm{G}}_{\mu\nu}\cdot{}^*\vec{\mathrm{G}}^{\mu\nu} + \frac{1}{2}D^\mu\vec{\phi}\cdot D_\mu\vec{\phi} - V(\phi) \tag{1.32}$$

where all the (e,θ)-dependence is now shown explicitly. We now define a complex parameter

$$\tau \equiv \frac{\theta}{2\pi} + i\frac{4\pi}{e^2}$$

whose imaginary part is positive since e is real. To write the lagrangian explicitly in terms of τ it is convenient to introduce the following complex linear combination:

$$\vec{\mathcal{G}}_{\mu\nu} \equiv \vec{\mathrm{G}}_{\mu\nu} + i^\star\vec{\mathrm{G}}_{\mu\nu} \tag{1.33}$$

It then follows that

$$\vec{\mathcal{S}}_{\mu\nu}\cdot\vec{\mathcal{G}}^{\mu\nu} = 2\vec{\mathrm{G}}_{\mu\nu}\cdot\vec{\mathrm{G}}^{\mu\nu} + 2i\vec{\mathrm{G}}_{\mu\nu}\cdot\times\vec{\mathrm{G}}^{\mu\nu}$$

whence the first two terms in the lagrangian (1.32) can be written simply as

$$-\frac{1}{32\pi}\,\mathrm{Im}\left(\tau\vec{\mathfrak{S}}_{\mu\nu}\cdot\vec{\mathfrak{G}}^{\mu\nu}\right) \tag{1.34}$$

Notice that because θ is an angular variable, it is only defined up to 2π. This means that physics is invariant under $\tau\mapsto\tau+1$. At $\theta=0$, the conjecture of electromagnetic duality says that $e\mapsto g=-4\pi/e$ is a symmetry. But this duality transformation is just $\tau\mapsto -1/\tau$. We are therefore tempted to strengthen the conjecture of electromagnetic duality to say that for arbitrary θ, the physics should depend on τ only modulo the transformations:

$$T:\tau\mapsto\tau+1$$
$$S:\tau\mapsto-\frac{1}{\tau}$$

Exercise

1.14((P)$SL(2,\mathbb{Z})$ and its action on the upper half-plane)

The group $SL(2,\mathbb{Z})$ of all 2×2 matrices with unit determinant and with integer entries acts naturally on the complex plane:

$$\begin{pmatrix} a & b \\ c & d\end{pmatrix}\cdot\tau = \frac{a\tau+b}{c\tau+d}$$

Prove that this action preserves the upper half-plane, so that if $\mathrm{Im}\,\tau>0$, so will its transform under $SL(2,\mathbb{Z})$. Prove that the matrices $\mathbb{1}$ and $-\mathbb{1}$ both act trivially (and are the only two matrices that do). Thus the action is not faithful, but it becomes faithful if we identify every matrix $M\in SL(2,\mathbb{Z})$ with $-M$. The resulting group is denoted $PSL(2,\mathbb{Z})\equiv SL(2,\mathbb{Z})/\{\pm\mathbb{1}\}$.

The operations S and T defined above are clearly invertible and hence generate a discrete group. Prove that they satisfy the following relations:

$$S^2 = \mathbb{1} \quad \text{and} \quad (ST)^3 = \mathbb{1}$$

Prove that the group generated by S and T subject to the above relations is a subgroup of $PSL(2,\mathbb{Z})$, by exhibiting matrices $\hat{S}$ and $\hat{T}$ whose action on τ coincides with the action of S and T. These matrices are not unique, since in going from $PSL(2,\mathbb{Z})$ to $SL(2,\mathbb{Z})$ we have to choose a sign. Nevertheless, for any choice of $\hat{S}$ and $\hat{T}$, prove that the following matrix identities are satisfied:

$$\hat{S}^2 = -\mathbb{1} \quad \text{and} \quad (\hat{S}\hat{T})^6 = \mathbb{1}$$

The matrices $\hat{S}$ and $\hat{T}$ thus generate a subgroup of $SL(2,\mathbb{Z})$. Prove that this subgroup is in fact the whole group, which implies that S and T generate all of $PSL(2,\mathbb{Z})$.

(Hint: if you get stuck look in Ser73].)

Is physics invariant under $SL(2,\mathbb{Z})$? Clearly this would be a bold conjecture, but no bolder than the original Montonen-Olive $\mathbb{Z}_2$ conjecture, for in fact the evidence for both is more or less the same. Indeed, as we now show the mass formula for BPS-states is invariant under $SL(2,\mathbb{Z})$. The mass of a BPS-state with charges (q, g) is given by the equality in formula (1.27). From formula (1.23) it follows that the allowed magnetic charges of the form $g = n_m 4\pi/e$, for some $n_m \in \mathbb{Z}$. As a consequence of the Witten effect, the allowed electric charges are given by $q = n_e e + n_m e\theta/2\pi$. The mass of BPS-states is then given by

$$M^2 = 4\pi a^2 \overrightarrow{\boldsymbol{n}}^t \cdot A(\tau) \cdot \overrightarrow{\boldsymbol{n}} \tag{1.35}$$

where $\overrightarrow{\boldsymbol{n}} = (n_e, n_m)^t \in \mathbb{Z} \times \mathbb{Z}$ and where

$$A(\tau) = \frac{1}{\operatorname{Im}\tau} \begin{pmatrix} 1 & \operatorname{Re}\tau \\ \operatorname{Re}\tau & |\tau|^2 \end{pmatrix}$$

Exercise 1.15 ($SL(2,\mathbb{Z})$-invariance of the mass formula)

Let

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{Z})$$

First prove that

$$A(M \cdot \tau) = \left(M^{-1}\right)^t \cdot A(\tau) \cdot M^{-1}$$

and as a consequence deduce that the mass formula is invariant provided that we also transform the charges:

$$\vec{n} \mapsto M \cdot \vec{n}$$

The improved Montonen-Olive conjecture states that physics is $SL(2,\mathbb{Z})$ invariant. If this is true, this means that the theories defined by two values of τ related by the action of $SL(2,\mathbb{Z})$ are physically equivalent, provided that we are willing to relabel magnetic and electric charges by that same $SL(2,\mathbb{Z})$ transformation.

The action of $PSL(2,\mathbb{Z})$ on the upper half-plane is well-known (see for example Serre's book [Ser73]). There is a fundamental domain D defined by

$$D = \left\{ \tau \in \mathbb{C} | \operatorname{Im}\tau > 0, |\operatorname{Re}\tau \Big| \leq \frac{1}{2}, \Big|\tau\Big| \geq 1 \right\} \tag{1.36}$$

which has the property that its orbit under $PSL(2,\mathbb{Z})$ span the whole upper half-plane and that no two points in its interior

$$\operatorname{Int} D = \left\{ \tau \in \mathbb{C} | \operatorname{Im}\tau > 0, |\operatorname{Re}\tau \Big| < \frac{1}{2}, \Big|\tau\Big| > 1 \right\}$$

are related by the action of $PSL(2,\mathbb{Z})$.

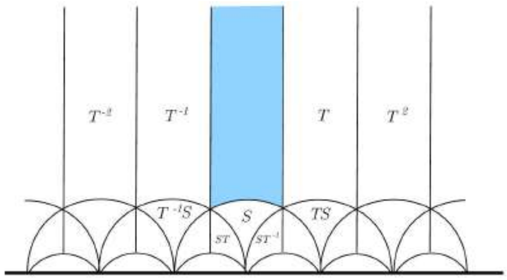


Figure 1.1: Fundamental domain (shaded) for the action of $PSL(2,\mathbb{Z})$ on the upper half plane, and some of its $PSL(2,\mathbb{Z})$ images.

Exercise 1.16 (Orbifold points in the fundamental domain D**)**

The fundamental domain D contains three "orbifold" points: $i, \omega = \exp(i\pi/3)$ and $-\bar{\omega} = \exp(2i\pi/3)$ which are fixed by some finite subgroup of $PSL(2,\mathbb{Z})$. Indeed, prove that i is fixed by the $\mathbb{Z}_2$-subgroup generated by S, whereas ω and $-\bar{\omega}$ are fixed respectively by the $\mathbb{Z}_3$-subgroups generated by TS and ST.

We end this section and this chapter with a discussion of the dyonic spectrum predicted by $SL(2,\mathbb{Z})$-duality. If we had believed in the electromagnetic $\mathbb{Z}_2$-duality, we could have predicted the existence of the BPS-monopoles from the knowledge of the existence of the massive vector bosons (and viceversa). But this is as far as we could have gone with $\mathbb{Z}_2$. On the other hand $SL(2,\mathbb{Z})$ has infinite order, and assuming that for all values of τ there are massive vector bosons in the spectrum, $SL(2,\mathbb{Z})$-duality predicts an infinite number of dyonic states. This assumption is not as innocent as it seems, as the Seiberg-Witten solution to pure $N = 2$ supersymmetric Yang-Mills demonstrates; but it seems to hold if we have $N = 4$ supersymmetry. But for now let us simply follow our noses and see what this assumption implies.

Let's assume then that for all values of τ there is a state with quantum numbers $\vec{n} = (1,0)^t$. The duality conjecture predicts the existence of one
state each with quantum numbers in the $SL(2,\mathbb{Z})$-orbit of $\vec{n}$:

$$M \cdot \vec{n} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix}$$

Because M has unit determinant, a and c are not arbitrary integers: there exist integers b and d such that $ad - bc = 1$. This means that a and c are coprime; that is, they don't have a common factor (other than 1). Indeed, if n were a common factor: $a = na'$ and $c = nc'$ for integers a' and c', and we would have that $n\,(a'd - bc') = 1$ which forces $n = 1$. We will now show that this arithmetic property of a and c actually translates into the stability of the associated dyonic state!

Exercise 1.17 (Properties of the mass matrix $A(\tau)$)

Notice that the matrix $A(\tau)$ in the mass formula (1.35) enjoys the following properties for all τ in the upper half-plane:

$$\det A(\tau) = 1 \quad \text{and} \quad A(\tau) \text{ is positive-definite.}$$

Prove that this latter property implies that the mass formula defines a distance function, so that in particular it obeys the triangle inequality. In other words, if we define $\|\vec{\boldsymbol{n}}\|^2 \equiv M^2_{\vec{\boldsymbol{n}}}$-that is, the Bogomol'nyi mass of a dyonic state with that charge assignment - then prove that

$$\|\vec{n} + \vec{m}\| \leq \|\vec{n}\| + \|\vec{m}\| \tag{1.37}$$

Now let's consider a dyonic state $\vec{\boldsymbol{q}} = (a,c)^t$. The triangle inequality (1.37) says that for any two dyonic states $\vec{\boldsymbol{n}}$ and $\vec{\boldsymbol{m}}$ which obey $\vec{\boldsymbol{n}} + \vec{\boldsymbol{m}} = \vec{\boldsymbol{q}}$, the mass of the $\vec{\boldsymbol{q}}$ is less than or equal to the sum of the masses of $\vec{\boldsymbol{n}}$ and $\vec{\boldsymbol{m}}$. But we claim that when a and c are coprime, the inequality is actually strict! Indeed, the inequality is only saturated when $\vec{\boldsymbol{n}}$ and $\vec{\boldsymbol{m}}$, and hence $\vec{\boldsymbol{q}}$, are collinear. But if this is the case, a and c must have a common factor. Assume for a contradiction that they don't. If $\vec{\boldsymbol{n}} = (p,q)^t$ and $\vec{\boldsymbol{m}} = (r,s)^t$, then we must have that both p and r are smaller in magnitude to a, and that q and s are smaller in magnitude to c. But collinearity means that $pc = qa$. Since a and c are relatively prime, it must be that a divides p so that there is some integer n such that $p = an$, which contradicts that p is smaller in magnitude to a. This also follows pictorially from the fact that a and c are coprime if and only if in the straight line from the origin to $\vec{\boldsymbol{q}} \in \mathbb{Z}^2 \subset \mathbb{R}^2$, $\vec{\boldsymbol{q}}$ is the first integral point. Therefore the dyonic state represented by $\vec{\boldsymbol{q}}$ is is a genuine stable state which cannot be interpreted as a bound state of other dyonic states with "smaller" charges.

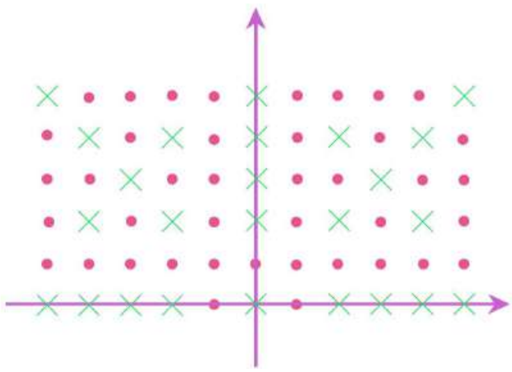


Figure 1.2: Dyonic spectrum predicted by $SL(2,\mathbb{Z})$ duality. Dots indicate dyons, crosses indicates holes in the dyonic spectrum. Only dyons with nonnegative magnetic charge are shown.

The dyonic states in the $SL(2,\mathbb{Z})$-orbit of $(1,0)$ can be depicted as follows:

Notice that for $c = 0$ we have the original state and its charge conjugate. For $c = 1$ we have the Julia-Zee dyons but with quantised electric charge: a can be an arbitrary integer. For $c = 2$, we have that a must be odd. Notice that in every rational direction (that is, every half-line with rational slope emanating from the origin) only the first integral point is present. As explained above these are precisely those points (m,n) whose coordinates are coprime.

As we will see in the context of $N = 4$ supersymmetric Yang-Mills theory, the dyonic spectrum is in one-to-one correspondence with square-integrable harmonic forms on monopole moduli space. This is a fascinating prediction: it says that there is an action of the modular group on the (L^2) cohomology of monopole moduli space.

*Contrast this with the case where the triangle inequality saturates. In this case, this means that the bound state of two dyons with "collinear" charges exhibits no net force between its constituents. Compare this with the discussion in section 1.3.2

Chapter 2

7 Special geometry and symplectic transformations by de Wit, Van Proeyen

Special Kähler manifolds are defined by coupling of vector multiplets to $N = 2$ supergravity. The coupling in rigid supersymmetry exhibits similar features. These models contain n vectors in rigid supersymmetry and $n + 1$ in supergravity, and n complex scalars. Apart from exceptional cases they are defined by a holomorphic function of the scalars. For supergravity this function is homogeneous of second degree in an $(n+1)$-dimensional projective space. Another formulation exists which does not start from this function, but from a symplectic $(2n)$- or $(2n+2)$-dimensional complex space. Symplectic transformations lead either to isometries on the manifold or to symplectic reparametrizations. Finally we touch on the connection with special quaternionic and very special real manifolds, and the classification of homogeneous special manifolds.

---

† Based on invited lectures delivered by AVP at the Spring workshop on String theory, Trieste, April 1994; to be published in the proceedings.

* Onderzoeksleider, N.F.W.O., Belgium

7.0.1 Introduction

In nonlinear sigma models, the spinless fields define a map from the d-dimensional Minkowskian space-time to some 'target space', whose metric is given by the kinetic terms of these scalars. Supersymmetry severely restricts the possible target-space geometries. The type of target space which one can obtain depends on d and on N, the latter indicating the number of independent supersymmetry transformations. The number of supersymmetry generators ('supercharges') is thus equal to N times the dimension of the (smallest) spinor representation. For realistic supergravity this number of supercharge components cannot exceed 32. As 32 is the number of components of a Lorentz spinor in $d = 11$ space-time dimensions, it follows that realistic supergravity theories can only exist for dimensions $d \leq 11$. For the physical $d = 4$ dimensional space-time, one can have supergravity theories with $1 \leq N \leq 8$.

Table 7: Restrictions on target-space manifolds according to the type of supergravity theory. The rows are arranged such that the number κ of supercharge components is constant. $\mathcal{M}$ refers to a general Riemannian manifold, SK to 'special Kähler', VSR to 'very special real' and Q to quaternionic manifolds.

κ	$d=2$	$d=3$	$d=4$	$d=5$	$d=6$
2	$N=1$ $\mathcal{M}$				
4	$N=2$ Kähler	$N=2$ Kähler	$N=1$ Kähler		
8	$N=4$ Q	$N=4$ Q	$N=2$ $SK \oplus Q$	$N=2$ $VSR \oplus Q$	$N=1$ $\emptyset \oplus Q$
16	... ...	... ...	$N=4$ $\frac{SO(6,n)}{SO(6)\otimes SO(n)} \otimes \frac{SU(1,1)}{U(1)}$	... ...	$\rightarrow$ $d=10$
32	... ...	... ...	$N=8$ $\frac{E_7}{SU(8)}$	... ...	$\rightarrow$ $d=11$

As clearly exhibited in table 7, the more supercharge components one has, the more restrictions one finds. When the number of supercharge components exceeds 8, the target spaces are restricted to symmetric spaces. For $\kappa = 16$ components, they are specified by an integer n, which specifies the number of vector multiplets. This row continues to $N = 1$, $d = 10$. Beyond 16 supercharge components there is no freedom left. The row with 32 supercharge components continues to $N = 1$, $d = 11$. Here we treat the case of 8 supercharge components. This is the highest value of N where the target space is not yet restricted to be a symmetric space, although supersymmetry has already fixed a lot of its structure. We will mostly be concerned with $N = 2$ in $d = 4$ dimensions. The target space factorizes into a quaternionic and a Kähler manifold of a particular type [1], called *special* [2]. The former contains the scalars of the hypermultiplets (multiplets without vectors). The latter contain the scalars in vector multiplets. Recently the special Kähler structure received a lot of attention, because it plays an important role in string compactifications. Also quaternionic manifolds appear in this context, and also here it is a restricted class of special quaternionic manifolds that is relevant. In lowest order of the string coupling constant these manifolds are even 'very special' Kähler and quaternionic, a notion that we will define below.

In the next section we describe the actions of $N = 2$ vector multiplets. First we consider rigid supersymmetry. We explain the fields in the multiplets, their description in superspace and how this leads to a holomorphic prepotential. Then we exhibit how the structure becomes more complicated in supergravity, where the space of physical scalars is embedded in a projective space. This became apparent by starting from the superconformal tensor calculus.

In section 7.0.3 we discuss the symplectic transformations, which play an important role in the recent developments of weak–strong coupling dualities. First we repeat the general idea (and elucidate it for S and T dualities), and then show what is the extra structure in $N = 2$ theories. There are two kind of applications, either as isometries of the manifolds (symmetries of the theory), or as equivalence relations of prepotentials (pseudo-symmetries). We illustrate both with explicit examples. These will also exhibit formulations without a prepotential, showing the need for a formulation that does not rely on the existence of a prepotential. This formulation is given at the end of the section. Some further results will be mentioned in section 7.0.4.

In all of this we confine ourselves to special geometry from a supersymmetry/supergravity perspective. The connection with the geometry of the moduli of Calabi-Yau spaces [2, 3, 4, 5, 6, 7] is treated in the lectures of Pietro Frè [8].

7.0.2 $N = 2$ actions

Table 8: Physical fields in $N = 2$, $d = 4$ actions

spin	pure SG	n vector m.	s hyperm.
2	1		
3/2	2		
1	1	n	
1/2		2n	2s
0		2n	4s

We briefly introduce special Kähler manifolds in the context of $N = 2$, $d = 4$ supergravity. As exhibited in table ??, the physical multiplets of supersymmetry are vector and hypermultiplets, which can be coupled to supergravity. In this section we will not consider the hypermultiplets. The scalar sector of the $N = 2$ supergravity-Yang-Mills theory in four spacetime dimensions defines the 'special Kähler manifolds'. Without supergravity we have $N = 2$ supersymmetric Yang-Mills theory, which we will treat first. The spinless fields parametrize then a similar type of Kähler manifolds. The vector potentials, which describe the spin-1 particles, are accompanied by complex scalar fields and doublets of spinor fields, all taking values in the Lie algebra associated with the group that can be gauged by the vectors. In the second subsection we will see what the consequences are of mixing the vectors in the vector multiplets with the one in the supergravity multiplet.

Rigid supersymmetry

The superspace contains the anticommuting coordinates θ^i_α and $\bar\theta_{\dot\alpha i}$ where $i = 1, 2$ and $\alpha, \dot\alpha$ are the spinor indices. The simplest superfields are, as in $N = 1$, the chiral superfields. They are defined by a constraint $\bar D^{\dot\alpha i}\Phi = 0$, where $\bar D^{\dot\alpha}$ is a covariant chiral superspace derivative, and Φ is a complex superfield. This constraint determines its structure*:

$$\begin{aligned}\Phi = {} & X + \theta^i_\alpha \lambda^\alpha_i \\ & + \epsilon_{ij}\theta^i_\alpha \sigma^{\alpha\beta}_{\mu\nu}\theta^j_\beta \mathcal{F}^{+\mu\nu} + \epsilon_{\alpha\beta}\theta^i_\alpha\theta^j_\beta Y_{ij} + \ldots \ ,\end{aligned} \tag{7.1}$$

where ... stands for terms cubic or higher in θ. New component fields can appear up to θ^4, leading to $8 + 8$ complex field components. All these fields do not form an irreducible representation of supersymmetry, but can be split into two sets of $8+8$ real fields transforming irreducibly. We restrict ourselves to the set containing the fields already exhibited in (7.1),

*We use $\mathcal{F}^\pm_{\mu\nu} = \frac{1}{2}\left(\mathcal{F}_{\mu\nu} \pm \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\mathcal{F}^{\rho\sigma}\right)$ with $\epsilon^{0123} = i$.

which leads to the vector multiplet (The others form a 'linear multiplet'). The reduction is accomplished by the additional constraint

$$D^{\alpha}_{(i}D^{\beta}_{j)}\Phi\,\epsilon_{\alpha\beta} = \epsilon_{ik}\epsilon_{j\ell}\bar{D}^{\dot{\alpha}(k}\bar{D}^{\ell)\dot{\beta}}\bar{\Phi}\,\epsilon_{\dot{\alpha}\dot{\beta}}\ , \tag{7.2}$$

which for instance implies that the symmetric tensor Y_{ij} satisfies a reality constraint: $Y_{ij} = \epsilon_{ik}\epsilon_{j\ell}\bar{Y}^{k\ell}$, so that it consists of only 3 real scalar fields. But more importantly, we also obtain a constraint on the antisymmetric tensor: $\partial^{\mu}\left(\mathcal{F}^{+}_{\mu\nu} - \mathcal{F}^{-}_{\mu\nu}\right) = 0$, which is the Bianchi identity, which implies that $\mathcal{F}$ is the field strength of a vector potential. All the terms $\ldots$ in (7.1) are determined in terms of the fields written down. Therefore the independent components of the vector multiplet are: $X^A, \lambda^{iA}, \mathcal{F}^{A}_{\mu\nu}, Y^{A}_{ij}$ (where $A = 1, ..., n$ denotes the possibility to include several multiplets). X^A and λ^{iA} will describe the physical scalars and spinors, $\mathcal{F}^A$ are the fields strengths of the vectors and Y^A will be auxiliary scalars in the actions which we will construct.

As we have a chiral superfield, an action can be obtained by integrating an arbitrary holomorphic function $F(\Phi)$ over chiral superspace. The action

$$\int d^4x \int d^4\theta\ F(\Phi)\ + c.c. \tag{7.3}$$

leads to the Lagrangian

$$\begin{aligned} \mathcal{L} = {} & g_{A\bar{B}}\partial_{\mu}X^{A}\partial_{\mu}\bar{X}^{B} + g_{A\bar{B}}\bar{\lambda}^{iA}\ \not{\partial}\lambda^{\bar{B}}_{i} + \\ & + \mathrm{Im}\ (F_{AB}\mathcal{F}^{-A}_{\mu\nu}\mathcal{F}^{-B}_{\mu\nu}) + \mathcal{L}_{\mathrm{Pauli}} + \mathcal{L}_{4-\mathrm{fermi}} \end{aligned} \tag{7.4}$$

where the latter two terms are the couplings of the vector fields to the spinors and the terms quartic in fermions, which we do not write explicitly here. The metric in target space is Kählerian: [9]

$$\begin{aligned} g_{A\bar{B}}(X, \bar{X}) &= \partial_A\partial_{\bar{B}}K(X, \bar{X}) \\ K(X, \bar{X}) &= i(\bar{F}_A(\bar{X})X^A - F_A(X)\bar{X}^A) \\ F_A(X) &= \partial_A F(X)\ ;\ \bar{F}_A(\bar{X}) = \partial_{\bar{A}}\bar{F}(\bar{X}). \end{aligned} \tag{7.5}$$

For $N = 1$ the Kähler potential could have been arbitrary. The presence of two independent supersymmetries implies that this Kähler metric, and even the complete action, depends on a holomorphic prepotential $F(X)$, where X denotes the complex scalar fields. Two different functions $F(X)$ may correspond to equivalent equations of motion and to the same geometry. From the equation* $g_{A\bar{B}} = 2\,\mathrm{Im}\ F_{AB}$, it follows that

$$F \approx F + a + q_A X^A + c_{AB}X^A X^B\ , \tag{7.6}$$

where a and q_A are complex numbers, and c_{AB} real†. But more relations can be derived from the symplectic transformations that we discuss shortly.

The fact that the metric is Kählerian implies that only curvature components with two holomorphic and two anti–holomorphic indices can be non–zero. In this case, these are determined by the third derivative of F:

$$R_{A\bar{B}C\bar{D}} = -F_{ACE}\,g^{E\bar{F}}\,\bar{F}_{\bar{F}\bar{B}\bar{D}}. \tag{7.7}$$

*Here and henceforth we use the convention where $F_{AB\ldots}$ denote multiple derivatives with respect to X of the holomorphic prepotential.

†In supergravity, or in the full quantum theory the q_A must be zero.

Vector multiplets coupled to supergravity

The general action for vector multiplets coupled to $N = 2$ supergravity was first derived using superconformal tensor calculus [1]. In that approach one starts from the $N = 2$ superconformal group, which is

$$SU(2, 2|N = 2) \supset SU(2, 2) \otimes U(1) \otimes SU(2). \tag{7.8}$$

The bosonic subgroup, which we exhibited, contains, apart from the conformal group in $d = 4$, also $U(1)$ and $SU(2)$ factors. The Kählerian nature of vector multiplet couplings and the quaternionic nature of hypermultiplet couplings is directly related to the presence of these two groups. The superconformal group is, however, mainly a useful tool for constructing actions which have just super-Poincaré invariance (see the reviews [10]). To make that transition, the dilatations, special conformal transformations and $U(1)\otimes SU(2)$ are broken by an explicit gauge fixing. The same applies to some extra S–supersymmetry in the fermionic sector.

To describe theories as exhibited in table ??, the following multiplets are introduced: (other possibilities, leading to equivalent physical theories, also exist, see [11, 10]). The *Weyl multiplet* contains the vierbein, the two gravitinos, and auxiliary fields. We introduce $n + 1$ *vector multiplets* :

$$\left(X^I, \lambda^{iI}, \mathcal{A}^I_\mu\right) \qquad \text{with} \qquad I = 0, 1, ..., n. \tag{7.9}$$

The extra vector multiplet labelled by $I = 0$ contains the scalar fields which are to be gauge–fixed in order to break dilations and the $U(1)$, the fermion to break the S–supersymmetry, and the vector which corresponds to the physical vector of the supergravity multiplet in table ??. Finally, there are $s + 1$ *hypermultiplets*, one of these contains only auxiliary fields and fields used for the gauge fixing of $SU(2)$. For most of this paper we will not discuss hypermultiplets ($s = 0$).

Under dilatations the scalars X^I transform with weight 1. On the other hand an action similar to (7.3) can only be constructed if $F(X)$ has Weyl weight 2. This leads to the important conclusion that for the coupling of vector multiplets to supergravity, one again starts from a holomorphic prepotential $F(X)$, this time of $n + 1$ complex fields, but now it must be a *homogeneous* function of degree two [1].

In the resulting action appears $-\frac{1}{2}i(\bar{X}^I F_I - X^I \bar{F}_I)eR$, where R is the space–time curvature. To have the canonical kinetic terms for the graviton, it is therefore convenient to impose as gauge fixing for dilatations the condition

$$i(\bar{X}^I F_I - \bar{F}_I X^I) = 1\,. \tag{7.10}$$

Therefore, the physical scalar fields parametrize an n-dimensional complex hypersurface, defined by the condition (7.10), while the overall phase of the X^I is irrelevant in view of a local (chiral) invariance. The embedding of this hypersurface can be described in terms of n complex coordinates z^A by letting X^I be proportional to some holomorphic sections $Z^I(z)$ of the projective space $P\mathbb{C}^{n+1}$ [12]. The bosonic part of the resulting action is (without gauging)

$$\begin{aligned} e^{-1}\mathcal{L} = & -\tfrac{1}{2}R + g_{\alpha\bar{\beta}}\partial_\mu z^\alpha \partial^\mu \bar{z}^{\bar{\beta}} \\ & -\text{Im}\ \left(\mathcal{N}_{IJ}(z,\bar{z})\mathcal{F}^{+I}_{\mu\nu}\mathcal{F}^{+J}_{\mu\nu}\right). \end{aligned} \tag{7.11}$$

The n-dimensional space parametrized by the z^α ($\alpha = 1, \ldots, n$) is a Kähler space; the Kähler metric $g_{\alpha\bar{\beta}} = \partial_\alpha \partial_{\bar{\beta}} K(z, \bar{z})$ follows from the Kähler potential

$$\begin{aligned} & e^{-K(z,\bar{z})} = i\bar{Z}^I(\bar{z})\, F_I(Z(z)) - iZ^I(z)\, \bar{F}_I(\bar{Z}(\bar{z})) \\ & X^I = e^{K/2} Z^I(z)\ , \qquad \bar{X}^I = e^{K/2}\bar{Z}^I(\bar{z}). \end{aligned} \tag{7.12}$$

The resulting geometry is known as *special* Kähler geometry [1, 2]. The curvature tensor associated with this Kähler space satisfies the characteristic relation [13]

$$R^{\alpha}{}_{\beta\gamma}{}^{\delta} = \delta^{\alpha}_{\beta}\delta^{\delta}_{\gamma} + \delta^{\alpha}_{\gamma}\delta^{\delta}_{\beta} - e^{2K}\mathcal{W}_{\beta\gamma\epsilon}\bar{\mathcal{W}}^{\epsilon\alpha\delta}\,, \tag{7.13}$$

where

$$\mathcal{W}_{\alpha\beta\gamma} = iF_{IJK}\big(Z(z)\big)\,\frac{\partial Z^I}{\partial z^\alpha}\frac{\partial Z^J}{\partial z^\beta}\frac{\partial Z^K}{\partial z^\gamma}\,. \tag{7.14}$$

A convenient choice of inhomogeneous coordinates z^α are the *special* coordinates, defined by

$$z^A = X^A/X^0, \qquad A = 1,\ldots,n, \tag{7.15}$$

or, equivalently,

$$Z^0(z) = 1\,, \qquad Z^A(z) = z^A\,. \tag{7.16}$$

The kinetic terms of the spin-1 gauge fields in the action are proportional to the symmetric tensor

$$\mathcal{N}_{IJ} = \bar{F}_{IJ} + 2i\frac{\mathrm{Im}(F_{IK})\,\mathrm{Im}(F_{JL})\,X^K X^L}{\mathrm{Im}(F_{KL})\,X^K X^L}\,. \tag{7.17}$$

This tensor describes the field-dependent generalization of the inverse coupling constants and so-called θ parameters.

We give here some examples of functions $F(X)$ and their corresponding target spaces, which will be useful later on:

$$F = -i\,X^0X^1 \qquad \frac{SU(1,1)}{U(1)} \tag{7.18}$$

$$F = (X^1)^3/X^0 \qquad \frac{SU(1,1)}{U(1)} \tag{7.19}$$

$$F = -4\sqrt{X^0(X^1)^3} \qquad \frac{SU(1,1)}{U(1)} \tag{7.20}$$

$$F = iX^I\eta_{IJ}X^J \qquad \frac{SU(1,n)}{SU(n)\otimes U(1)} \tag{7.21}$$

$$F = \frac{d_{ABC}X^AX^BX^C}{X^0} \qquad \text{‘very special’} \tag{7.22}$$

The first three functions give rise to the manifold $SU(1,1)/U(1)$. However, the first one is not equivalent to the other two as the manifolds have a different value of the curvature [14]. The latter two are, however, equivalent by means of a symplectic transformation as we will show below. In the fourth example η is a constant non-degenerate real symmetric matrix. In order that the manifold has a non-empty positivity domain, the signature of this matrix should be $(+-\cdots-)$. So not all functions $F(X)$ allow a non-empty positivity domain. The last example, defined by a real symmetric tensor d_{ABC}, defines a class of special Kähler manifolds, which we will denote as ‘very special’ Kähler manifolds. This class of manifolds is important in the applications discussed below.

7.0.3 Symplectic transformations

The symplectic transformations are a generalization of the electro-magnetic duality transformations. We first recall the general formalism for arbitrary actions with coupled spin-0 and spin-1 fields, and then come to the specific case of $N = 2$.

Pseudo-symmetries in general

We consider general actions of spin-1 fields with field strengths $\mathcal{F}^{\Lambda}_{\mu\nu}$ (now labelled by $\Lambda = 1, ..., m$) coupled to scalars. The general form of the kinetic terms of the spin 1 fields is

$$\begin{aligned}\mathcal{L}_1 &= \tfrac{1}{4}(\text{Im } \mathcal{N}_{\Lambda\Sigma})\mathcal{F}^{\Lambda}_{\mu\nu}\mathcal{F}^{\mu\nu\Sigma} \\ &\quad -\tfrac{i}{8}(\text{Re } \mathcal{N}_{\Lambda\Sigma})\epsilon^{\mu\nu\rho\sigma}\mathcal{F}^{\Lambda}_{\mu\nu}\mathcal{F}^{\Sigma}_{\rho\sigma} \\ &= \tfrac{1}{2}\text{Im } \left(\mathcal{N}_{\Lambda\Sigma}\mathcal{F}^{+\Lambda}_{\mu\nu}\mathcal{F}^{+\mu\nu\Sigma}\right)\end{aligned} \tag{7.23}$$

We define

$$\begin{aligned}G^{\mu\nu}_{+\Lambda} &\equiv 2i\frac{\partial\mathcal{L}}{\partial\mathcal{F}^{+\Lambda}_{\mu\nu}} = \mathcal{N}_{\Lambda\Sigma}\mathcal{F}^{+\Sigma\,\mu\nu} \\ G^{\mu\nu}_{-\Lambda} &\equiv -2i\frac{\partial\mathcal{L}}{\partial\mathcal{F}^{-\Lambda}_{\mu\nu}} = \bar{\mathcal{N}}_{\Lambda\Sigma}\mathcal{F}^{-\Sigma\,\mu\nu}.\end{aligned} \tag{7.24}$$

The equations for the field strengths can then be written as

$$\begin{aligned}\partial^{\mu}\text{Im } \mathcal{F}^{+\Lambda}_{\mu\nu} &= 0 \qquad \text{Bianchi identities} \\ \partial_{\mu}\text{Im } G^{\mu\nu}_{+\Lambda} &= 0 \qquad \text{Equations of motion}\end{aligned}$$

This set of equations is invariant under $GL(2m, \mathbb{R})$ transformations:

$$\begin{pmatrix}\widetilde{\mathcal{F}}^{+} \\ \widetilde{G}_{+}\end{pmatrix} = \mathcal{S}\begin{pmatrix}\mathcal{F}^{+} \\ G_{+}\end{pmatrix} = \begin{pmatrix}A & B \\ C & D\end{pmatrix}\begin{pmatrix}\mathcal{F}^{+} \\ G_{+}\end{pmatrix}. \tag{7.25}$$

However, the $G_{\mu\nu}$ are related to the $\mathcal{F}_{\mu\nu}$ as in (7.24). The previous transformation implies

$$\begin{aligned}\widetilde{G}^{+} &= (C + D\mathcal{N})F^{+} \\ &= (C + D\mathcal{N})(A + B\mathcal{N})^{-1}\widetilde{F}^{+}.\end{aligned} \tag{7.26}$$

Therefore the new tensor $\mathcal{N}$ is

$$\boxed{\widetilde{\mathcal{N}} = (C + D\mathcal{N})(A + B\mathcal{N})^{-1}} \tag{7.27}$$

This tensor should be symmetric, as it is the second derivative of the action with respect to the field strength. This request leads to the equations which determine that $\mathcal{S} \in Sp(2m, \mathbb{R})$, i.e.

$$\mathcal{S}^T\Omega\mathcal{S} = \Omega \qquad \text{where} \qquad \Omega = \begin{pmatrix}0 & \mathbb{1} \\ -\mathbb{1} & 0\end{pmatrix}$$

$$\text{or} \begin{cases}A^TC - C^TA = 0 \\ B^TD - D^TB = 0 \\ A^TD - C^TB = \mathbb{1}\end{cases}. \tag{7.28}$$

Some remarks are in order: First, these transformations act on the field strengths. They generically rotate electric into magnetic fields and vice versa. Such rotations, which are called duality transformations, because in four space-time dimensions electric and magnetic fields are dual to each other in the sense of Poincaré duality, cannot be implemented on the vector potentials, at least not in a local way. Therefore, the use of these symplectic transformations is only legitimate for zero gauge coupling constant. From now on, we deal exclusively with Abelian gauge groups. Secondly, the Lagrangian is not an invariant if C and B are not zero:

$$\begin{aligned}&\text{Im } \widetilde{\mathcal{F}}^{+\Lambda}\widetilde{G}_{+\Lambda} = \text{Im } \left(\mathcal{F}^{+}G_{+}\right) \\ &+\text{Im } \left(2\mathcal{F}^{+}(C^TB)G_{+} + \mathcal{F}^{+}(C^TA)\mathcal{F}^{+}\right. \\ &\left.+G_{+}(D^TB)G_{+}\right).\end{aligned} \tag{7.29}$$

If $C \neq 0, B = 0$ it is invariant up to a four–divergence. Thirdly, the transformations can also act on dyonic solutions of the field equations and the vector $\begin{pmatrix} q_m^\Lambda \\ q_{e\,\Lambda} \end{pmatrix}$ of magnetic and electric charges transforms also as a symplectic vector. The Schwinger-Zwanziger quantization condition restricts these charges to a lattice with minimal surface area proportional to λ. Invariance of this lattice restricts the symplectic transformations to a discrete subgroup:

$$\mathcal{S} \in Sp(2m, \mathbb{Z}). \tag{7.30}$$

Finally, the transformations with $B \neq 0$ will be non–perturbative. This can be seen from the fact that they do not leave the purely electric charges invariant, or from the fact that (7.27) shows that these transformations invert $\mathcal{N}$ which plays the role of the gauge coupling constant.

Pseudo–symmetries and proper symmetries

The transformations described above, change the matrix $\mathcal{N}$, which are gauge coupling constants of the spin-1 fields. This can be compared to diffeomorphisms of the scalar manifold $z \to \hat{z}(z)$ which change the metric (which is the coupling constant matrix for the kinetic energies of the scalars) and $\mathcal{N}$:

$$\hat{g}_{\alpha\beta}(\hat{z}(z))\frac{\partial \hat{z}^\alpha}{\partial z^\gamma}\frac{\partial \hat{z}^\beta}{\partial z^\delta} = g_{\gamma\delta}(z)\ ; \hat{\mathcal{N}}(\hat{z}(z)) = \mathcal{N}(z).$$

Both these diffeomorphisms and symplectic reparametrizations are *'Pseudo–symmetries':* [15]

$$D_{pseudo} = Diff(\mathcal{M}) \times Sp(2m, \mathbb{R}). \tag{7.31}$$

They leave the action form invariant, but change the coupling constants and are thus not invariances of the action.

If $\hat{g}_{\alpha\beta}(z) = g_{\alpha\beta}(z)$ then the diffeomorphisms become isometries of the manifold, and proper symmetries of the scalar action. If these isometries are combined with symplectic transformations such that

$$\hat{\tilde{\mathcal{N}}}(z) = \mathcal{N}(z)\ , \tag{7.32}$$

then this is a *proper symmetry*. These are invariances of the equations of motion (but not necessarily of the action as not all transformations can be implemented locally on the gauge fields). To extend the full group of isometries of the scalar manifold to proper symmetries, one thus has to embed this isometry group in $Sp(2m; \mathbb{R})$, and arrives at the following situation:

$$D_{prop} = Iso(\mathcal{M}) \subset Iso(\mathcal{M}) \times Iso(\mathcal{M}) \subset D_{pseudo}$$

Let us illustrate how S and T dualities, treated in Sen's lectures [16], fit in this scheme as proper symmetries. The action he treats occurs in $N = 4$ supergravity. The scalars are $\lambda = \lambda_1 + i\lambda_2$ and a symmetric matrix M, satisfying $M\eta M = \eta^{-1}$ where $\eta = \eta^T$ is the metric of $O(6, 22)$. Their coupling to the spin-1 fields is encoded in the matrix

$$\mathcal{N} = \lambda_1\eta + i\lambda_2\eta M\eta. \tag{7.33}$$

The transformations on the scalars should lead to (7.27) with (7.28). Let us first consider this for the T *dualities*. These are transformations of $O(6, 22)$:

$$\widetilde{\mathcal{F}}^+ = A\mathcal{F}^+\ ; \qquad \widetilde{M} = AMA^T\ , \tag{7.34}$$

(λ is invariant) where $\eta = A^T\eta A$. This leads to $\widetilde{\mathcal{N}} = \left(A^T\right)^{-1}\mathcal{N}A^{-1}$, which is of the form (7.27), identifying $D = \left(A^T\right)^{-1}$. The matrices C and B are zero, which indicates that these symmetries are realised perturbatively.

For the S *dualities*, M is invariant. These transformations are determined by the integers s,r,q,p such that $sp-qr=1$:

$$\widetilde{\mathcal{F}}^+ = s\mathcal{F}^+ + r\eta^{-1}\mathcal{N}\mathcal{F}^+ \ ; \ \widetilde{\lambda} = \frac{p\lambda+q}{r\lambda+s}.$$

This leads to $\widetilde{\mathcal{N}} = (p\mathcal{N}+q\eta)(r\eta^{-1}\mathcal{N}+s)$, which is of the required form upon the identification

$$\mathcal{S} = \begin{pmatrix} s\mathbb{1} & r\eta^{-1} \\ q\eta & p\mathbb{1} \end{pmatrix}. \tag{7.35}$$

Now, B and C are non-zero, which shows the non-perturbative aspect of the S-duality.

Symplectic transformations in $N=2$

In $N=2$ the tensor $\mathcal{N}$ is determined by the function F as explained in section 7.0.2. The definitions of $\mathcal{N}$ in rigid and local supersymmetry can be written in a clarifying way as follows*

$$\begin{array}{ll} \text{rigid SUSY} & \text{SUGRA} \\ \partial_{\bar{C}}\bar{F}_A = \mathcal{N}_{AB}\partial_{\bar{C}}\bar{X}^B & \partial_{\bar{\gamma}}\bar{F}_I = \mathcal{N}_{IJ}\partial_{\bar{\gamma}}\bar{X}^J \\ & F_I = \mathcal{N}_{IJ}X^J \end{array} \tag{7.36}$$

From this definition it is easy to see that $\mathcal{N}$ transforms in the appropriate way if we define

$$\begin{array}{ll} V = \begin{pmatrix} X^A \\ F_A \end{pmatrix} & V = \begin{pmatrix} X^I \\ F_I \end{pmatrix} \\ U_C = \begin{pmatrix} \partial_C X^A \\ \partial_C F_A \end{pmatrix} & U_\alpha = \begin{pmatrix} \partial_\alpha X^I \\ \partial_\alpha F_I \end{pmatrix} \end{array} \tag{7.37}$$

(and their complex conjugates) as symplectic vectors in the two cases. They thus transform as in (7.25). With this identification in mind, we can reconsider the kinetic terms of the scalars. Then it is clear that the Kähler potentials (7.5) and (7.12), and the constraint (7.10) are symplectic invariants. This will lead to a new formulation of special geometry in section 7.0.3.

When we start from a prepotential $F(X)$, the F_I are the derivatives† of F. The expression $\tilde{X}^I = A^I{}_J X^J + B^{IJ}F_J(X)$ expresses the dependence of the new coordinates $\tilde{X}$ on the old coordinates X. If this transformation is invertible ‡, the $\tilde{F}_I$ are again the derivatives of an new function $\tilde{F}(\tilde{X})$ of the new coordinates,

$$\tilde{F}_I(\tilde{X}) = \frac{\partial\tilde{F}(\tilde{X})}{\partial\tilde{X}^I}. \tag{7.38}$$

The integrability condition which implies this statement is equivalent to the condition that $\mathcal{S}$ is a symplectic matrix. In the supergravity case, one can obtain $\tilde{F}$ due to the homogeneity:

$$\tilde{F}(\tilde{X}(X)) = \frac{1}{2}V^T\begin{pmatrix} C^TA & C^TB \\ D^TA & D^TB \end{pmatrix}V. \tag{7.39}$$

*For the rigid case, here $\partial_{\bar{C}}\bar{X}^B = \delta^B_{\bar{C}}$, but this definition is also applicable when we take derivatives w.r.t. arbitrary coordinates $z^\alpha(X)$. For the local case one regards $(\partial_{\bar{\gamma}}\bar{F}_I, F_I)$ as an $n+1$ by $n+1$ matrix to see how this defines the matrix $\mathcal{N}$.

†The remarks below are written with indices I,J as in the supergravity case, but can be applied as well in rigid supersymmetry replacing these indices by A,B.

‡The full symplectic matrix is always invertible, but this part may not be. In rigid supersymmetry, the invertibility of this transformation is necessary for the invertibility of $\mathcal{N}$, but in supergravity we may have that the $\tilde{X}^I$ do not form an independent set, and then $\tilde{F}$ can not be defined. See below.

Hence we obtain a new formulation of the theory, and thus of the target-space manifold, in terms of the function $\tilde{F}$.

We have to distinguish two situations:
1. The function $\tilde{F}(\tilde{X})$ is different from $F(\tilde{X})$, even taking into account (7.6). In that case the two functions describe equivalent classical field theories. We have a *pseudo symmetry*. These transformations are called symplectic reparametrizations [4]. Hence we may find a variety of descriptions of the same theory in terms of different functions F.
2. If a symplectic transformation leads to the same function F (again up to (7.6)), then we are dealing with a *proper symmetry*. As explained above, this invariance reflects itself in an isometry of the target-space manifold. Henceforth these symmetries are called 'duality symmetries', as they are generically accompanied by duality transformations on the field equations and the Bianchi identities. The question remains whether the duality symmetries comprise all the isometries of the target space, i.e. whether

$$Iso(\mathcal{M}) \subset Sp(2(n+1), \mathbb{R}). \tag{7.40}$$

We investigated this question in [17] for the very special Kähler manifolds, and found that in that case one does obtain the complete set of isometries from the symplectic transformations. For generic special Kähler manifolds no isometries have been found that are not induced by symplectic transformations, but on the other hand there is no proof that these do not exist.

Examples (in supergravity)

We present here some examples of symplectic reparametrizations and duality symmetries in the context of $N = 2$ supergravity. First consider (7.19). If we apply the symplectic transformation

$$\mathcal{S} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/3 \\ 0 & 0 & 1 & 0 \\ 0 & -3 & 0 & 0 \end{pmatrix} \tag{7.41}$$

one arrives, using (7.39), at (7.20). So this is a symplectic reparametrization, and shows the equivalence of the two forms of F as announced above.

On the other hand consider

$$\mathcal{S} = \begin{pmatrix} 1+3\epsilon & \mu & 0 & 0 \\ \lambda & 1+\epsilon & 0 & 2\mu/9 \\ 0 & 0 & 1-3\epsilon & -\lambda \\ 0 & -6\lambda & -\mu & 1-\epsilon \end{pmatrix} \tag{7.42}$$

for infinitesimal ϵ, μ, λ. Then F is invariant. On the scalar field $z = X^1/X^0$, the transformations act as

$$\delta z = \lambda - 2\epsilon z - \mu z^2/3. \tag{7.43}$$

They form an $SU(1,1)$ isometry group of the scalar manifold. The domain were the metric is positive definite is $\mathrm{Im}\, z > 0$. This shows the identification of the manifold as the coset space in (7.19), (7.20).

As a second example, consider (7.18). Using (7.17) one obtains the matrix $\mathcal{N}$ which determines (again with $z = X^1/X^0$)

$$e^{-1}\mathcal{L}_1 = -\tfrac{1}{2}\mathrm{Re}\, \left[z \left(F^{+0}_{\mu\nu}\right)^2 + z^{-1}\left(F^{+1}_{\mu\nu}\right)^2 \right]. \tag{7.44}$$

This appears also in pure $N = 4$ supergravity in the so–called '$SO(4)$ formulation' [18]. Consider now the symplectic mapping [19]

$$\mathcal{S} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} , \tag{7.45}$$

leading to the transformations

$$\widetilde{X}^0 = X^0 \qquad \widetilde{X}^1 = -F_1 = iX^0 \tag{7.46}$$

$$\widetilde{F}_0 = F_0 \qquad \widetilde{F}_1 = X^1 . \tag{7.47}$$

This is an example where the transformation between $\tilde{X}$ and X is not invertible. Using (7.39), we obtain $\tilde{F} = 0$. However, $A + B\mathcal{N}$ is invertible, and we can compute $\tilde{N}$ using (7.27), leading to

$$e^{-1}\mathcal{L}_1 = -\tfrac{1}{2}\text{Re}\ \left[z\left(F^{+0}_{\mu\nu}\right)^2 + z\left(F^{+1}_{\mu\nu}\right)^2 \right] . \tag{7.48}$$

(We performed here a symplectic transformation, but no diffeomorphism. We are still using the same variable z). This is the form familiar from the '$SU(4)$ formulation' of pure $N = 4$ supergravity [20]. This shows that there are formulations which can not be obtained directly from a superspace action.

In the final example, we will show that this particular formulation can be the most useful one. For that we consider the manifold

$$\frac{SU(1,1)}{U(1)} \otimes \frac{SO(r,2)}{SO(r) \otimes SO(2)} . \tag{7.49}$$

This is the only special Kähler manifold which is a product of two factors [21]. Therefore it appears in string theory where the first factor contains the dilaton-axion. The first formulation of this class of manifolds used a function F of the type (7.22): $F(X) = \frac{1}{X^0} X^S X^r X^t \eta_{rt}$, where η_{rt} is the constant diagonal metric with signature $(+, -, \ldots, -)$ [13]. In this parametrization only an $SO(r - 1)$ subgroup of $SO(r, 2)$ is linearly realized (residing in A and D of (7.25)). From a string compactification point of view one does not expect this. The full $SO(r, 2)$ should be a perturbative symmetry, as it is realized in the $N = 4$ theory described by Sen [22, 16]. In the search for better parametrizations, by means of a symplectic reparametrization a function F of the square root type was discussed in [23] which has $SO(r)$ linearly realized. However, the solution was found in [19], and was not based on a function F at all. The symplectic vector V contains then

$$F_I = S\,\eta_{IJ}\,X^J , \tag{7.50}$$

where S is one of the coordinates (representing the first factor of (7.49)), and the X^I satisfy the constraint $X^I\,\eta_{IJ}\,X^J = 0$, where η_{IJ} is the $SO(2, r)$ metric. For additional details on this example, see also [24], where the perturbative corrections to the vector multiplet couplings are considered in the context of the $N = 2$ heterotic string vacua. This important example shows that under certain circumstances one needs a formulation that does not rely on the existence of a function F.

Coordinate independent description

We want to be able to use more general coordinates than the special ones which appeared naturally in the superspace approach, and also to set up a formulation of the theory in which the symplectic structure is evident. First we will formulate this for **the rigid case** [25].

We start by introducing the symplectic vector $V \in \mathbb{C}^{2n}$, as in (7.37), where now the F_A are no longer the derivative of a function F, but n independent components. Then consider functions $V(z)$, parametrized by n coordinates z^α ($\alpha = 1, ..., n$), which will be the coordinates on the special manifold. The choice of special coordinates introduced before, corresponds to $X^A(z) = z^\alpha$, $F_A(z) = \frac{\partial F}{\partial X^A}(X(z))$. By taking now derivatives with respect to z^α one obtains U_α analogous to the U_A in (7.37).

We define as metric on the special manifold

$$g_{\alpha\bar\beta} = i\, U_\alpha^T\, \Omega\, \bar U_{\bar\beta} = i\, \langle U_\alpha, \bar U_{\bar\beta}\rangle \ , \tag{7.51}$$

where we introduced a symplectic inner product $\langle V, W\rangle \equiv V^T \Omega W$. The constraints which define the rigid special geometry can be formulated on the $2n \times 2n$ matrix

$$\mathcal{V} \equiv \begin{pmatrix} U_\alpha^T \\ \bar U^{\alpha\, T} \end{pmatrix} \equiv \begin{pmatrix} \partial_\alpha X^A & \partial_\alpha F_A \\ g^{\alpha\bar\beta}\partial_{\bar\beta}\bar X^A & g^{\alpha\bar\beta}\partial_{\bar\beta}\bar F_A \end{pmatrix} . \tag{7.52}$$

This matrix should satisfy $\mathcal{V}\Omega\mathcal{V}^T = -i\Omega$ and

$$\mathcal{D}_\alpha \mathcal{V} = \mathcal{A}_\alpha \mathcal{V} \qquad \text{with} \qquad \mathcal{A}_\alpha = \begin{pmatrix} 0 & C_{\alpha\beta\gamma} \\ 0 & 0 \end{pmatrix}$$

for a symmetric $C_{\alpha\beta\gamma}$ (being F_{ABC} in special coordinates); and $\mathcal{D}$ contains the Levi-Civita connection. The integrability condition of this constraint then implies the form of the curvature: $R_{\alpha\bar\beta\gamma\bar\delta} = -C_{\alpha\gamma\epsilon}\bar C_{\bar\beta\bar\delta\bar\epsilon}g^{\epsilon\bar\epsilon}$ (compare this with (7.7)). The formulation can even be simplified in terms of a vielbein $e_\alpha^A \equiv \partial_\alpha X^A$ (being the unit matrix in special coordinates). Then the connection $\hat\Gamma_{\alpha\beta}^\gamma = e_A^\gamma \partial_\beta e_\alpha^A$ is flat, and there are holomorphic constraints

$$\hat{\mathcal{V}} \equiv \begin{pmatrix} e_\alpha^A & \partial_\alpha F_A \\ 0 & e_A^\alpha \end{pmatrix}$$

$$\partial_\alpha \hat{\mathcal{V}} = \hat{\mathcal{A}}_\alpha \hat{\mathcal{V}} \qquad \text{with} \qquad \hat{\mathcal{A}}_\alpha = \begin{pmatrix} \hat\Gamma_{\alpha\beta}^\gamma & -iC_{\alpha\beta\gamma} \\ 0 & -\hat\Gamma_{\alpha\gamma}^\beta \end{pmatrix}$$

For **Supergravity** a similar definition of special geometry is possible. This formulation was first given in the context of a treatment of the moduli space of Calabi-Yau three-folds [2, 5, 7]. The particular way in which we present it here is explained in more detail in [26]. Now the symplectic vectors have $2(n+1)$ components. We first impose the constraint (7.10), which is written in a symplectic way as $\langle \bar V, V\rangle \equiv \bar V^T \Omega V = -i$. Then we define n holomorphic symplectic sections, parametrized by z^α, which are proportional to V:

$$V(z,\bar z) = e^{\frac{1}{2}K(z,\bar z)} v(z) \ , \tag{7.53}$$

and the proportionality constant defines the Kähler potential. These equations are then invariant under 'Kähler transformations'

$$\begin{aligned} &v(z) \to e^{f(z)}\, v(z) \\ &K(z,\bar z) \to K(z,\bar z) - f(z) - \bar f(\bar z) \\ &V \to e^{\frac{1}{2}(f(z)-\bar f(\bar z))}\, V . \end{aligned} \tag{7.54}$$

for which $\partial_\alpha K$ and $\partial_{\bar\alpha} K$ play the role of connections. Then special geometry is defined, using* $U_\alpha = \mathcal{D}_\alpha V$, with one additional constraint:

$$\langle U_\alpha, U_\beta\rangle = 0. \tag{7.55}$$

*The connection contains now the Levi–Civita one and the Kähler connection related to (7.54): $\mathcal{D}_\alpha X = \left(\partial_\alpha + \frac{1}{2}(\partial_\alpha K)\right) X$.

Usually the $F_I(z)$ are functions which depend on $X^I(z)$. Then one has $F_I = \partial_I F$, and the scaling symmetry implies that F is a holomorphic function homogeneous of 2nd degree in X^I. But e.g. with (7.50) this is not the case.

To make contact with the Picard-Fuchs equations in Calabi-Yau manifolds, a similar formulation as for the rigid case is useful. This is obtained by defining the $(2n+2)\times(2n+2)$ matrix

$$\mathcal{V} = \begin{pmatrix} V \\ \bar{U}^\alpha \\ \bar{V} \\ U_\alpha \end{pmatrix}\,, \tag{7.56}$$

which satisfies $\mathcal{V}\,\Omega\,\mathcal{V}^{\mathrm{T}} = i\Omega$. One then introduces a connection such that the constraints are [27]

$$\mathcal{D}_\alpha \mathcal{V} = \mathcal{A}_\alpha \mathcal{V}\,, \qquad \mathcal{D}_{\bar\alpha} \mathcal{V} = \mathcal{A}_{\bar\alpha} \mathcal{V}\,. \tag{7.57}$$

$$\text{with e.g. } \mathcal{A}_\alpha = \begin{pmatrix} 0 & 0 & 0 & \delta_\alpha^\gamma \\ 0 & 0 & \delta_\alpha^\beta & 0 \\ 0 & 0 & 0 & 0 \\ 0 & C_{\alpha\beta\gamma} & 0 & 0 \end{pmatrix}. \tag{7.58}$$

The integrability conditions lead to the curvature tensor

$$R_{\alpha\bar\beta\delta\bar\gamma} = g_{\alpha\bar\beta}g_{\delta\bar\gamma} + g_{\alpha\bar\gamma}g_{\delta\bar\beta} - C_{\alpha\delta\epsilon}C_{\bar\beta\bar\gamma\bar\epsilon}g^{\epsilon\bar\epsilon}. \tag{7.59}$$

7.0.4 Further results and conclusions

Special geometry is not confined to Kähler manifolds. There exist a **c** map, which can be obtained either from dimensional reduction of the field theory to 3 dimensions, or from superstring compactification mechanisms [4]. This maps special Kähler manifolds to a subclass of the quaternionic manifolds, which are then called special quaternionic. As already mentioned, a subclass of special manifolds are the 'very special' ones. These can be obtained from dimensional reduction of actions in 5 dimensions, characterised by a symmetric tensor d_{ABC} [28]. This mapping is called the **r** map [29], and the manifolds in the 5-dimensional theory are called 'very special real' manifolds. These concepts were very useful in the classification of homogeneous [30] and symmetric [14] special manifolds. It turned out that homogeneous special manifolds are in one-to-one correspondence to realizations of real Clifford algebras with signature $(q+1,1)$ for real, $(q+2,2)$ for Kähler, and $(q+3,3)$ for quaternionic manifolds. A study of the full set of isometries could be done systematically in these models. All this has been summarised in [26].

For string theory the implications of special geometry in the rigid theories for the moduli spaces of Riemann surfaces [25], and in the supergravity theories for Calabi-Yau spaces [2, 3, 4, 5, 6, 7] is extremely useful for obtaining non-perturbative results [31, 25, 19]. For these results we refer to [8] and to [32], where many more aspects of special manifolds in the context of topological theories, Landau-Ginzburg theories, etc. are discussed.

Acknowledgements

This review is a result of many collaborations. We want to thank all our collaborators and especially Sergio Ferrara, Anna Ceresole, Riccardo D'Auria, Pietro Frè, Marco Billó and Paolo Soriani.

8 Duality in supersymmetry for Children by JM Figueroa-O'Farrill

In this chapter we discuss the intimate relation between supersymmetry and the Bogomol'nyi bound. The effect of supersymmetry is two-fold: first of all, it enforces the bound since this is a property of unitary representations of the supersymmetry algebra; but it also protects the bound against quantum corrections, guaranteeing that if a state saturates the bound classically, it does so quantum mechanically. This last assertion follows because, as we will see, supersymmetry multiplets corresponding to BPS-states are smaller than the multiplets of states where the Bogomol'nyi bound is not saturated.

We first discuss the supersymmetry algebra and its representations. For definiteness we shall work in four dimensions, but much of what we'll say can (and will) be used in dimensions other than four. It will be while studying (massive) representations with central charges that we will see the mechanism by which the Bogomol'nyi bound follows from the algebra. We then illustrate this fact by studying a particular example: $N = 2$ supersymmetric Yang-Mills in four dimensions. We define this theory by dimensional reduction from $N = 1$ supersymmetric Yang-Mills in six dimensions. This theory admits a Higgs mechanism by which the gauge symmetry is broken to $U(1)$ while preserving supersymmetry. The higgsed spectrum falls into a massless gauge multiplet corresponding to the unbroken $U(1)$ and two massive short multiplets. From the structure of the short $N = 2$ multiplets we can deduce that the $N = 2$ supersymmetry algebra admits central charges and, moreover, that the multiplets containing the massive vector bosons must saturate the mass bound. We will also see that this theory admits BPS-like solutions, which are shown to break one half of the supersymmetries. This implies that the BPS-monopole belongs to a short multiplet and suggests that the bound which follows abstractly from the supersymmetry algebra agrees with the Bogomol'nyi bound for dyons given by equation (1.27). This is shown to be case. Nevertheless the short multiplets containing the massive vector
bosons and those containing the BPS-monopole have different spins, whence $N = 2$ supersymmetric Yang-Mills does not yet seem to be a candidate for a theory which is (Montonen-Olive) self-dual. This problem will be solved for $N = 4$ supersymmetric Yang-Mills, which we study as the dimensional reduction of ten-dimensional $N = 1$ supersymmetric Yang-Mills. At a formal level, $N = 4$ supersymmetric Yang-Mills is qualitatively very similar to the $N = 2$ theory; except that we will see that the short multiplets which contain the solitonic and the fundamental BPS-states have the same spin. This prompts the question whether $N = 4$ supersymmetric Yang-Mills is self-dual - a conjecture that we will have ample opportunity to test as the lectures progress.

8.1 2.1 The super-Poincaré algebra in four dimensions

In this section we will briefly review the supersymmetric extension of the fourdimensional Poincaré algebra. There are plenty of good references available so we will be brief. We will follow for the most part the conventions in [Soh85], to where we refer the reader for the relevant references on supersymmetry.

8.1.1 2.1.1 Some notational remarks about spinors

The Lorentz group in four dimensions, $SO(1,3)$ in our conventions, is not simply-connected and therefore, strictly speaking, has no spinorial representations. In order to consider spinorial representations we must look to the corresponding spin group $\mathrm{Spin}(1,3)$ which happens to be isomorphic to $SL(2,\mathbb{C})$ - the group of 2×2-complex matrices with unit determinant. From its very definition, $SL(2,\mathbb{C})$ has a natural two-dimensional complex representation, which we shall call $\mathbb{S}$. More precisely, $\mathbb{S}$ is the vector space $\mathbb{C}^2$ with the natural action of $SL(2,\mathbb{C})$. If $u \in \mathbb{S}$ has components $u_\alpha = (u_1, u_2)$ relative to some fixed basis, and $M \in SL(2,\mathbb{C})$, the action of M on u is defined simply by $(Mu)_\alpha = M_\alpha{}^\beta u_\beta$. We will abuse the notation and think of the components u_α as the vector and write $u_\alpha \in \mathbb{S}$.

This is not the only possible action of $SL(2,\mathbb{C})$ on $\mathbb{C}^2$, though. We could also define an action by using instead of the matrix M, its complex conjugate $\bar{M}$, its inverse transpose $(M^t)^{-1}$ or its inverse hermitian adjoint $\left(M^\dagger\right)^{-1}$, since they all obey the same group multiplication law. These choices correspond, respectively to the conjugate representation $\overline{\mathbb{S}}$, the dual representation $\mathbb{S}^*$, and the conjugate dual representation $\overline{\mathbb{S}}^*$. We will use the following notation: if $u_\alpha \in \mathbb{S}$, then $u_{\dot{\alpha}} \in \overline{\mathbb{S}}, u^\alpha \in \mathbb{S}^*$ and $u^{\dot{\alpha}} \in \overline{\mathbb{S}}^*$. These representations
are not all different, however. Indeed, we have that $\mathbb{S} \cong \mathbb{S}^*$ and $\overline{\mathbb{S}} \cong \overline{\mathbb{S}}^*$, which follows from the existence of $\epsilon_{\alpha\beta}$: an $SL(2,\mathbb{C})$-invariant tensor (since $\epsilon_{\alpha\beta} \mapsto M_\alpha{}^{\alpha'} M_\beta{}^{\beta'} \epsilon_{\alpha'\beta'} = (\det M)\epsilon_{\alpha\beta}$ and $\det M = 1$) which allows us to raise and lower indices in an $SL(2,\mathbb{C})$-covariant manner: $u^\alpha = \epsilon^{\alpha\beta} u_\beta$, and $u^{\dot{\beta}} = u_{\dot{\alpha}} \epsilon^{\dot{\alpha}} \dot{\beta}$. We use conventions where $\epsilon_{12} = 1$ and $\epsilon_{\dot{1}\dot{2}} = -1$.

Because both the Lie algebra $sl(2,\mathbb{C})$ (when viewed as a real Lie algebra) and $su(2) \times su(2)$ are real forms of the same complex Lie algebra, one often employs the notation (j, j') for representations of $SL(2,\mathbb{C})$, where j and j' are the spins of the two $su(2)$'s. In this notation the trivial one dimensional representation is denoted $(0,0)$, whereas $\mathbb{S} = \left(\frac{1}{2}, 0\right)$. The two $su(2)$'s are actually not independent but are related by complex conjugation, hence $\overline{\mathbb{S}} = \left(0, \frac{1}{2}\right)$. In general, complex conjugation will interchange the labels. If the labels are the same, say $\left(\frac{1}{2}, \frac{1}{2}\right)$, complex conjugation sends the representation to itself and it makes sense to restrict to the sub-representation which is fixed by complex conjugation. This is a real representation and in the case of the $\left(\frac{1}{2}, \frac{1}{2}\right)$ representation of $SL(2,\mathbb{C})$, it coincides with the defining representation of the Lorentz group $SO(1,3)$: that is, the vector representation.

Indeed, given a 4 -vector $P_\mu = \left(p_0, \overrightarrow{\mathbf{p}}\right)$ we can turn it into a bispinor as follows:

$$\sigma \cdot P \equiv \sigma^\mu P_\mu = \begin{pmatrix} p_0 + p_3 & p_1 - ip_2 \\ p_1 + ip_2 & p_0 - p_3 \end{pmatrix}$$

where $\sigma^\mu = (\mathbb{1}, \overrightarrow{\boldsymbol{\sigma}})$ with $\overrightarrow{\boldsymbol{\sigma}}$ the Pauli matrices:

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Since the Pauli matrices are hermitian, so will be $\sigma \cdot P$ provided P_μ is real. The Pauli matrices have indices $\sigma^\mu_{\alpha\dot{\alpha}}$, which shows how $SL(2,\mathbb{C})$ acts on this space. If $M \in SL(2,\mathbb{C})$, then the action of M on such matrices is given by $\sigma \cdot P \mapsto M\sigma \cdot PM^\dagger$. This action is linear and preserves both the hermiticity of $\sigma \cdot P$ and the determinant $\det(\sigma \cdot P) = P^2 = p_0^2 - \overrightarrow{\mathbf{p}} \cdot \overrightarrow{\mathbf{p}}$, just as we expect of Lorentz transformations. We can summarise this discussion by saying that the $\sigma^\mu_{\alpha\dot{\alpha}}$ are Clebsch-Gordon coefficients intertwining between the "vector" and the $\left(\frac{1}{2}, \frac{1}{2}\right)$ representations of $SL(2,\mathbb{C})$. Notice also that both M and $-M$ at the same way on bispinors, which reiterates the fact that $SL(2,\mathbb{C})$ is the double-cover of the Lorentz group $SO(1,3)$.

Finally we discuss the adjoint representation of the Lorentz group, which is generated by antisymmetric tensors $L_{\mu\nu} = -L_{\nu\mu}$. In terms of bispinors, such an $L_{\mu\nu}$ becomes a pair $\left(L_{\alpha\beta}, \bar{L}_{\dot{\alpha}\dot{\beta}}\right)$ where $L_{\alpha\beta} = L_{\beta\alpha}$ and similarly for $\bar{L}_{\dot{\alpha}\dot{\beta}}$. In other words, $L_{\mu\nu}$ transforms as the $(1,0) \oplus (0,1)$ representation of $SL(2,\mathbb{C})$: notice that we need to take the direct sum because the representation is real.

8.1.2 2.1.2 The Coleman-Mandula and Haag-LopuszańskiSohnius theorems

Back in the days when symmetry was everything, physicists spent a lot of time trying to unify the internal symmetries responsible for the observed particle spectrum and the Poincaré group into the same group: the holy grail being the so-called relativistic quark model. However their hopes were dashed by the celebrated no-go theorem of Coleman and Mandula. In a nutshell, this theorem states that the maximal Lie algebra of symmetries of the S-matrix of a unitary local relativistic quantum field theory obeying some technical but reasonable assumptions (roughly equivalent to demanding that the S-matrix be analytic), is a direct product of the Poincaré algebra with the Lie algebra of some compact internal symmetry group. Since Lie algebras of compact Lie groups are reductive: that is, the direct product of a semisimple and an abelian Lie algebras, the largest Lie algebra of symmetries of the S-matrix is a direct product: Poincaré $\times$ semisimple $\times$ abelian. In particular this implies that multiplets of the internal symmetry group consist of particles with the same mass and the same spin or helicity.

If all one-particle states are massless, then the symmetry is enhanced to conformal $\times$ semisimple $\times$ abelian; but the conclusions are unaltered: there is no way to unify the spacetime symmetries and the internal symmetries in a nontrivial way.

A wise person once said that inside every no-go theorem there is a "yesgo" theorem waiting to come out, □ and the Coleman-Mandula theorem is no exception. The trick consists, not in trying to relax some of the assumptions on the S-matrix of the field theory, but in redefining the very notion of symmetry to encompass Lie superalgebras. In a classic paper Haag, Lopuzański and Sohnius re-examined the result of Coleman and Mandula in this new light and found the most general Lie superalgebra of symmetries of an S matrix. The Coleman-Mandula theorem applies to the bosonic sector of the Lie superalgebra, so this is given again by Poincaré $\times$ reductive. In terms of representations of $SL(2,\mathbb{C})$, these generators transform according to the $(0,0)$, $\left(\frac{1}{2},\frac{1}{2}\right)$, $(0,1)$ and $(1,0)$ representations. The singlets are the internal symmetry generators which we will denote collectively by B_ℓ. The $\left(\frac{1}{2},\frac{1}{2}\right)$ generators correspond to the translations $P_{\alpha\dot{\alpha}}$, and the $(1,0)$ and $(0,1)$ generators are the Lorentz generators: $L_{\alpha\beta}$ and $\bar{L}_{\dot{\alpha}\dot{\beta}}$.

The novelty lies in the fermionic sector, which is generated by spinorial charges $Q_{\alpha I}$ in the $\left(\frac{1}{2},0\right)$ representation of $SL(2,\mathbb{C})$ and their hermitian adjoints $\bar{Q}^I_{\dot{\alpha}}=(Q_{\alpha I})^\dagger$ in the $\left(0,\frac{1}{2}\right)$. Here I is a label running from 1 to some positive integer N. The Lie superalgebra generated by these objects is called the N-extended super-Poincaré algebra. The important Lie brackets are given by

$$\begin{array}{ll}
[B_\ell,Q_{\alpha I}]=b_{\ell I}{}^J Q_{\alpha J} & \left[B_\ell,\bar{Q}^I_{\dot{\alpha}}\right]=-\bar{b}_\ell{}^I{}_J\bar{Q}^J_{\dot{\alpha}} \\
[P_{\alpha\dot{\alpha}},Q_{\beta I}]=0 & \left[P_{\alpha\dot{\alpha}},Q^I_{\dot{\alpha}}\right]=0 \\
\{Q_{\alpha I},Q_{\beta J}\}=2\epsilon_{\alpha\beta}Z_{IJ} & \left\{\bar{Q}^I_{\dot{\alpha}},\bar{Q}^J_{\dot{\beta}}\right\}=-2\bar{\epsilon}_{\dot{\alpha}\dot{\beta}}Z^{IJ} \\
\left\{Q_{\alpha I},\bar{Q}^J_{\dot{\alpha}}\right\}=2\delta_I^J P_{\alpha\dot{\alpha}} & [Z_{IJ},\ \text{anything}\ ]=0
\end{array} \tag{2.1}$$

where $Z_{IJ}=z_{IJ}{}^m B_m$, $Z^{IJ}=(Z_{IJ})^\dagger$ and the coefficients $b_{\ell I}{}^J$ and $z_{IJ}{}^m$ must obey:

$$b_{\ell I}{}^K z_{KJ}{}^m+b_{\ell J}{}^K z_{IK}{}^m=0 \tag{2.2}$$

This last condition is nontrivial and constraints the structure of that part of the internal symmetry group which acts nontrivially on the spinorial charges of the supersymmetry algebra, what we will call the internal automorphism group of the supersymmetry algebra. In the absence of central charges, the internal automorphism group of the supersymmetry algebra is $U(N)$, but in the presence of the central charges, it gets restricted generically to $USp(N)$, since condition (2.2) can be interpreted as the invariance under the internal automorphism group

*[1] and a wise guy said that we should call it a "go-go" theorem

of each of the antisymmetric forms $z_{IJ}{}^m$, for each fixed value of m. Notice that $Z_{IJ} = -Z_{JI}$, whence central charges requires $N \geq 2$.

The above Lie superalgebra is the most general symmetry of a local relativistic S-matrix in a theory describing point-particles. In the presence of extended objects: strings or, more generally, p-branes, the supersymmetry algebra receives extra terms involving topological conserved charges. These charges are no longer central since they fail to commute with the Lorentz generators; nevertheless they still commute with the spinorial charges and with the momentum generators. We will see an example of this later on when we discuss the six-dimensional $N = 1$ supersymmetry algebra.

It is sometimes convenient, especially when considering supersymmetry algebras in dimensions other than 4, where there is no analogue to the isomorphism $\mathrm{Spin}(1,3) \cong SL(2,\mathbb{C})$, to work with 4 -spinors. We can assemble the spinorial changes $Q_{\alpha I}$ and $\bar{Q}^I_{\dot{\alpha}}$ into a Majorana spinor: $Q_I = \left(Q_{\alpha I}, \bar{Q}^{\dot{\alpha} I}\right)^t$. The Dirac (= Majorana) conjugate is given by $\bar{Q}_I = \left(Q^\alpha{}_I, \bar{Q}^I_{\dot{\alpha}}\right)$, and the relevant bit of the supersymmetry algebra is now given by

$$\left\{Q_I, \bar{Q}_J\right\} = 2\delta_{IJ}\gamma^\mu P_\mu + 2i\left(\operatorname{Im} Z_{IJ} + \gamma_5 \operatorname{Re} Z_{IJ}\right) \tag{2.3}$$

where our conventions are such that

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix} \tag{2.4}$$

and $\bar{\sigma}^\mu = (\mathbb{1}, -\vec{\boldsymbol{\sigma}})$.

8.2 2.2 Unitary representations of the supersymmetry algebra

The construction of unitary representations of the super-Poincaré algebra can be thought of as a mild extension of the construction of unitary representations of the Poincaré algebra. Because the Lorentz group is simple but noncompact, any nontrivial unitary representation is infinite-dimensional. The irreducible unitary representations are simply given by classical fields in Minkowski space subject to their equations of motion. Indeed the KleinGordon and Dirac equations, among others, can be understood as irreducibility constraints on the fields. The method of construction for the Poincaré algebra is originally due to Wigner and was greatly generalised by Mackey. The method consists of inducing the representation from a finite-dimensional unitary representation of some compact subgroup. Let us review this briefly.

8.2.1 2.2.1 Wigner's method and the little group

The Poincaré algebra has two casimir operators: P^2 and W^2, where $W^\mu = \frac{1}{2}\epsilon^{\mu\nu\lambda\rho} P_\nu L_{\lambda\rho}$ is the Pauli-Lubansky vector. By Schur's lemma, on an irreducible representation they must both act as multiplication by scalars. Let's focus on P^2. On an irreducible representation $P^2 = M^2$, where M is the "rest-mass" of the particle described by the representation. With our choice of metric, physical masses are real, whence $M^2 \geq 0$. We can thus distinguish two kinds of representations: massless for which $M^2 = 0$ and massive for which $M^2 > 0$.

Wigner's method starts by choosing a nonzero momentum k_μ on the massshell: $k^2 = M^2$. That is, this is a character (that is, a one-dimensional representation) of the translation subalgebra generated by the P_μ. We let G_k denote the subgroup of the Lorentz group (or rather of its double-cover $SL(2,\mathbb{C})$) which leaves k_μ invariant. G_k is known as the little group. Wigner's method, which we will not describe in any more detail than this, consists in inducing a representation of the Poincaré group from a finite-dimensional unitary representation of the little group. This is done by boosting the representation to fields on the mass shell and then Fourier transforming to yield fields on Minkowski space subject to their equations of motion.

In extending this method to the super-Poincaré algebra all that happens is that now the Lie algebra of the little group gets extended by the spinorial supersymmetry charges, since these commute with P_μ and hence stabilise the chosen 4 -vector.

We will need to know about the structure of the little groups before introducing supersymmetry. The little group happens to be different for massive and for massless representations, as the next exercise asks you to show.

Exercise

2.1 (The little groups for positive-energy particles)

Let k_μ be a 4 -vector obeying $k_0 > 0, k^2 = M^2 \geq 0$. Prove that the little group of k_μ is isomorphic to:

SU(2), for $M^2 > 0$; $\widetilde{E}_2$, for $M^2 = 0$,

where $E_2 \cong SO(2) \ltimes \mathbb{R}^2$, is the two-dimensional euclidean group and $\widetilde{E}_2 \cong \mathrm{Spin}(2) \ltimes \mathbb{R}^2$ is its double cover.

(Hint: argue that two momenta k_μ which are Lorentz-related have isomorphic little groups. Then choose a convenient k_μ in each case, examine the action of $SL(2,\mathbb{C})$ on the bispinor $\sigma^\mu k_\mu$, and identify those $M \in SL(2,\mathbb{C})$ for which $M\sigma \cdot k M^\dagger = \sigma \cdot k$.)

The reason why we have restricted ourselves to positive-energy representations in this exercise, is that unitary representations of the supersymmetry algebra have non-negative energy. Indeed, for an arbitrary momentum k_μ, the supersymmetry algebra becomes

$$\left\{Q_{\alpha I}, \bar{Q}_{\dot{\alpha}}^J\right\} = 2\delta_I^J \begin{pmatrix} k_0+k_3 & k_1-ik_2 \\ k_1+ik_2 & k_0-k_3 \end{pmatrix}$$

Therefore the energy k_0 of any state $|k\rangle$ with momentum k_μ can be written as follows (for a fixed but otherwise arbitrary I)

$$\begin{aligned} k_0 \| |k\rangle \|^2 &= \langle k | k_0 | k\rangle \\ &= \frac{1}{2}\left\langle k\left|\left\{Q_{1I}, \bar{Q}_{\dot{1}}^I\right\}\right| k\right\rangle + \frac{1}{2}\left\langle k\left|\left\{Q_{2I}, \bar{Q}_{\dot{2}}^I\right\}\right| k\right\rangle \\ &= \frac{1}{2}\left\|Q_{1I}|k\rangle\right\|^2 + \frac{1}{2}\left\|Q_{2I}|k\rangle\right\|^2 + \frac{1}{2}\left\|\left(Q_{1I}\right)^\dagger|k\rangle\right\|^2 + \frac{1}{2}\left\|\left(Q_{2I}\right)^\dagger|k\rangle\right\|^2 \end{aligned}$$

whence k_0 is positive, unless $|k\rangle$ is annihilated by all the supersymmetry charges.

8.2.2 2.2.2 Massless representations

We start by considering massless representations. As shown in Exercise 2.1, the little group for the momentum k_μ of a massless particle is noncompact. Therefore its finite-dimensional unitary representations must all come from its maximal compact subgroup Spin(2) and be trivial on the translation subgroup $\mathbb{R}^2$. The unitary representations of Spin(2) are one-dimensional and indexed by a number $\lambda \in \frac{1}{2}\mathbb{Z}$ called the helicity. For CPT-invariance of the spectrum, it may be necessary to include both helicities $\pm\lambda$, but clearly all this does is double the states and we will not mention this again except to point out that some supersymmetry multiplets are CPT-self-conjugate.

Let's choose $k_\mu = (E, 0, 0, E)$, with $E > 0$. Then

$$\sigma^\mu k_\mu = \begin{pmatrix} 2E & 0 \\ 0 & 0 \end{pmatrix}$$

and the supersymmetry algebra becomes

$$\left\{Q_{\alpha I}, \bar{Q}^J_{\dot{\alpha}}\right\} = 4E\delta_I^J \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

In particular this means that $\left\{Q_{2I}, \bar{Q}^J_{\dot{2}}\right\} = 0$. Because $\bar{Q}^J_{\dot{2}} = (Q_{2J})^\dagger$, it follows that in a unitary representation $Q_{2I} = 0$ for all I. Indeed for any state $|\psi\rangle$,

$$0 = \left\langle \psi \left| \left\{ Q_{2I}, (Q_{2I})^\dagger \right\} \right| \psi \right\rangle = \| Q_{2I} |\psi\rangle \|^2 + \| (Q_{2I})^\dagger |\psi\rangle \|^2$$

Plugging this back into the supersymmetry algebra (2.1) we see that $Z_{IJ} = \frac{1}{2}\{Q_{1I}, Q_{2J}\} = 0$, so that there are no central charges for massless representations.

Let us now introduce $q_I \equiv (1/2\sqrt{E})Q_{1I}$, in terms of which the supersymmetry algebra becomes

$$\left\{q_I, q_J^\dagger\right\} = \delta_{IJ} \quad \{q_I, q_J\} = \left\{q_I^\dagger, q_J^\dagger\right\} = 0$$

We immediately recognise this is as a Clifford algebra corresponding to a 2 N dimensional pseudo-euclidean space with signature (N, N). The irreducible representations of such Clifford algebras are well-known. We simply start with a Clifford vacuum $|\Omega\rangle$ satisfying

$$q_I |\Omega\rangle = 0 \quad \text{for all } I = 1, \ldots, N$$

and we act repeatedly with the $q_I^\dagger$. Since $\left\{q_I^\dagger, q_J^\dagger\right\} = 0$, we obtain a 2^{N^N} dimensional representation spanned by the vectors: $q^\dagger_{I_1} q^\dagger_{I_2} \cdots q^\dagger_{I_p} |\Omega\rangle$, where $1 \leq I_1 < I_2 < \cdots < I_p \leq N$, and $p = 0, \ldots, N$.

The Clifford vacuum actually carries quantum numbers corresponding to the momentum k and also to the helicity: $|\Omega\rangle = |k, \lambda\rangle$. It may also contain quantum numbers corresponding to the internal symmetry generators B_ℓ, but we ignore them in what follows.

Exercise 2.2 (Helicity content of massless multiplets)

Paying close attention to the helicity of the supersymmetry charges, prove that $Q_{1\,\mathrm{I}}$ raises the helicity by $\frac{1}{2}$, whereas Q_{2I} lowers it by the same amount. Deduce that the massless supersymmetry multiplet of helicity λ contains the following states:

States	Helicity	Number
$\|k, \lambda\rangle$	λ	1
$q_I^\dagger \|k, \lambda\rangle$	$\lambda + \frac{1}{2}$	N
$q_J^\dagger q_I^\dagger \|k, \lambda\rangle$	$\lambda + 1$	$\binom{N}{2}$
$\vdots$	$\vdots$	$\vdots$
$q^\dagger_{I_1} q^\dagger_{I_2} \cdots q^\dagger_{I_p} \|k, \lambda\rangle$	$\lambda + p/2$	$\binom{N}{p}$
$\vdots$	$\vdots$	$\vdots$
$q_1^\dagger q_2^\dagger \cdots q_N^\dagger \|k, \lambda\rangle$	$\lambda + N/2$	1

Particularly interesting cases are the CPT-self-conjugate massless multiplets. First notice that CPT-self-conjugate multiplets can only exist for N even. For $N = 2$ we have the helicity $\lambda = -\frac{1}{2}$ multiplet, whose spectrum consists of

Helicity	$-1/2$	0	$1/2$
Number	1	2	1

Then we have the $N = 4$ gauge multiplet which has $\lambda = -1$ and whose spectrum is given by:

Helicity	-1	$-1/2$	0	$1/2$	1
Number	1	4	6	4	1

.

Pure (that is, without matter) $N = 4$ supersymmetric Yang-Mills in fourdimensions consists of several of these multiplets-one for each generator of the gauge algebra. Finally, the third interesting case is the $N = 8$ supergravity multiplet with $\lambda = -2$ and spectrum given by:

Helicity	-2	$-3/2$	-1	$-1/2$	0	$1/2$	1	$3/2$	2
Number	1	8	28	56	70	56	28	8	1

8.2.3 2.2.3 Massive representations

We now consider massive representations. As shown in Exercise 2.1, the little group for the momentum k_μ of a massive particle is $SU(2)$. Its finitedimensional irreducible unitary representations are well-known: they are indexed by the spin s, where $2s$ is a non-negative integer, and have dimension $2s + 1$.

A massive particle can always be boosted to its rest frame, so that we can choose a momentum $k_\mu = (M, 0, 0, 0)$. Then

$$\sigma^\mu k_\mu = \begin{pmatrix} M & 0 \\ 0 & M \end{pmatrix}$$

and the supersymmetry algebra becomes

$$\left\{Q_{\alpha I}, \bar{Q}^J_{\dot{\alpha}}\right\} = 2M\delta_I^J \mathbb{1}_{\alpha\dot{\alpha}}$$

No central charges

In the absence of central charges, $\{Q_{\alpha I}, Q_{\beta J}\} = 0$. Thus we can introduce $q_{\alpha I} \equiv (1/\sqrt{2M})Q_{\alpha,I}$, in terms of which the supersymmetry algebra is again a Clifford algebra:

$$\left\{q_{\alpha I}, q^\dagger_{\beta J}\right\} = \delta_{IJ}\delta_{\alpha\beta} \quad \{q_{\alpha I}, q_{\beta J}\} = \left\{q^\dagger_{\alpha I}, q^\dagger_{\beta J}\right\} = 0 \tag{2.5}$$

but where now the underlying pseudo-euclidean space is $4N$-dimensional with signature $(2N, 2N)$. The unique irreducible representation of such a Clifford algebra is now 2^{2N}-dimensional and it is built just as before from a Clifford vacuum by acting successively with the $q^\dagger_{\alpha I}$.

However unlike the case of massless representations, the Clifford vacuum is now degenerate since it carries spin: for spin s the Clifford vacuum is really a $(2s + 1)$-dimensional $SU(2)$ multiplet. Notice that for fixed $I, q^\dagger_{\alpha I}$ transforms as a $SU(2)$-doublet of spin $\frac{1}{2}$. This must be taken into account when determining the spin content of the states in the supersymmetry multiplet. Instead of simply adding the helicities like in the massless case, now we must use the Clebsch-Gordon series to add the spins.

Exercise 2.3 (Highest spin in the multiplet)

Prove that the highest spin in the multiplet will be carried by states of the form $q_1^\dagger q_{12}^\dagger \cdots q_{1N}^\dagger$ acting on the Clifford vacuum, and that their spin is $s + N/2$.

For example, if $N = 1$ and $s = 0$, then we find the following spectrum: $|k, 0\rangle$ with spin 0, $\left(q_1^\dagger |k, 0\rangle, q_2^\dagger |k, 0\rangle\right)$ with spin 1/2 and $q_1^\dagger q_2^\dagger |k, 0\rangle$ which has spin 0 too. The supersymmetric field theory describing this multiplet consists of a scalar field, a pseudo-scalar field, and a Majorana fermion: it is the celebrated Wess-Zumino model and the multiplet is known as the massive Wess-Zumino multiplet. Another example that will be important to us is the $N = 2$ multiplets with spins $s = 0$ and $s = 1/2$, which we leave as an exercise.

Exercise

2.4 (Massive $N = 2$ multiplets with $s = 0$ and $s = 1/2$)

Work out the spin content of the massive $N = 2$ multiplets without central charges and with spins $s = 0$ and $s = 1/2$. Show that for $s = 0$ the spin content is $\left(0^5, \frac{1}{2}^4, 1\right)$ in the obvious notation, and for $s = 1/2$ it is given by $\left(3/2, 1^4, \frac{1}{2}^6, 0^4\right)$.

Adding central charges

Adding central charges changes the nature of the supersymmetry algebra, which now becomes

$$\left\{Q_{\alpha I}, \bar{Q}_{\dot{\alpha}}^J\right\} = 2M\delta_I^J \mathbb{1}_{\alpha\dot{\alpha}} \quad \left\{Q_{\alpha I}, Q_{\beta J}\right\} = 2\epsilon_{\alpha\beta} Z_{IJ}$$

Because Z_{IJ} is antisymmetric, we can rotate the $Q_{\alpha I}$ unitarily-which is an automorphism of the first of the above brackets-in such a way that Z_{IJ} takes a standard form:

$$Z_{IJ} = \begin{pmatrix} 0 & z_1 & & & & & & \\ -z_1 & 0 & & & & & & \\ & & 0 & z_2 & & & & \\ & & -z_2 & 0 & & & & \\ & & & & \ddots & & & \\ & & & & & \ddots & & \\ & & & & & & 0 & z_{N/2} \\ & & & & & & -z_{N/2} & 0 \end{pmatrix}$$

where the z_i can be chosen to be real and non-negative. To simplify the discussion we have assumed that N is even, but one should keep in mind that if N is odd, there will of course be a zero 1×1 block in the above normal form for Z_{IJ}.

Let us break up the index I into a pair (A, i) where $A = 1, 2$ and $i = 1, \ldots, N/2$. In terms of these indices the supersymmetry algebra can be rewritten as

$$\left\{Q_{\alpha Ai}, Q_{\beta Bj}^\dagger\right\} = 2M\delta_{ij}\delta_{\alpha\beta}\delta_{AB} \quad \left\{Q_{\alpha Ai}, Q_{\beta Bj}\right\} = 2\epsilon_{\alpha\beta}\epsilon_{AB}\delta_{ij} z_i$$

Define the following linear combinations

$$S_{\alpha i}^{\pm} \equiv \frac{1}{2}\left(Q_{\alpha 1i} \pm \epsilon^{\dot{\alpha}\dot{\beta}} \bar{Q}_{\dot{\beta}}^{2i}\right)$$

where we have raised the spinor index in the second term in order to preserve covariance under the little group $SU(2)$. In terms of these generators, the algebra is now:

$$\left\{S_{\alpha i}^{\pm}, \left(S_{\beta j}^{\pm}\right)^\dagger\right\} = \delta_{\alpha\beta}\delta_{ij}\left(M \pm z_i\right)$$

with all other brackets being zero. Notice that acting on any state $|\psi\rangle$,

$$\begin{aligned}(M \pm z_i)\, \| |\psi\rangle \|^2 &= \langle \psi \,|(M \pm z_i)|\, \psi \rangle \\ &= \left\langle \psi \left| \left\{ S^{\pm}_{1i}, (S^{\pm}_{1i})^{\dagger} \right\} \right| \psi \right\rangle \\ &= \| S^{\pm}_{1i} |\psi\rangle \, \|^2 + \| \left(S^{\pm}_{1i}\right)^{\dagger} |\psi\rangle \|^2\end{aligned}$$

from where it follows that $M \pm z_i \geq 0$ for all i, or

$$M \geq |z_i| \qquad \text{for all } i = 1, \ldots, N/2 \tag{2.6}$$

which is reminiscent of the Bogomol'nyi bound (1.27). Notice that this bound is an unavoidable consequence of having a unitary representation of the supersymmetry algebra. Therefore provided that supersymmetry is not broken quantum-mechanically, the bound will be maintained.

Suppose that $M > z_i$ for all i. Then we can define $q^{\pm}_{\alpha i} \equiv \left(1/\sqrt{M \pm z_i}\right) S^{\pm}_{\alpha i}$, in terms of which the supersymmetry algebra is again given by equation (2.5) once we recombine the indices $(\pm, i)$ into I. Therefore we are back in the case of massive representations without central charges, at least as far as the dimension of the representations is concerned.

Suppose instead that some of the z_i saturate the bound (2.6): $z_i = M$ for $i = 1, \ldots, q \leq N/2$. Then a similar argument as in the discussion of the massless representations allows us to conclude that the $2q$ generators $S^{-}_{\alpha i}$for $i = 1, \ldots, q$ act trivially and can be taken to be zero. The remaining $2N - 2q$ generators obey a Clifford algebra whose unique irreducible representation has dimension 2^{2N-2q}. Notice that the smallest representation occurs when all central charges saturate the bound (2.6), in which case all the $S^{-}_{\alpha i} = 0$ and we are left only with 2^N states, just as in the case of a massless multiplet. These massive multiplets are known as short multiplets.

For example, in $N = 2$ there is only one $z = z_1$. If $z < M$ the massive multiplet contains $2^4 = 16$ states, whereas if $z = M$ the short multiplet only contains $2^2 = 4$ states. For $N = 4$, there are two z_i. If both $z_i < M$, then the massive multiplet has $2^8 = 256$ states, whereas if both $z_i = M$, then the short multiplet contains only $2^4 = 16$ states. Half-way we find the case when exactly one of the $z_i = M$, in which case the dimension of the multiplet is $2^6 = 64$. Strictly speaking we shouldn't call these numbers the dimension of the multiplet, but rather the degeneracy, since it may be that the Clifford vacuum is degenerate, in which case the dimension of the supersymmetry multiplet is the product of what we've been calling the dimension of the multiplet and that of the Clifford vacuum. Let us work out some examples. We first work out the case of $N = 2$ and spins $s = 0$ and $s = \frac{1}{2}$ in the following exercise.

Exercise

2.5 (Short $N = 2$ multiplets with $s = 0$ and $s = 1/2$)

Prove that the spin contents of the short multiplet with $s = 0$ is $\left(\frac{1}{2}, 0^2\right)$ and that of the short multiplet with $s = 1/2$ is $\left(1, \frac{1}{2}^2, 0\right)$. Compare with the results of Exercise 2.4, which are the spin contents when the central charge does not saturate the bound. We will see that the $s = 0$ multiplet contains the BPS-monopole, whereas the $s = 1/2$ multiplet contains the massive vector bosons.

Next we take a look at the short $N = 4$ multiplets with $s = 0$. These will be the important ones when we discuss $N = 4$ supersymmetric Yang-Mills.

Exercise

2.6 (Short $N = 4$ multiplets with $s = 0$)

Prove that the spin content of the $N = 4$ short multiplet with $s = 0$ is $\left(1, \frac{1}{2}^4, 0^5\right)$, which totals the expected 16 states. As we will see later, this will be the multiplet containing both the BPS-monopole and the massive vector boson.

This difference in the dimension of representations for which the bound (2.6) is saturated is responsible for the fact that if a multiplet saturates the bound classically, it will continue to do so when perturbative quantum corrections are taken into account. This is because perturbative quantum corrections do not alter the number of degrees of freedom, hence a short multiplet (that is, one which saturates the bound) cannot all of a sudden undergo the explosion in size required to obey the bound strictly.

8.3 2.3 $N = 2$ Supersymmetric Yang-Mills

The supersymmetric bound (2.6) for massive representations with central charges may seem a little abstract, but it comes to life in particular field theoretical models, where we can explicitly calculate the central charges in terms of the field variables. We will see this first of all in pure $N = 2$ supersymmetric Yang-Mills, which embeds the bosonic part of the Georgi-Glashow model. This result is due to Witten and Olive WO78.

We could simply write the action down and compute the supersymmetry algebra directly as was done in WO78, but it is much more instructive to derive it by dimensional reduction from the $N = 1$ supersymmetric Yang-Mills action in six dimensions. This derivation of $N = 2$ supersymmetric Yang-Mills by dimensional reduction was first done in DHdV78, and the six-dimensional computation of the central charges was first done in Oli79.

That there should be a $N = 1$ supersymmetric Yang-Mills theory in six dimensions is not obvious: unlike its nonsupersymmetric counterpart, supersymmetric Yang-Mills theories only exist in a certain number of dimensions. Of course one can always write down the Yang-Mills action in any dimension and then couple it to fermions, but supersymmetry requires a delicate balance between the bosonic and fermionic degrees of freedom. A gauge field in d dimensions has $d-2$ physical degrees of freedom corresponding to the transverse polarisations. The number of degrees of freedom of a fermion field depends on what kind fermion it is, but it always a power of 2. An unconstrained Dirac spinor in d dimensions has $2^{d/2}$ or $2^{(d-1)/2}$ real degrees of freedom, for d even or odd respectively: a Dirac spinor has $2^{d/2}$ or $2^{(d-1)/2}$
complex components but the Dirac equation cuts this number in half. In even dimensions, one can further restrict the spinor by imposing that it be chiral or Weyl. This cuts the number of degrees of freedom by two. Alternatively, in some dimensions (depending on the signature of the metric) one can impose a reality or Majorana condition which also halves the number of degrees of freedom. For a lorentzian metric of signature $(1, d-1)$, Majorana spinors exist for $d \equiv 1, 2, 3, 4 \bmod 8$. When $d \equiv 2 \bmod 8$ one can in fact impose that a spinor be both Majorana and Weyl, cutting the number of degrees of freedom in four. The next exercise asks you to determine in which dimensions can supersymmetric Yang-Mills theory exist based on the balance between bosonic and fermionic degrees of freedom.

Exercise 2.7 ($N = 1$ supersymmetric Yang-Mills)

Verify via a counting of degrees of freedom that $N = 1$ supersymmetric Yang-Mills can exist only in the following dimensions and with the following types of spinors:

d	Spinor
3	Majorana
4	Majorana or Weyl
6	Weyl
10	Majorana-Weyl

It is a curious fact that these are precisely the dimensions in which the classical superstring exists. Unlike superstring theory, in which only the ten-dimensional theory survives quantisation, it turns out that supersymmetric Yang-Mills theory exists in each of these dimensions. Although we are mostly concerned with four-dimensional field theories in these notes, the six-dimensional and ten-dimensional theories are useful tools since upon dimensional reduction to four dimensions they yield $N = 2$ and $N = 4$ supersymmetric Yang-Mills, respectively.

8.3.1 2.3.1 $N = 1 d = 6$ supersymmetric Yang-Mills

We start by setting some conventions. We will let uppercase Latin indices from the beginning of the alphabet $A, B, \ldots$ take the values $0, 1, 2, 3, 5, 6$. Our metric η_{AB} is "mostly minus"; that is, with signature $(1, 5)$. We choose the following explicit realisation of the Dirac matrices:

$$\Gamma_\mu = \begin{pmatrix} 0 & \gamma_\mu \\ \gamma_\mu & 0 \end{pmatrix} \quad \Gamma_5 = \begin{pmatrix} 0 & \gamma_5 \\ \gamma_5 & 0 \end{pmatrix} \quad \Gamma_6 = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}$$

where $\mu = 0, 1, 2, 3$, and where γ_μ are defined in (2.4) and $\gamma_5 = \gamma_0\gamma_1\gamma_2\gamma_3$. The Γ_A obey the Clifford algebra

$$\{\Gamma_A, \Gamma_B\} = 2\eta_{AB}\mathbb{1}$$

Weyl spinors are defined relative to Γ_7, which is defined by

$$\Gamma_7 = \Gamma_0\Gamma_1 \cdots \Gamma_6 = \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}$$

We now write down the action for $N = 1$ supersymmetric Yang-Mills. We will take the gauge group to be $SO(3)$ for definiteness, but it should be clear that the formalism is general. As before we will identify the Lie algebra so(3) with $\mathbb{R}^3$ but we will now drop the arrows on the vectors to unclutter the notation, hoping it causes no confusion. The lagrangian density is given by

$$\mathcal{L} = -\frac{1}{4}\mathrm{G}^{AB} \cdot \mathrm{G}_{AB} + \frac{i}{2}\bar{\Psi} \cdot \Gamma^A \overleftrightarrow{D}_A \Psi \tag{2.7}$$

where

$$\begin{aligned} \mathrm{G}_{AB} &= \partial_A \mathrm{W}_B - \partial_B \mathrm{W}_A - e\, \mathrm{W}_A \times \mathrm{W}_B \\ D_A\Psi &= \partial_A\Psi - e\, \mathrm{W}_A \times \Psi \end{aligned}$$

and where Ψ is a complex Weyl spinor obeying $\Gamma_7\Psi = -\Psi$. The Dirac conjugate spinor is defined by $\bar{\Psi} = \Psi^\dagger\Gamma_0$, and obeys $\bar{\Psi}\Gamma_7 = \bar{\Psi}$. Finally we have used the convenient shorthand $\overleftrightarrow{D}_A$ to mean

$$\bar{\Psi} \cdot \Gamma^A \overleftrightarrow{D}_A \Psi = \bar{\Psi} \cdot \Gamma^A D_A \Psi - D_A\bar{\Psi} \cdot \Gamma^A \Psi$$

The action defined by (2.7) is manifestly gauge invariant, but it is also invariant under supersymmetry. Let $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ be two constant anticommuting Weyl spinors of the same chirality as Ψ. Let us define the following transformations:

$$\delta\mathrm{W}_A = i\overline{\boldsymbol{\alpha}}\Gamma_A\Psi \qquad \bar{\Psi}\mathrm{W}_A = -i\overline{\boldsymbol{\Psi}}\Gamma_A\boldsymbol{\beta}$$

$$\delta\boldsymbol{\Psi} = 0 \qquad \bar{\delta}\boldsymbol{\Psi} = \frac{1}{2}\mathrm{G}^{AB}\Gamma_{AB}\boldsymbol{\beta}$$

$$\delta\bar{\Psi} = -\frac{1}{2}\overline{\boldsymbol{\alpha}}\mathrm{G}^{AB}\Gamma_{AB} \qquad \bar{\delta}\bar{\Psi} = 0$$

where $\Gamma_{AB} = \frac{1}{2}(\Gamma_A\Gamma_B - \Gamma_B\Gamma_A)$. We should remark that there is only one supersymmetry in our theory: $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are chiral. That is, there is only one spinorial charge Q, in terms of which the transformations δ and $\bar{\delta}$ defined above can be understood as follows:

$$\delta\phi = [\overline{\boldsymbol{\alpha}}Q, \phi] \quad \text{and} \quad \bar{\delta}\phi = [\bar{Q}\boldsymbol{\beta}, \phi]$$

for any field ϕ. Notice that it follows from this that the action of $\bar{\delta}$ can be deduced from that of δ as follows: $\bar{\delta}\phi = \left(\delta\phi^\dagger\right)^\dagger$ (apart from the obvious change of $\boldsymbol{\alpha}$ to $\boldsymbol{\beta}$, of course). Keep in mind that we have chosen the Lie algebra structure constants to be real, whence the generators are antihermitian.

We claim that $\boldsymbol{\mathcal{L}}$ is invariant under δ and $\bar{\delta}$ above up to a divergence. In order to derive the supersymmetry current, we will actually take $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ to depend on the position and simply vary the lagrangian density. We expect a total divergence plus a term with the current multiplying the derivative of the parameter. The calculations will take us until the end of the section and are contained in the following set of exercises.

Exercise 2.8 (Supersymmetry variation of $\mathcal{L}$)

Prove first of all that for any derivation δ,

$$\delta \mathrm{G}_{AB} = D_A \delta \mathrm{W}_B - D_B \delta \mathrm{W}_A$$

and conclude that the variation of the bosonic part of the action $\boldsymbol{\mathcal{L}}_\mathrm{b}$ is given by

$$\delta\boldsymbol{\mathcal{L}}_\mathrm{b} = -i\mathrm{G}^{AB} \cdot D_A\left(\overline{\boldsymbol{\alpha}}\Gamma_B\Psi\right) \quad \text{and} \quad \bar{\delta}\boldsymbol{\mathcal{L}}_\mathrm{b} = i\mathrm{G}^{AB} \cdot D_A\left(\bar{\Psi}\Gamma_B\boldsymbol{\beta}\right)$$

Next we tackle the fermions. Prove the following identities:

$$\begin{aligned}\delta\left(D_A\Psi\right) &= -ie\left(\overline{\boldsymbol{\alpha}}\Gamma_A\Psi\right) \times \Psi \\ \delta\left(D_A\bar{\Psi}\right) &= -\frac{1}{2}D_A\left(\overline{\boldsymbol{\alpha}}\mathrm{G}^{BC}\Gamma_{BC}\right) - ie\left(\overline{\boldsymbol{\alpha}}\Gamma_A\Psi\right) \times \bar{\Psi}\end{aligned}$$

and

$$\begin{aligned}\bar{\delta}\left(D_A\boldsymbol{\Psi}\right) &= \frac{1}{2}D_A\left(\mathrm{G}^{BC}\Gamma_{BC}\boldsymbol{\beta}\right) + ie\left(\overline{\boldsymbol{\Psi}}\Gamma_A\boldsymbol{\beta}\right) \times \boldsymbol{\Psi} \\ \bar{\delta}\left(D_A\overline{\boldsymbol{\Psi}}\right) &= ie\left(\overline{\boldsymbol{\Psi}}\Gamma_A\boldsymbol{\beta}\right) \times \overline{\boldsymbol{\Psi}},\end{aligned}$$

and conclude that the variation of the fermionic part of the action $\boldsymbol{\mathcal{L}}_\mathrm{f}$ is given by

$$\delta\boldsymbol{\mathcal{L}}_\mathrm{f} = \frac{i}{4}D_A\left(\overline{\boldsymbol{\alpha}}\mathrm{G}^{BC}\right) \cdot \Gamma_{BC}\Gamma^A\boldsymbol{\Psi} - \frac{i}{4}\overline{\boldsymbol{\alpha}}\mathrm{G}^{BC} \cdot \Gamma_{BC}\Gamma^A D_A\Psi + e\bar{\Psi} \cdot \left(\overline{\boldsymbol{\alpha}}\Gamma^A\boldsymbol{\Psi}\right) \times \Gamma_A\boldsymbol{\Psi}$$

and

$$\bar{\delta}\boldsymbol{\mathcal{L}}_\mathrm{f} = \frac{i}{4}\overline{\boldsymbol{\Psi}}\Gamma^A\Gamma^{BC} \cdot D_A\left(\mathrm{G}_{BC}\boldsymbol{\beta}\right) - \frac{i}{4}D_A\overline{\boldsymbol{\Psi}}\Gamma^A\Gamma^{BC} \cdot \mathrm{G}_{BC}\boldsymbol{\beta} - e\overline{\boldsymbol{\Psi}} \cdot \left(\overline{\boldsymbol{\Psi}}\Gamma^A\boldsymbol{\beta}\right) \times \Gamma_A\boldsymbol{\Psi}$$

Supersymmetry invariance demands, in particular, that the fermion trilinear terms in $\delta\boldsymbol{\mathcal{L}}_\mathrm{f}$ should cancel. This requires a Fierz rearrangement, and this is as good a time as any to discuss this useful technique. Writing explicitly the Lie algebra indices on the fermions, the trilinear terms in $\delta\boldsymbol{\mathcal{L}}_\mathrm{f}$ become

$$e\epsilon_{abc}\left(\overline{\boldsymbol{\alpha}}\Gamma^A\boldsymbol{\Psi}^a\right)\left(\overline{\boldsymbol{\Psi}}^c\Gamma_A\boldsymbol{\Psi}^b\right) \tag{2.9}$$

Let us focus on the expression $\boldsymbol{\Psi}^a \overline{\boldsymbol{\Psi}}^c$. This is a bispinor. Since spinors in six dimensions have 8 components, bispinors form a 64-dimensional vector space spanned by the antisymmetrised products of Γ-matrices:

$\mathbb{1}$, Γ_A, Γ_{AB}, Γ_{ABC}, Γ_{ABCD}, Γ_{ABCDE} and Γ_{ABCDEF}, or equivalently

$$\mathbb{1}, \quad \Gamma_A, \quad \Gamma_{AB}, \quad \Gamma_{ABC}, \quad \Gamma_{AB}\Gamma_7, \quad \Gamma_A\Gamma_7 \quad \text{and} \quad \Gamma_7.$$

(Notice that antisymmetrisation is defined by

$$\Gamma_{A_1 A_2 \cdots A_p} = \Gamma_{[A_1} \Gamma_{A_2} \cdots \Gamma_{A_p]} = \frac{1}{p!} \sum_{\sigma \in \mathfrak{S}_p} \operatorname{sign} \sigma \Gamma_{A_{\sigma(1)}} \Gamma_{A_{\sigma(2)}} \cdots \Gamma_{A_{\sigma(p)}}$$

so that it has "strength one.")

We will let $\{M_\Lambda\}$ denote collectively these matrices. The above basis is orthogonal relative to the inner product defined by the trace:

$$\operatorname{tr} M_\Lambda M_{\Lambda'} = c_\Lambda \delta_{\Lambda\Lambda'}$$

which allows us to expand

$$\Psi^a \bar{\Psi}^c = \sum_\Lambda b_\Lambda{}^{ac} M_\Lambda$$

and to compute the coefficients $b_\Lambda{}^{ac}$ simply by taking traces. Remembering that $\boldsymbol{\Psi}^a$ are anticommuting, we find

$$b_\Lambda{}^{ac} = -\frac{1}{c_\Lambda} \left(\bar{\Psi}^c M_\Lambda \Psi^a \right) \tag{2.10}$$

Exercise 2.9 (A Fierz rearrangement)

Using the above formula and computing the relevant traces, prove that

$$\boldsymbol{\Psi}^a \overline{\boldsymbol{\Psi}}^c = -\frac{1}{8} \left(\overline{\boldsymbol{\Psi}}^c \Gamma_A \boldsymbol{\Psi}^a \right) \Gamma^A \left(\mathbb{1} + \Gamma_7 \right) - \frac{1}{48} \left(\overline{\boldsymbol{\Psi}}^c \Gamma_{ABC} \boldsymbol{\Psi}^a \right) \Gamma^{ABC}$$

(Hint: use the fact that $\Gamma_7 \Psi = -\boldsymbol{\Psi}$ to discard from the start many of the terms in the general Fierz expansion.)

We now use this Fierz rearrangement to rewrite the trilinear term (2.9) as follows:

$$-\frac{1}{8} e \epsilon_{abc}) \, \overline{\boldsymbol{\alpha}} \Gamma^A \Gamma^B \left(\mathbb{1} + \Gamma_7 \right) \Gamma_A \boldsymbol{\Psi}^b \big) \left(\overline{\boldsymbol{\Psi}}^c \Gamma_B \boldsymbol{\Psi}^a \right)$$
$$- \frac{1}{48} e \epsilon_{abc} \left(\overline{\boldsymbol{\alpha}} \Gamma^A \Gamma^{BCD} \Gamma_A \Psi^b \right) \left(\bar{\Psi}^c \Gamma_{BCD} \Psi^a \right)$$

which using that Ψ is Weyl, can be simplified to

$$-\frac{1}{4} e \epsilon_{abc}) \, \overline{\boldsymbol{\alpha}} \Gamma^A \Gamma^B \Gamma_A \boldsymbol{\Psi}^b \big) \left(\overline{\boldsymbol{\Psi}}^c \Gamma_B \Psi^a \right)$$
$$- \frac{1}{48} e \epsilon_{abc} \left(\overline{\boldsymbol{\alpha}} \Gamma^A \Gamma^{BCD} \Gamma_A \Psi^b \right) \left(\bar{\Psi}^c \Gamma_{BCD} \Psi^a \right)$$

Exercise 2.10 (Some Γ-matrix identities)

Prove the following two identities:

$$\Gamma^A \Gamma^B \Gamma_A = -4\Gamma^B \quad \text{and} \quad \Gamma^A \Gamma^{BCD} \Gamma_A = 0 \tag{2.11}$$

and deduce that the trilinear terms cancel exactly. The above identities are in fact the minor miracle that makes supersymmetric Yang-Mills possible in six dimensions.

Up to a divergence, the remaining terms in the supersymmetric variation of the lagrangian density $\mathcal{L}$ are then:

$$\delta\mathcal{L} = -i\mathrm{G}^{AB} \cdot D_A \left(\overline{\boldsymbol{\alpha}} \Gamma_B \Psi\right) - \frac{i}{2} \mathrm{G}^{BC} \cdot \overline{\boldsymbol{\alpha}} \Gamma_{BC} \Gamma^A D_A \Psi$$

and

$$\bar{\delta}\mathcal{L} = i\mathrm{G}^{AB} \cdot D_A \left(\overline{\boldsymbol{\Psi}} \Gamma_B \boldsymbol{\beta}\right) - \frac{i}{2} \mathrm{G}_{BC} \cdot D_A \overline{\boldsymbol{\Psi}} \Gamma^A \Gamma^{BC} \boldsymbol{\beta}$$

Exercise 2.11 (... and the proof of supersymmetry invariance)
Prove the following identity between Dirac matrices

$$\Gamma_{AB} \Gamma_C = \Gamma_{ABC} + \eta_{BC} \Gamma_A - \eta_{AC} \Gamma_B$$

and use it to rewrite the supersymmetric variations of $\mathcal{L}$ as

$$\delta\mathcal{L} = \frac{i}{2} \partial_A \overline{\boldsymbol{\alpha}} \mathrm{G}^{BC} \cdot \Gamma_{BC} \Gamma^A \boldsymbol{\Psi} \quad \text{and} \quad \bar{\delta}\mathcal{L} = \frac{i}{2} \overline{\boldsymbol{\Psi}} \Gamma^A \Gamma_{BC} \cdot \mathrm{G}^{BC} \partial_A \boldsymbol{\beta} \tag{2.12}$$

again up to divergences and where we have used the Bianchi identity in the form $\Gamma_{ABC} D^A \mathrm{G}^{BC} = 0$. This proves the invariance of $\mathcal{L}$ under the supersymmetry transformations (2.8).

From (2.12) we can read the expression for the supersymmetry currents:

$$J^A = \frac{i}{2} \mathrm{G}^{BC} \cdot \Gamma_{BC} \Gamma^A \Psi \quad \text{and} \quad \bar{J}^A = \frac{i}{2} \overline{\boldsymbol{\Psi}} \Gamma^A \Gamma_{BC} \cdot \mathrm{G}^{BC}$$

As usual the spinorial supersymmetry charge is the space integral of the zero component of the current. Provided we already knew that $\mathcal{L}$ is supersymmetric, there is a more economical way to derive the expression of the supercurrent. This uses the fact that the supercurrent is part of a supersymmetry multiplet.

Exercise 2.12 (The supersymmetry multiplet)

Prove that the lagrangian density (2.7) is invariant under the transformation $\boldsymbol{\Psi} \mapsto \exp(i\theta)\boldsymbol{\Psi}, \overline{\boldsymbol{\Psi}} \mapsto \exp(-i\theta)\overline{\boldsymbol{\Psi}}$, and that the corresponding Noether current is given by $j_A = \overline{\boldsymbol{\Psi}} \cdot \Gamma_A \boldsymbol{\Psi}$. Prove that

$$\delta j_A = i\overline{\boldsymbol{\alpha}} J_A \quad \text{and} \quad \bar{\delta} j_A = -i\bar{J}_A \boldsymbol{\beta}$$

The supersymmetry multiplet also contains the energy-momentum tensor, alone or in combination with other topological currents that may appear in the right hand side of $\{Q, \bar{Q}\}$ in the supersymmetry algebra. We will use this later to compute the supersymmetry algebra corresponding to six-dimensional supersymmetric Yang-Mills. But first we perform the dimensional reduction to four dimensions.

8.3.2 2.3.2 From $N = 1$ in $d = 6$ to $N = 2$ in $d = 4$

Let us single out two of the coordinates (x^5, x^6) in six dimensions and assume that none of our fields depend on them: $\partial_5 \equiv \partial_6 \equiv 0$. This breaks $SO(1,5)$ Lorentz invariance down to $SO(1,3) \times SO(2)$. Let us therefore decompose our six-dimensional fields in a way that reflects this. In fact, we will at first ignore the $SO(2)$ invariance and focus only on the behaviour of the components of the six-dimensional fields under the action of $SO(1,3)$. The gauge field W_A breaks up into a vector W_μ, a pseudo-scalar $\mathrm{P} = \mathrm{W}_5$ and a scalar $S = W_6$. In terms of these fields, the field-strength breaks up as $G_{\mu\nu}$, $\mathrm{G}_{\mu 5} = D_\mu \mathrm{P}, \mathrm{G}_{\mu 6} = D_\mu \mathrm{S}$ and $\mathrm{G}_{56} = e\, \mathrm{S} \times \mathrm{P}$. Meanwhile, the Weyl spinor breaks up as $\Psi = \binom{\psi}{0}$, where ψ is an unconstrained (complex) Dirac spinor. The covariant derivative of the spinor then breaks up as $(D_\mu \boldsymbol{\psi}, -e\mathrm{P} \times \boldsymbol{\psi}, -e\, \mathrm{S} \times \boldsymbol{\psi})$.

The lagrangian density now becomes $\mathcal{L} = \mathcal{L}_\mathrm{b} + \mathcal{L}_\mathrm{f}$ where

$$\mathcal{L}_\mathrm{b} = -\frac{1}{4}\mathrm{G}_{\mu\nu} \cdot \mathrm{G}^{\mu\nu} + \frac{1}{2} D_\mu \mathrm{P} \cdot D^\mu \mathrm{P} + \frac{1}{2} D_\mu \mathrm{S} \cdot D^\mu \mathrm{S} - \frac{1}{2} e^2 \|\mathrm{P} \times \mathrm{S}\|^2$$

and

$$\mathcal{L}_\mathrm{f} = i\overline{\boldsymbol{\psi}} \cdot \gamma^\mu D_\mu \boldsymbol{\psi} + ie\overline{\boldsymbol{\psi}} \cdot \gamma_5 \mathrm{P} \times \boldsymbol{\psi} + ie\overline{\boldsymbol{\psi}} \cdot \mathrm{S} \times \boldsymbol{\psi} \tag{2.13}$$

where we see that P is indeed as pseudo-scalar as claimed. $\mathcal{L}$ is the lagrangian density of $N = 2$ supersymmetric Yang-Mills theory in four dimensions. The supersymmetry parameter $\boldsymbol{\alpha}$, which in the six-dimensional theory is a Weyl spinor, becomes upon dimensional reduction a Dirac spinor. But in four dimensions the supersymmetry parameters are Majorana, hence this gives rise to $N = 2$ supersymmetry. One can see this explicitly by breaking up the supersymmetry parameter into its Majorana components: simply choose a Majorana representation and split it into its real and imaginary parts. Each of these spinors is Majorana and generates one supersymmetry.

Let us first do this with $\boldsymbol{\psi}$. The next exercise shows the resulting fermion action in a Majorana basis.

Exercise 2.13 ($\mathcal{L}$ in a Majorana basis)

In a Majorana basis, let us split ψ as follows:

$$\psi = \frac{1}{\sqrt{2}} (\psi_1 - i\psi_2)$$

Prove that relative to $\boldsymbol{\psi}_\alpha, \alpha = 1, 2$, the fermionic part $\mathcal{L}_\mathrm{f}$ of the lagrangian density becomes (up to a total derivative)

$$\mathcal{L}_\mathrm{f} = \frac{i}{2}\overline{\boldsymbol{\psi}}_1 \cdot \gamma^\mu D_\mu \boldsymbol{\psi}_1 + \frac{i}{2}\overline{\boldsymbol{\psi}}_2 \cdot \gamma^\mu D_\mu \boldsymbol{\psi}_2 + e\overline{\boldsymbol{\psi}}_1 \cdot \gamma_5 \mathrm{P} \times \boldsymbol{\psi}_2 + e\overline{\boldsymbol{\psi}}_1 \cdot \mathrm{S} \times \boldsymbol{\psi}_2 \tag{2.14}$$

(Hint: you may find useful the following identities for anticommuting Majorana spinors in four dimensions:

$$\bar{\chi}\lambda = \bar{\lambda}\chi \quad \bar{\chi}\gamma_5\lambda = -\bar{\lambda}\gamma_5\chi \quad \bar{\chi}\gamma_\mu\lambda = -\bar{\lambda}\gamma_\mu\chi \tag{2.15}$$

which you are encouraged to prove!)

We can do the same with the supersymmetry transformations (2.8), as the next exercise asks you to show.

Exercise 2.14 (Explicit $N=2$ supersymmetry transformations)

Show that in a Majorana basis, the dimensional reduction of the supersymmetry transformations (2.8) becomes:

$$\begin{aligned}
\delta_1 \mathrm{W}_\mu &= i\bar{\alpha}\gamma_\mu\boldsymbol{\psi}_1 + \bar{\alpha}\gamma_\mu\boldsymbol{\psi}_2 \\
\delta_2 \mathrm{W}_\mu &= -\bar{\alpha}\gamma_\mu\boldsymbol{\psi}_1 + i\bar{\alpha}\gamma_\mu\boldsymbol{\psi}_2 \\
\delta_1\mathrm{P} &= i\bar{\alpha}\gamma_5\boldsymbol{\psi}_1 + \bar{\alpha}\gamma_5\boldsymbol{\psi}_2 \\
\delta_2\mathrm{P} &= -\bar{\alpha}\gamma_5\boldsymbol{\psi}_1 + i\bar{\alpha}\gamma_5\boldsymbol{\psi}_2 \\
\delta_1 \mathrm{S} &= i\bar{\alpha}\boldsymbol{\psi}_1 + \bar{\alpha}\psi_2 \\
\delta_2 \mathrm{S} &= -\bar{\alpha}\boldsymbol{\psi}_1 + i\bar{\alpha}\boldsymbol{\psi}_2 \\
\delta_1\boldsymbol{\psi}_1 &= -D^\mu\left(\mathrm{S}+\mathrm{P}\gamma_5\right)\gamma_\mu\alpha + \frac{1}{2}e(\mathrm{S}\times\mathrm{P})\gamma_5\alpha + \frac{1}{2}\mathrm{G}^{\mu\nu}\gamma_{\mu\nu}\alpha \\
\delta_1\psi_2 &= 0 \\
\delta_2\boldsymbol{\psi}_1 &= 0 \\
\delta_2\boldsymbol{\psi}_2 &= -D^\mu\left(\mathrm{S}+\mathrm{P}\gamma_5\right)\gamma_\mu\alpha + \frac{1}{2}e(\mathrm{S}\times\mathrm{P})\gamma_5\alpha + \frac{1}{2}\mathrm{G}^{\mu\nu}\gamma_{\mu\nu}\alpha
\end{aligned}$$

The $SO(2)$ invariance of (2.14) can be made manifest by rewriting $\mathcal{L}_\mathrm{f}$ explicitly in terms of the $SO(2)$ invariant tensors $\delta_{\alpha\beta}$ and $\epsilon_{\alpha\beta}$. In fact, using the identities (2.15), one can rewrite (2.14) as:

$$\mathcal{L}_\mathrm{f} = \frac{i}{2}\delta^{\alpha\beta}\overline{\boldsymbol{\psi}}_\alpha\cdot\gamma^\mu D_\mu\boldsymbol{\psi}_\beta + \frac{e}{2}\epsilon^{\alpha\beta}\overline{\boldsymbol{\psi}}_\alpha\cdot(\gamma_5\mathrm{P}+\mathrm{S})\times\boldsymbol{\psi}_\beta$$

The $SO(2)$ transformation properties of the four-dimensional fields can be succinctly written as follows:

$$\begin{aligned}
\mathrm{S}+i\mathrm{P} &\mapsto e^{-i\mu}(\mathrm{S}+i\mathrm{P}) \\
\psi &\mapsto e^{\mu\gamma_5/2}\psi \\
\mathrm{W}_\mu &\mapsto \mathrm{W}_\mu
\end{aligned} \tag{2.16}$$

Exercise 2.15 (The $SO(2)$ Noether current)

Prove that the Noether current associated with the $SO(2)$ transformations (2.16) is given by

$$j_\mu^5 = \mathrm{P}\cdot D_\mu\mathrm{S} - \mathrm{S}\cdot D_\mu\mathrm{P} + \frac{i}{2}\bar{\psi}\cdot\gamma_5\gamma_\mu\psi$$

Notice that this current contains the axial current, hence the notation.

Problem: Is it anomalous in this theory?

8.3.3 2.3.3 Higgsed $N=2$ supersymmetric Yang-Mills

The hamiltonian density corresponding to the $N=2$ supersymmetric YangMills theory defined by (2.13) is given by $\mathcal{H}=\mathcal{H}_\mathrm{b}+\mathcal{H}_\mathrm{f}$. We focus on the bosonic part:

$$\begin{aligned}
\mathcal{H}_\mathrm{b} = \frac{1}{2}\left\|\mathrm{E}_i\right\|^2 + \frac{1}{2}\left\|D_0\,\mathrm{S}\right\|^2 + \frac{1}{2}\left\|D_0\mathrm{P}\right\|^2 \\
+ \frac{1}{2}\left\|\,\mathrm{B}_i\right\|^2 + \frac{1}{2}\left\|D_i\,\mathrm{S}\right\|^2 + \frac{1}{2}\left\|D_i\mathrm{P}\right\|^2 + \frac{1}{2}e^2\|\mathrm{P}\times\mathrm{S}\|^2
\end{aligned}$$

Demanding that the energy of a given field configuration be finite doesn't necessarily imply that P and S acquire non-zero vacuum expectation valuesfor the term $\|\mathrm{P}\times\mathrm{S}\|^2$ is already zero provided that $\mathrm{P} \times \mathrm{S} = 0$, which for so (3) means that they be parallel. Indeed, except for that term and the extra field, $\mathcal{H}_\mathrm{b}$ is nothing but the energy density (1.13) of (the bosonic part of) the Georgi-Glashow model in the limit of vanishing potential. We could add a potential term $\lambda\left(\|\mathrm{P}\|^2 + \|\mathrm{S}\|^2 - a^2\right)^2$ to the lagrangian (2.13) to force S and P to acquire a nonzero vacuum expectation value, but such a term would break supersymmetry. Nevertheless we could then take the limit $\lambda \downarrow 0$ while keeping the nonzero vacuum expectation values of S and P. This restores the supersymmetry provided that $\langle \mathrm{S} \rangle$ and $\langle \mathrm{P} \rangle$ are parallel, which would be the supersymmetric version of the Prasad-Sommerfield limit. Since the potential depends only on the $SO(2)$ invariant combination $\|\mathrm{P}\|^2 + \|\mathrm{S}\|^2, SO(2)$ is preserved and we could use this symmetry to choose $\langle \mathrm{P} \rangle = 0$ and $\langle \mathrm{S} \rangle = \boldsymbol{a}$, where $\boldsymbol{a}$ is a fixed vector with $\|\boldsymbol{a}\|^2 = a^2$.

Exercise 2.16 (The perturbative spectrum of the model)

We can analyse the perturbative spectrum of the model around such a vacuum in exactly the same way as we did in Exercise 1.4. Choosing for example the unitary gauge $\boldsymbol{a} = a\boldsymbol{e}_3$, show that there are now two massive multiplets $(\psi^{\pm}, W_\mu^{\pm}, P^{\pm})$ of mass $M_W = ae\lambda$, and a massless gauge multiplet corresponding to the unbroken $U(1) : \left(\psi^3, W_\mu^3, S^3, P^3\right)$. Prove that the massless gauge multiplet is actually made out of two massless multiplets with helicities $\lambda = -1$ and $\lambda = 0$.

Now watch carefully: something curious has happened. From the analysis in section 2.2.3, we know that the generic massive representations of $N = 2$
supersymmetry are sixteen-fold degenerate, and from Exercise 2.3 we know that they must have a state with spin 3/2. Yet the massive multiplets which have arisen out of higgsing the model contain maximum spin 1 and are only four-fold degenerate. This is only possible if the $N = 2$ supersymmetry algebra in this model has central charges and these charges saturate the bound! Indeed, the only way to reconcile the above spectrum with the structure of massive representations of the $N = 2$ supersymmetry algebra studied in section 2.2.3 is if it corresponds to the short multiplet with spin $s = 1/2$ studied in Exercise 2.5. In the next section we will actually compute the supersymmetry algebra for this model and we will see that the central charges are precisely the electric and magnetic charges relative to the unbroken $U(1)$. But before doing this let us check that the BPS-monopole is actually a solution of $N = 2$ supersymmetric Yang-Mills.

8.3.4 2.3.4 $N = 2$ avatar of the BPS-monopole

We now show that this $N = 2$ supersymmetric Yang-Mills theory admits BPSmonopole solutions. We look for static solutions, so we put $\mathrm{W}_0 = 0$. Since the fermion equations of motion are linear, we can always put $\boldsymbol{\psi} = 0$ at the start. Applying supersymmetry transformations to such a solution, we will be able to generate solutions with nonzero fermions. Similarly, using the $SO(2)$ invariance we can look for a solution with $\mathrm{P} = 0$, and then obtain solutions with nonzero P by acting with $SO(2)$. Having made these choices, we are left with W_i and S, which is precisely the spectrum of the bosonic part of the Georgi-Glashow model provided we identify S and ϕ. Furthermore, not just the spectrum, but also the lagrangian density agrees, with potential set to zero, of course. Therefore the BPS-monopole given by (1.15) with H and K given in Exercise 1.12 is a solution of $N = 2$ supersymmetric Yang-Mills. If we now apply an $SO(2)$ rotation to this solution, we find the following BPS-monopole solution:

$$
\begin{aligned}
\psi &= \mathrm{W}_0 = 0 \\
\mathrm{S}^a &= \alpha \frac{r^a}{er^2} H(\xi) \\
\mathrm{P}^a &= \beta \frac{r^a}{er^2} H(\xi) \\
\mathrm{W}_i^a &= \epsilon_{aij} \frac{r^j}{er^2} (K(\xi) - 1)
\end{aligned} \tag{2.17}
$$

where as before $\xi = aer$, where H and K are the same functions in Exercise 1.12, and $\alpha^2 + \beta^2 = 1$. Putting $\beta = 0$ we recover the BPS-monopole and anti-monopole for $\alpha = \pm 1$, respectively-a result first obtained in DHdV78.

Since (2.17) is a solution of the field equations of a supersymmetric theory, supersymmetry transformations map solutions to solutions. Hence starting with (2.17) we can try to generate solutions with nonzero fermions by performing a supersymmetry transformation. We will actually assume a more general solution than the one above.

Exercise 2.17 (Supersymmetric BPS-monopoles)

Prove that any BPS-monopole, that is, any static solution (W_i, ϕ) of the Bogomol'nyi equation (1.28), can be thought of as an $N=2$ BPS-monopole by setting $\mathrm{S} = \alpha \boldsymbol{\phi}, \mathrm{P} = \beta \phi$ and $\boldsymbol{\psi} = 0$, with $\alpha^2 + \beta^2 = 1$.

We will then take one such $N=2$ BPS-monopole as our starting point and try to generate other solutions via supersymmetry transformations. The supersymmetry transformation laws on the four-dimensional fields can be read off from those given in (2.8) for the six-dimensional fields. Since we start with a background in which $\psi = 0$, the bosonic fields are invariant under supersymmetry. The supersymmetry transformation law of the fermion $\boldsymbol{\psi}$ is given by

$$
\delta \boldsymbol{\psi} = \left(\frac{1}{2} G^{\mu\nu} \gamma_{\mu\nu} - D_\mu \boldsymbol{\phi} \gamma^\mu (\alpha + \beta \gamma_5) \right) \epsilon
$$

where ϵ is an unconstrained (complex) Dirac spinor. Because the solution is static - $\mathrm{W}_0 = 0$ and all fields are time-independent-the above can be rewritten as

$$
\delta \boldsymbol{\psi} = \left(\frac{1}{2} G_{ij} \gamma_{ij} + D_i \phi \gamma_i (\alpha + \beta \gamma_5) \right) \epsilon
$$

For definiteness we will assume that $(\mathrm{W}_i, \boldsymbol{\phi})$ describe a BPS-monopole (as opposed to an anti-monopole) so that $D_i \phi = +\frac{1}{2} \epsilon_{ijk} G_{jk}$. Then we can rewrite the above transformation law once more as

$$
\delta \boldsymbol{\psi} = D_k \boldsymbol{\phi} \left(\frac{1}{2} \epsilon_{ijk} \gamma_{ij} + \gamma_k (\alpha + \beta \gamma_5) \right) \epsilon \tag{2.18}
$$

Exercise

2.18 (More γ-matrix identities)

Prove the following identity:

$$
\frac{1}{2} \epsilon_{ijk} \gamma_{ij} = -\gamma_0 \gamma_5 \gamma_k \tag{2.19}
$$

Exercise 2.19 (Some euclidean γ-matrices)

Let $\bar{\gamma}_i \equiv \gamma_0\gamma_i$ for $i = 1, 2, 3$, and let $\bar{\gamma}_4 = \gamma_0 (\alpha + \beta\gamma_5)$. Prove that they generate a euclidean Clifford algebra. Define $\bar{\gamma}_5 \equiv \bar{\gamma}_1\bar{\gamma}_2\bar{\gamma}_3\bar{\gamma}_4 = \gamma_0 (\alpha\gamma_5 - \beta)$. Prove that $\bar{\gamma}_5$ is hermitian and that $\bar{\gamma}_5^2 = \mathbb{1}$.

In terms of these euclidean Clifford algebra, and using (2.19), we can rewrite (2.18) as

$$\delta\boldsymbol{\psi} = \gamma_5\bar{\gamma}_k D_k\boldsymbol{\phi} (\mathbb{1} - \bar{\gamma}_5) \epsilon$$

Notice that $\frac{1}{2} (\mathbb{1} \pm \bar{\gamma}_5)$ is a projector. If we denote $\epsilon_\pm = \frac{1}{2} (\epsilon \pm \bar{\gamma}_5\epsilon)$, then the supersymmetric variation of $\boldsymbol{\psi}$ in a BPS-monopole background is given simply by

$$\delta\boldsymbol{\psi} = 2\gamma_5\bar{\gamma}_k D_k\boldsymbol{\phi}\epsilon_-$$

This means that if ϵ has negative chirality relative to $\bar{\gamma}_5$, then we don't generate new solutions, yet if ϵ has positive chirality, then we do. Equivalently, supersymmetry transformations with negative chirality parameters preserve the solution, whereas those with positive chirality parameters break it.

Exercise

2.20 (BPS-monopoles break one half of the supersymmetry) Prove that the (± 1)-eigenspaces of $\bar{\gamma}_5$ have the same dimension. Conclude that the projector $\frac{1}{2} (\mathbb{1} \pm \bar{\gamma}_5)$ projects out precisely one half of spinors.

As a corollary of the above exercise we see that supersymmetric BPSmonopoles break half the supersymmetries.

Notice that the parameter ϵ, being an unconstrained Dirac spinor has 4 complex (or 8 real) components, whereas $\epsilon_\pm$only has 2 complex (or 4 real) components. Hence we expect that the BPS-monopole belongs to a fourfold degenerate multiplet. From our study in section 2.2.3 of massive representations of the $N = 2$ supersymmetry algebra, we know that those massive multiplets preserving half the supersymmetries are necessarily short, and from Exercise 2.5 we see that the $k = 1$ BPS-monopole given by (2.17) generates a short multiple with spin $s = 0$. This multiplet contains two "particles" of spin 0 and one of spin 1/2, yet none of spin 1. Therefore although as we will see in the next section, $N = 2$ supersymmetry solves the first of the problems with the Montonen-Olive conjecture mentioned at the end of section 1.4.1, it still does not address the second problem satisfactorily. As we will see later, the solution of this problem requires $N = 4$ supersymmetry.

8.3.5 2.3.5 The supersymmetry bound is the Bogomol'nyi bound

The Bogomol'nyi bound (1.27) can be suggestively rewritten as

$$M^2 - (aq)^2 - (ag)^2 \geq 0$$

which is begging us to add two spatial dimensions to our spacetime and interpret the above inequality as the positivity of mass. As explained in section 2.2.1, the positivity of the mass is a consequence of unitarity and the supersymmetry algebra. Therefore it would make sense to look for a six-dimensional supersymmetric explanation of the Bogomol'nyi bound. The explanation of WO78 used the central charges in four-dimensional $N = 2$
supersymmetric Yang-Mills theory, and as we have seen this theory comes induced from six-dimensional $N = 1$ supersymmetric Yang-Mills via dimensional reduction. It would make sense therefore to look for a direct sixdimensional explanation. This was done to a large extent in Oli79 and we will now review this.

The above heuristics suggest that we think of the electric and magnetic charges as momenta in the two extra spatial dimensions. However it isn't hard to see that this interpretation is not quite correct. If one computes the energy-momentum tensor T_{AB} of the six-dimensional supersymmetric YangMills theory, and from there the momenta $P_A = T_{0A}$, then the positivity of mass formula in six-dimensions:

$$M^2 \geq P_5^2 + P_6^2 \tag{2.20}$$

where $M^2 = P^\mu P_\mu$ is the four-dimensional mass, does not agree with the Bogomol'nyi bound (1.27). In fact one finds that the magnetic charge does not appear. What is wrong then? Simply that we have assumed that it is P_μ which appears in the right hand side of $\{Q, \bar{Q}\}$ in the supersymmetry algebra, when in fact it is $P_\mu + Z_\mu$, where Z_μ can be interpreted as the topological charge due to the presence of a string-like source in six-dimensions. We now find out what Z_μ is by computing the supersymmetry algebra. We first do this in six dimensions and then reduce down to four.

The supersymmetry algebra in six dimensions

We start by noticing that the space integral of $\delta\bar{\delta}j_0$ is equal to $\overline{\boldsymbol{\alpha}}\{Q, \bar{Q}\}\boldsymbol{\beta}$, whence it is enough to compute $\delta\bar{\delta}j_A$, which we naturally leave as an exercise.

Exercise 2.21 (Supersymmetric variation of the supercurrent)

Prove that

$$\begin{aligned}\delta\bar{\delta}j_A &= -i\delta\bar{J}_A\boldsymbol{\beta}\\ &= -\frac{1}{4}\mathrm{G}^{BC}\cdot\mathrm{G}^{EF}\overline{\boldsymbol{\alpha}}\Gamma_{BC}\Gamma_A\Gamma_{EF}\boldsymbol{\beta} + i\left(\overline{\boldsymbol{\alpha}}\Gamma_C D_B\boldsymbol{\Psi}\right)\cdot\left(\bar{\Psi}\Gamma_A\Gamma^{BC}\boldsymbol{\beta}\right)\end{aligned}$$

The fermion bilinear term has to be Fierzed, but we will not be concerned with the fermions in what follows: we are interested in computing the "momenta" in classical configurations like the BPS-monopole, where the fermions have been set to zero. Of course, it would be a good exercise in Γ-matrix algebra to compute the fermionic terms, not that there is little Γ-matrix algebra to be done. In fact, prove that setting $\Psi = 0, \delta\bar{\delta}j_A$ is given by

$$\delta\bar{\delta}j_A = 2\overline{\boldsymbol{\alpha}}\left(-\frac{1}{8}\epsilon_{BCAEFD}\mathrm{G}^{BC}\cdot\mathrm{G}^{EF} + \mathrm{G}^{BC}\cdot\mathrm{G}_{CA}\eta_{BD} + \frac{1}{4}\mathrm{G}^{BC}\cdot\mathrm{G}_{BC}\eta_{AD}\right)\Gamma^D\boldsymbol{\beta} \tag{2.21}$$

(Hint: use that $\Gamma_{ABCDE} = -\epsilon_{ABCDEF}\Gamma^F\Gamma_7$ (prove it!) and use the fact that $\Gamma_7\boldsymbol{\beta} = -\boldsymbol{\beta}$.)
We see that there are two very different tensors appearing in the righthand-side of $\delta\bar{\delta}j_A$:

$$\Theta_{AB} = -\frac{1}{8}\epsilon_{ABCDEF}\mathrm{G}^{CD}\cdot\mathrm{G}^{EF} \tag{2.22}$$

$$T_{AB} = \mathrm{G}_A{}^C\cdot\mathrm{G}_{CB} + \frac{1}{4}\mathrm{G}^{CD}\cdot\mathrm{G}_{CD}\eta_{AB} \tag{2.23}$$

Notice that T_{AB} is symmetric, whereas Θ_{AB} is antisymmetric. In fact, T_{AB} is (the bosonic part of) the energy-momentum tensor of the six-dimensional theory.

Exercise

2.22 (The symmetric gauge-invariant energy momentum tensor) Prove that the energy-momentum tensor of the six-dimensional supersymmetric Yang-Mills theory is given by

$$T_{AB} + \frac{i}{2}\bar{\Psi} \cdot \Gamma_{(A} \overleftrightarrow{D}_{B)} \Psi - \eta_{AB} \frac{i}{2} \bar{\Psi} \cdot \Gamma^{C} \overleftrightarrow{D}_{C} \boldsymbol{\Psi}$$

Prove that T_{AB} is gauge-invariant and that it is conserved on-shell.

(Hint: Vary $\mathcal{L}_{\mathrm{b}}$ with respect to an infinitesimal translation $x^A \mapsto x^A + \varepsilon^A(x)$ and determine the associated Noether current, which after symmetrisation is the energy-momentum tensor, by definition.)

Defining P_A to be the space integral of T_{0A} and Z_A the space integral of Θ_{0A}, we see that the supersymmetry algebra becomes $\{Q, \bar{Q}\}$:

$$\{Q, \bar{Q}\} = 2\Gamma^A \left(P_A + Z_A\right)$$

Only the first of these terms is to be interpreted as the momentum, the other term is associated with a topologically conserved current.

Exercise 2.23 (The topological current)

Prove that Θ_{AB} is gauge invariant and that it is conserved off-shell, that is, without imposing the equations of motion. This means that it is a topological current.

(Hint: show that $\Theta_{AB} = \partial^C \Xi_{ABC}$ where Ξ_{ABC} is totally antisymmetric, though not gauge invariant.)

The supersymmetry algebra in four dimensions

It is now time to dimensionally reduce the supersymmetry algebra. The following exercise asks you to compute P_5, P_6, Z_5, and Z_6 (with fermions put to zero) after dimensional reduction.

Exercise 2.24 (The "momenta" in the extra dimensions)

Prove that

$$\begin{aligned}
T_{05} &= -D_i \mathrm{P} \cdot \mathrm{G}_{0i} - e(\mathrm{P} \times \mathrm{S}) \cdot D_0 \mathrm{~S} \\
T_{06} &= -D_i \mathrm{~S} \cdot \mathrm{G}_{0i} + e(\mathrm{P} \times \mathrm{S}) \cdot D_0 \mathrm{P} \\
\Theta_{05} &= \frac{1}{2} \epsilon_{ijk} \mathrm{G}_{ij} \cdot D_k \mathrm{~S} \\
\Theta_{05} &= -\frac{1}{2} \epsilon_{ijk} \mathrm{G}_{ij} \cdot D_k \mathrm{P}
\end{aligned}$$

Using the Bianchi identity $\epsilon_{ijk} D_i G_{jk} = 0$, we can rewrite Θ_{05} and Θ_{06} as follows:

$$\Theta_{05} = \frac{1}{2} \partial_i \left(\epsilon_{ijk} \mathrm{G}_{jk} \cdot \mathrm{S}\right) \quad \text{and} \quad \Theta_{05} = -\frac{1}{2} \partial_i \left(\epsilon_{ijk} \mathrm{G}_{jk} \cdot \mathrm{P}\right)$$

whereas using the equations of motion (for zero fermions)

$$-D_i \mathrm{G}_{0i} + e \mathrm{P} \times D_0 \mathrm{P} + e \mathrm{~S} \times D_0 \mathrm{~S} = 0$$

we can rewrite T_{05} and T_{06} as follows:

$$T_{05} = -\partial_i \left(\mathrm{G}_{0i} \cdot \mathrm{P}\right) \quad \text{and} \quad T_{05} = -\partial_i \left(\mathrm{G}_{0i} \cdot \mathrm{S}\right)$$

We see that all the densities are divergences, whence their space integrals only receive contribution from spatial infinity:

$$P_5 + Z_5 = \int_{\Sigma_\infty} \left(-\mathrm{P} \cdot \mathrm{G}_{0i} + \frac{1}{2}\epsilon_{ijk}\ \mathrm{S} \cdot \mathrm{G}_{jk} \right) d\Sigma_i$$
$$P_6 + Z_6 = \int_{\Sigma_\infty} \left(-\mathrm{S} \cdot \mathrm{G}_{0i} - \frac{1}{2}\epsilon_{ijk}\mathrm{P} \cdot \mathrm{G}_{jk} \right) d\Sigma_i$$

To interpret these integrals we can proceed in either of two ways. The fastest way is to use the $SO(2)$ invariance of the theory to choose $\mathrm{P} = 0$ and $\|\mathrm{S}\|^2 = a^2$ at spatial infinity. Comparing with (1.24) and (1.25), we see that $P_5 + Z_5 = ag$ and $P_6 + Z_6 = -aq$. The same reasoning follows without having to use $SO(2)$ invariance, as the next exercise shows.

Exercise 2.25 (The effective electromagnetic field strength)

Define the following field strength:

$$F_{\mu\nu} \equiv \frac{1}{a}\left(\mathrm{S} \cdot \mathrm{G}_{\mu\nu} + \mathrm{P} \cdot {}^\star\mathrm{G}_{\mu\nu}\right) \tag{2.24}$$

Prove that in the "Higgs vacuum" it obeys Maxwell's equations, and deduce that $P_5 + Z_5 = ag$ and $P_6 + Z_6 = -aq$ where g and q are, respectively, the magnetic and electric charges of this electromagnetic field.

(Hint: Compare with Exercise 1.8).

To prove that (2.20) is the Bogomol'nyi bound (1.27), we can proceed in two ways. We can exploit the $SO(2)$ invariance of the supersymmetry algebra in order to set $\mathrm{P} = 0$, and then notice that $Z_\mu = 0$. Using the fact that P_μ is indeed the honest momentum of the theory, namely the space integral of $T_{0\mu}$, and plugging the expressions for $P_A + Z_A$ into (2.20), we finally arrive at the Bogomol'nyi bound (1.27)!

Alternatively we can deduce that $Z_\mu = 0$ without having to set $\mathrm{P} = 0$. This is the purpose of the following exercise.

Exercise

2.26 (The space components of the topological charge)

Prove that Θ_{0i} is given by

$$\Theta_{0i} = \epsilon_{ijk}\partial_j \left(\mathrm{P} \cdot D_k\ \mathrm{S}\right)$$

whence Z_i is given by

$$Z_i = \epsilon_{ijk} \int_{\Sigma_\infty} \left(\mathrm{P} \cdot D_k\ \mathrm{S}\right) d\Sigma_j$$

Prove that this vanishes for a finite-energy configuration.

(Hint: Notice that for a solution of the Bogomol'nyi equation $\mathrm{S} = \alpha\phi, \mathrm{P} = \beta\phi$ with $\alpha^2 + \beta^2 = 1, \mathrm{P} \cdot D_k\ \mathrm{S} = \frac{1}{2}\alpha\beta\partial_k\|\boldsymbol{\phi}\|^2$, and that the derivative ∂_k is tangential to Σ_∞ due to the ϵ_{ijk}. Since $\|\phi\|^2 = a^2$ on Σ_∞, its tangential derivative vanishes.)

Define the following complex linear combinations of fields (cf. 1.33)):

$$\mathcal{G}_{\mu\nu} = \mathrm{G}_{\mu\nu} + i^\star\mathrm{G}_{\mu\nu}$$
$$\boldsymbol{\Phi} = \mathrm{S} + i\mathrm{P}$$

in terms of which the effective electromagnetic field strength defined in 2.24, becomes

$$F_{\mu\nu} = \frac{1}{a}\,\mathrm{Re}\left(\overline{\mathbf{\Phi}}\cdot\mathcal{G}_{\mu\nu}\right)$$

Under an infinitesimal $SO(2)$ transformation, $\delta\overline{\mathbf{\Phi}} = i\overline{\mathbf{\Phi}}$, and because $i\mathcal{S}_{\mu\nu} = {}^\star\mathcal{G}_{\mu\nu}$, we can write

$$\delta F_{\mu\nu} = -{}^\star F_{\mu\nu}$$

In other words, $SO(2)$ transformations become infinitesimal duality transformations in the effective electromagnetic theory.

Problem: Are anomalies responsible for the breaking of this symmetry in the quantum theory?

8.4 2.4 $N = 4$ Supersymmetric Yang-Mills

We saw in Exercise 2.7 that 10 is the largest dimension in which $N = 1$ supersymmetric Yang-Mills theory can exist and that for it to exist we must impose that the spinors be both Weyl and Majorana-conditions which, luckily for us, can be simultaneously satisfied in ten-dimensional Minkowski space. In
this section we will prove that this theory exists and that upon dimensional reduction to four dimensions yields a gauge theory with $N = 4$ supersymmetry. This theory admits Higgs phenomena and has room to embed the BPS-monopole and indeed, any solution of the Bogomol'nyi equation, just as in the $N = 2$ theory discussed in the previous section. We will see that both the massive fundamental states (e.g., vector bosons) and the solitonic states (e.g., BPS-monopoles) belong to isomorphic (short) multiplets saturating the supersymmetry mass bound which once again will be shown to agree with the Bogomol'nyi bound for dyons.

8.4.1 2.4.1 $N = 1$ $d = 10$ supersymmetric Yang-Mills

We start by setting up some conventions. We will let indices $A, B, \ldots$ from the start of the Latin alphabet run from 0 to 9. (No confusion should arise from the fact that in the previous section the very same indices only reached 6.) The metric η_{AB} is mostly minus and the 32×32 matrices $\{\Gamma_A\}$ obey the Clifford algebra $\{\Gamma_A, \Gamma_B\} = 2\eta_{AB}\mathbb{1}$. We let $\Gamma_{11} \equiv \Gamma_0\Gamma_1\cdots\Gamma_9$; it obeys $\Gamma_{11}^2 = \mathbb{1}$. We shall also need the charge conjugation matrix $\mathcal{C}$, which obeys $\mathcal{C}^t = -\mathcal{C}$ and $(\mathcal{C}\Gamma_A)^t = \mathcal{C}\Gamma_A$, from where it follows that $(\Gamma_A)^t = -\mathcal{C}\Gamma_A\mathcal{C}^{-1}$.

$N = 1$ supersymmetric Yang-Mills theory is defined by the following lagrangian density:

$$\mathcal{L} = -\frac{1}{4}\mathrm{G}^{AB}\cdot\mathrm{G}_{AB} + \frac{i}{2}\bar{\Psi}\cdot\Gamma^A D_A\Psi \tag{2.25}$$

where

$$\mathrm{G}_{AB} = \partial_A\,\mathrm{W}_B - \partial_B\,\mathrm{W}_A - e\,\mathrm{W}_A\times\mathrm{W}_B$$
$$D_A\Psi = \partial_A\Psi - e\,\mathrm{W}_A\times\Psi$$

and where $\mathbf{\Psi}$ is a complex Majorana-Weyl spinor obeying $\Gamma_{11}\mathbf{\Psi} = -\mathbf{\Psi}$ (Weyl) and $\overline{\mathbf{\Psi}} = \mathbf{\Psi}^\dagger\Gamma_0 = \mathbf{\Psi}^t\mathcal{C}$ (Majorana).

Exercise 2.27 (The Majorana condition)

Prove that the Majorana condition above relates $\mathbf{\Psi}$ and its complex conjugate $\mathbf{\Psi}^*$:

$$\Psi^* = \mathcal{C}\Gamma_0\Psi$$

whence it can be considered a reality condition on the spinor.

The action defined above is clearly gauge invariant. We claim that it is also invariant under the following supersymmetry:

$$\begin{aligned} \delta \mathrm{W}_A &= i\overline{\boldsymbol{\alpha}}\Gamma_A\Psi = -i\overline{\boldsymbol{\Psi}}\Gamma_A\boldsymbol{\alpha} \\ \delta\boldsymbol{\Psi} &= \frac{1}{2}\mathrm{G}^{AB}\Gamma_{AB}\boldsymbol{\alpha} \\ \delta\bar{\Psi} &= -\frac{1}{2}\overline{\boldsymbol{\alpha}}\Gamma_{AB}\mathrm{G}^{AB} \end{aligned} \tag{2.26}$$

where $\boldsymbol{\alpha}$ is a constant anticommuting Majorana-Weyl spinor of the same chirality as Ψ, and where the second and third relations above imply each other.

The proof that the action is invariant under the supersymmetry transformations (2.26) is very similar to the analogous statement for the sixdimensional theory, so we will not be as verbose.

We start by varying the action with respect to (2.26). We don't take $\boldsymbol{\alpha}$ to be constant in order to be able to read off the form of the supersymmetry current from the variation of the lagrangian density. We will have proven invariance if we can show that up to a divergence, the variation of (2.25) is proportional to the derivative of $\overline{\boldsymbol{\alpha}}$-the coefficient being the supersymmetry current. Varying the lagrangian density we encounter two kinds of terms: terms linear in the fermions, and a term trilinear in the fermions and without derivatives, coming from the variation of the gauge field inside the covariant derivative acting on the fermions.

Before getting into the computation, it is useful to derive some properties of Majorana-Weyl fermions, which are left as an instructive exercise.

Exercise 2.28 (Properties of Majorana and Weyl fermions)

Let $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ be anticommuting Majorana fermions in ten dimensions. Prove that

$$\overline{\boldsymbol{\alpha}}\Gamma_{A_1A_2\cdots A_k}\boldsymbol{\beta} = (-)^{k(k+1)/2}\overline{\boldsymbol{\beta}}\Gamma_{A_1A_2\cdots A_k}\boldsymbol{\alpha} \tag{2.27}$$

If, in addition, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are Weyl and of the same chirality, then prove that

$$\overline{\boldsymbol{\alpha}} \text{ (even number of } \Gamma s) \ \boldsymbol{\beta} = 0.$$

(Hint: It may prove useful to first prove the identity

$$(\mathfrak{C}\Gamma_{A_1A_2\cdots A_k})^t = -(-)^{k(k+1)/2}\mathcal{C}\Gamma_{A_1A_2\cdots A_k} \tag{2.28}$$

which will play a role also later on.)

We now vary the lagrangian density.

Exercise 2.29 (Varying the lagrangian density)

Prove that supersymmetric variation of the lagrangian density $\mathcal{L}$ is given, up to a divergence, by:

$$\begin{aligned} \delta\mathcal{L} = \frac{i}{2}D_C\mathbf{G}_{AB}\cdot\overline{\boldsymbol{\alpha}}\Gamma^{AB}\Gamma^C\boldsymbol{\Psi} + iD^A\mathrm{G}_{AB}\cdot\overline{\boldsymbol{\alpha}}\Gamma^B\boldsymbol{\Psi} \\ + \frac{i}{2}\mathrm{G}_{AB}\cdot\partial_C\overline{\boldsymbol{\alpha}}\Gamma^{AB}\Gamma^C\boldsymbol{\Psi} + \frac{1}{2}e\overline{\boldsymbol{\Psi}}\Gamma^A\cdot((\overline{\boldsymbol{\alpha}}\Gamma_A\boldsymbol{\Psi})\times\boldsymbol{\Psi}) \end{aligned}$$

(Hint: Integrate by parts and use the identity (2.27) repeatedly.)

Using the Bianchi identity in the form $\Gamma^{ABC} D_C G_{AB} = 0$, it is easy to prove that the first two terms in the above expression for $\delta\mathcal{L}$ cancel out, leaving the trilinear terms and the term involving the supersymmetry current:

$$\delta\mathcal{L} = \partial_A \overline{\boldsymbol{\alpha}} J^A + \frac{1}{2} e \overline{\boldsymbol{\Psi}} \Gamma^A \cdot ((\overline{\boldsymbol{\alpha}} \Gamma_A \boldsymbol{\Psi}) \times \boldsymbol{\Psi})$$

where the supersymmetry current J^A is given by

$$J^A = \frac{i}{2} G_{BC} \cdot \Gamma^{BC} \Gamma^A \Psi \tag{2.29}$$

Finally we tackle the trilinear terms, which as usual are the trickier ones. Just as in the six-dimensional theory, their vanishing will be seen to be a property of some identities between the Γ-matrices. Writing the SO(3) indices explicitly, we find that these terms are given by

$$\frac{1}{2} e \epsilon_{abc} \overline{\boldsymbol{\alpha}} \Gamma_A \boldsymbol{\Psi}^a \overline{\boldsymbol{\Psi}}^c \Gamma^A \times \boldsymbol{\Psi}^b \tag{2.30}$$

and we once again must use the Fierz identities to expand the bi-spinor $\Psi^a \bar{\Psi}^c$.

Exercise 2.30 (A ten-dimensional Fierz identity)

Prove that

$$\begin{aligned}\boldsymbol{\Psi}^a \overline{\boldsymbol{\Psi}}^c = & -\frac{1}{32} \overline{\boldsymbol{\Psi}}^c \Gamma_A \boldsymbol{\Psi}^a \Gamma^A \left(\mathbb{1} + \Gamma_{11}\right) \\ & + \frac{1}{32 \cdot 3!} \overline{\boldsymbol{\Psi}}^c \Gamma_{ABC} \boldsymbol{\Psi}^a \Gamma^{ABC} \left(\mathbb{1} + \Gamma_{11}\right) - \frac{1}{32 \cdot 5!} \overline{\boldsymbol{\Psi}}^c \Gamma_{ABCDE} \boldsymbol{\Psi}^a \Gamma^{ABCDE}.\end{aligned}$$

Using the results of Exercise 2.28 - in particular equation (2.27)-and taking into account the antisymmetry of ϵ_{abc} we see that only the first and last terms on the right-hand side of the above Fierz identity contribute to (2.30).

Exercise 2.31 (Some more Γ-matrix identities)

Prove the following identities between ten-dimensional Γ-matrices:

$$\Gamma^A \Gamma^B \Gamma_A = -8 \Gamma^B \quad \text{and} \quad \Gamma^F \Gamma^{ABCDE} \Gamma_F = 0 \tag{2.31}$$

and use them to deduce that the trilinear terms (2.30) cancel exactly.
(Compare these identities with those in Exercise 2.10.)

8.4.2 2.4.2 Reduction to $d = 4 : N = 4$ supersymmetric YangMills

We now dimensionally reduce the $d = 10 \quad N = 1$ supersymmetric Yang-Mills theory described in the previous section down to four dimensions. From
now on we will let uppercase indices from the middle of the Latin alphabet: $I, J, K, \ldots$ run from 1 to 3 inclusive. It will be convenient to break up the ten-dimensional coordinates as $x^A = \left(x^\mu, x^{3+I}, x^{6+J}\right)$, and by dimensional reduction we simply mean that we drop the dependence of the fields on $\left(x^{3+I}, x^{6+J}\right) : \partial_{3+I} \equiv \partial_{6+J} \equiv 0$.

We also need to decompose the ten-dimensional Γ-matrices. This is done as follows:

$$\begin{aligned}\Gamma^\mu &= \gamma^\mu \otimes \mathbb{1}_4 \otimes \sigma_3 \quad\quad C = C \otimes \mathbb{1}_4 \otimes \mathbb{1}_2 \\ \Gamma^{3+I} &= \mathbb{1}_4 \otimes \alpha^I \otimes \sigma_1 \;\; \Gamma^{6+J} = i\gamma_5 \otimes \beta^J \otimes \sigma_3\end{aligned}$$

where C is the charge conjugation matrix in four-dimensional Minkowski space obeying $C^t = -C$ and $(C\gamma_\mu)^t = C\gamma_\mu$; and $\{\alpha^I\}$ and $\{\beta^J\}$ are 4×4 real antisymmetric matrices satisfying the following algebraic relations:

$$[\alpha^I, \alpha^J] = -2\epsilon^{IJK}\alpha^K \qquad \{\alpha^I, \alpha^J\} = -2\delta^{IJ}\mathbb{1}_4$$
$$[\beta^I, \beta^J] = -2\epsilon^{IJK}\beta^K \; \{\beta^I, \beta^J\} = -2\delta^{IJ}\mathbb{1}_4$$
$$[\alpha^I, \beta^J] = 0;$$

and where $\mathbb{1}_n$ denotes the $n \times n$ unit matrix. (From now on we will drop the subscript when the dimension is clear from the context.) In the above decomposition, Γ_{11} takes the form:

$$\Gamma_{11} = -\mathbb{1} \otimes \mathbb{1} \otimes \sigma_2$$

We can find an explicit realisation for the matrices α^I and β^J as follows. Because they are real antisymmetric 4×4 matrices, they belong to so(4). Their commutation relations say that they each generate an so(3) subalgebra and moreover that these two so(3) subalgebras commute. Happily $so(4) \cong so(3) \times so(3)$, so that all we have to find is an explicit realisation of this isomorphism. This is found as follows. We say that a matrix A in so(4) is self-dual (respectively antiselfdual), if its entries obey $A_{ij} = \frac{1}{2}\epsilon_{ijkl}A_{kl}$ (respectively, $A_{ij} = -\frac{1}{2}\epsilon_{ijkl}A_{kl}$). The next exercise asks you to show that the subspaces of $so(4)$ consisting of (anti)self-dual matrices define commuting subalgebras.

Exercise

2.32($so(4) \cong so(3) \times so(3)$ explicitly)

Prove that the commutator of two (anti)self-dual matrices in so(4) is (anti)selfdual, and that the commutator of a self-dual matrix and an antiselfdual matrix in so(4) vanishes.

(Hint: Either compute this directly or use the fact that the "duality" operation is so(4) invariant since ϵ_{ijkl} is an so(4)-invariant tensor, whence its eigenspaces are ideals.)

Using this result we can now find a explicit realisation for the α^I and the β^J: we simply find a basis for the (anti)self-dual matrices in $so(4)$. This is the purpose of the next exercise.

Exercise

2.33 (Explicit realisation for α^I and β^J)

Prove that a matrix A in so(4) is (anti)self-dual if its entries are related in the following way:

$$A_{12} = \pm A_{34} \quad A_{13} = \mp A_{24} \quad A_{14} = \pm A_{23}$$

where the top signs are for the self-dual case and the bottom signs for the antiselfdual case. Conclude that explicit bases for the (anti)self-dual matrices are given by:

$$e_1^+ = i\sigma_2 \otimes \mathbb{1} = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix} \quad e_1^- = \mathbb{1} \otimes i\sigma_2 = \begin{pmatrix} i\sigma_2 & 0 \\ 0 & i\sigma_2 \end{pmatrix}$$
$$e_2^+ = \sigma_3 \otimes i\sigma_2 = \begin{pmatrix} i\sigma_2 & 0 \\ 0 & -i\sigma_2 \end{pmatrix} \qquad e_2^- = i\sigma_2 \otimes \sigma_3 = \begin{pmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{pmatrix}$$
$$e_3^+ = \sigma_1 \otimes i\sigma_2 = \begin{pmatrix} 0 & i\sigma_2 \\ i\sigma_2 & 0 \end{pmatrix} \qquad e_3^- = i\sigma_2 \otimes \sigma_1 = \begin{pmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{pmatrix}$$

where the $\{e_I^+\}$are self-dual and the $\{e_I^-\}$are antiselfdual. Prove that $\alpha^I = e_I^\pm$ and $\beta^J = e_J^\mp$ is a valid realisation.

Exercise

2.34 (The fundamental representation of $su(4)$)

Either abstractly or using the above explicit realisation, prove that the fifteen (4×4)-matrices:

$$A^{IJ} = \epsilon^{IJK}\alpha^K \quad B^{IJ} = \epsilon^{IJK}\beta^K \quad C^{IJ} = i\left\{\alpha^I, \beta^J\right\} \tag{2.33}$$

are antihermitian and generate the su(4) Lie algebra. This is the fundamental representation of su(4).

The result of the above exercise and the above decomposition of the ten-dimensional Γ-matrices mean that we have broken up a ten-dimensional spinor index (running from 1 to 32) into three indices: a four-dimensional spinor index (running from 1 to 4), an internal su(4) index in the fundamental representation (i.e., also running from 1 to 4), and an internal $su(2)$ index also in the fundamental representation. We have chosen the above decomposition of the Γ-matrices because it possesses two immediate advantages:

Because of the form of Γ_{11}, a Weyl spinor in ten-dimensions gives rise to a unconstrained Dirac spinor in four-dimensions with values in the fundamental representation of $su(4)$. In other words, the chirality condition only affects the internal $su(2)$ space and does not constrain the other degrees of freedom; and

Because of the form of the charge conjugation matrix, the Majorana condition in ten dimensions becomes the Majorana condition in four dimensions.

Thus we see immediately that a Majorana-Weyl spinor in ten-dimensions yields a quartet of Majorana spinors in four-dimensions, or equivalently a Majorana spinor in four-dimensions with values in the fundamental representation of su(4).

This $su(4)$ is a "flavour" index of the four-dimensional theory; that is, $su(4)$ is a global symmetry of $N = 4$ supersymmetric Yang-Mills theory, not a gauge symmetry. Of course, this flavour symmetry is nothing but the residual Lorentz symmetry in ten-dimensions which upon dimensional reduction to four-dimensions breaks down to $SO(1,3) \times SO(6)$. The Lie algebras of $SO(6)$ and $SU(4)$ are isomorphic. In fact, $SU(4) \cong \mathrm{Spin}(6)$, the universal covering group of $SO(6)$; and the four dimensional representations of $SU(4)$ are precisely the spinorial representations of $Spin(6)$ under which the supersymmetric charges transform.

We now want to write down the four-dimensional action obtained by the above dimensional reduction. We define the scalar fields $\mathrm{S}_I = \mathrm{W}_{3+I}$ and pseudoscalar fields $\mathrm{P}_J = \mathrm{W}_{6+J}$. Together with the four-dimensional gauge fields W_μ they comprise the bosonic field content of the four-dimensional theory. As mentioned above, the chirality condition on a ten-dimensional spinor can be easily imposed. Let us write

$$\Psi = \boldsymbol{\psi} \otimes \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\end{pmatrix}$$

where $\boldsymbol{\psi}$ is a quartet of unconstrained Dirac spinors in four dimensions. From the form of Γ_{11} it is easy so see that $\Gamma_{11}\Psi = -\boldsymbol{\psi} \otimes \frac{1}{\sqrt{2}}\sigma_2\binom{1}{i} = -\boldsymbol{\Psi}$. The Majorana condition says that $\boldsymbol{\psi}$ is Majorana in four dimensions. Naturally, all fields are in the adjoint representation of the gauge group $SO(3)$.

In order to write down the action we need to dimensionally reduce the Dirac operator and the gauge field-strength. We find that G_{AB} breaks up as $\mathrm{G}_{\mu\nu}, \mathrm{G}_{\mu,3+I} = D_\mu \mathrm{S}_I, \mathrm{G}_{\mu,6+I} = D_\mu \mathrm{P}_J, \mathrm{G}_{3+I,3+J} = -e\ \mathrm{S}_I \times \mathrm{S}_J, \mathrm{G}_{3+I,6+J} = -e\ \mathrm{S}_I \times \mathrm{P}_J$, and $\mathrm{G}_{6+I,6+J} = -e\mathrm{P}_I \times \mathrm{P}_J$. This allows us to write the bosonic part of the lagrangian density immediately:

$$\begin{aligned}\mathcal{L}_{\mathrm{b}} = -\frac{1}{4}\mathrm{G}^{\mu\nu}\cdot\mathrm{G}_{\mu\nu} + \frac{1}{2}D_\mu\mathrm{S}_I\cdot D^\mu\mathrm{S}_I + \frac{1}{2}D_\mu\mathrm{P}_J\cdot D^\mu\mathrm{P}_J \\ -\frac{1}{4}e^2\parallel\mathrm{S}_I\times\mathrm{S}_J\|^2 - \frac{1}{4}e^2\,\|\mathrm{P}_I\times\mathrm{P}_J\|^2 - \frac{1}{2}e^2\parallel\mathrm{S}_I\times\mathrm{P}_J\|^2\end{aligned} \tag{2.34}$$

The fermionic part of the action requires a bit more work, but it is nevertheless straightforward, and is left as an exercise.

Exercise

2.35 (The fermionic terms in the lagrangian)

Using the explicit form of the Γ-matrices, prove that the term $\frac{i}{2}\overline{\boldsymbol{\Psi}}\cdot\Gamma^A D_A\boldsymbol{\Psi}$ in equation (2.25), becomes

$$\mathcal{L}_{\mathrm{f}} = \frac{i}{2}\overline{\boldsymbol{\psi}}\cdot\gamma^\mu D_\mu\boldsymbol{\psi} + \frac{e}{2}\overline{\boldsymbol{\psi}}\cdot\big(\big(\alpha^I\;\mathrm{S}_I + \beta^J\mathrm{P}_J\gamma_5\big)\times\boldsymbol{\psi}\big) \tag{2.35}$$

from where we see that indeed we were justified in calling S_I scalars and P_J pseudoscalars.

We now write down the supersymmetry transformations. In ten dimensions, the parameter of the supersymmetry transformation is a MajoranaWeyl spinor. As we have seen, upon dimensional reduction, such a spinor yields a quartet of Majorana spinors in four-dimensions. Therefore the fourdimensional theory will have $N = 4$ supersymmetry. Indeed the lagrangian density $\mathcal{L}_{\mathrm{b}} + \mathcal{L}_{\mathrm{f}}$, understood in four dimensions, defines $N = 4$ supersymmetric Yang-Mills theory.

Since the supersymmetry parameter $\boldsymbol{\alpha}$ is a Majorana-Weyl spinor obeying $\Gamma_{11}\boldsymbol{\alpha} = -\boldsymbol{\alpha}$, we can write it as $\boldsymbol{\alpha} = \boldsymbol{\epsilon}\otimes\frac{1}{\sqrt{2}}\binom{1}{i}$, where $\boldsymbol{\epsilon}$ is a quartet of four-dimensional anticommuting Majorana spinors.

Exercise 2.36 ($N = 4$ supersymmetry transformations)

Expand equation (2.26) in this reparametrisation to obtain the following supersymmetry transformations for the four-dimensional fields:

$$\begin{aligned}\delta\mathrm{W}_\mu &= i\overline{\boldsymbol{\epsilon}}\gamma_\mu\boldsymbol{\psi}\\ \delta\mathrm{S}_I &= \overline{\boldsymbol{\epsilon}}\alpha^I\boldsymbol{\psi}\\ \delta\mathrm{P}_J &= \overline{\boldsymbol{\epsilon}}\gamma_5\beta^J\boldsymbol{\psi}\\ \delta\boldsymbol{\psi} &= \frac{1}{2}\mathrm{G}^{\mu\nu}\gamma_{\mu\nu}\boldsymbol{\epsilon} + iD_\mu\mathrm{S}_I\gamma^\mu\alpha^I\boldsymbol{\epsilon} + iD_\mu\mathrm{P}_J\gamma^\mu\gamma_5\beta^J\boldsymbol{\epsilon}\\ &\quad - e\,(\,\mathrm{S}_I\times\mathrm{P}_J)\,\gamma_5\alpha^I\beta^J\boldsymbol{\epsilon} + \frac{1}{2}e\epsilon_{IJK}\,(\,\mathrm{S}_I\times\mathrm{S}_J)\,\alpha^K\boldsymbol{\epsilon} + \frac{1}{2}e\epsilon_{IJK}\,(\mathrm{P}_I\times\mathrm{P}_J)\,\beta^K\boldsymbol{\epsilon}\end{aligned}$$

Finally we have the $su(4)$ invariance of the action.

Exercise 2.37 (su(4) invariance)

Prove that, for every choice of constant parameters (a_{IJ}, b_{IJ}, c_{IJ}) where $a_{IJ} = -a_{JI}$ and $b_{IJ} = -b_{JI}$, the following transformations are a symmetry of $N = 4$ supersymmetric Yang-Mills theory:

$$
\begin{aligned}
\delta \mathrm{W}_\mu &= 0 \\
\delta \mathrm{S}_I &= 2a_{IJ}\ \mathrm{S}_J + 2c_{IJ}\mathrm{P}_J \\
\delta \mathrm{P}_I &= 2b_{IJ}\mathrm{P}_J - 2c_{JI}\ \mathrm{S}_J \\
\delta \boldsymbol{\psi} &= -\frac{1}{2}a_{IJ}A^{IJ}\boldsymbol{\psi} - \frac{1}{2}b_{IJ}B^{IJ}\boldsymbol{\psi} + \frac{i}{2}c_{IJ}C^{IJ}\gamma_5\boldsymbol{\psi}
\end{aligned}
$$

where A^{IJ}, B^{IJ} and C^{IJ} are the $su(4)$ generators in the fundamental representation given by equation (2.33).

(Hint: You may save some time by first showing that these transformations are induced from Lorentz transformations in ten dimensions, and then using the Lorentz invariance of the ten-dimensional action.)

8.4.3 2.4.3 Monopoles and gauge bosons in $N = 4$ supersymmetric Yang-Mills

In section 2.3.4, we saw how any BPS-monopole could be thought of as a solution to the equations of motion of $N = 2$ supersymmetric Yang-Mills by setting the fermions to zero and aligning the scalar fields properly. Moreover we saw that such solutions break one half of the supersymmetry, so that these $N = 2$ BPS-monopoles naturally belong to a short multiplet. In fact, they belong to the short multiplet with spin $s = 0$. On the other hand, we had seen in section 2.3.3 that after higgsing, the perturbative spectrum of the theory arranged itself in a massless vector multiplet corresponding to the unbroken $U(1)$ and massive short multiplets with spin $s = \frac{1}{2}$ containing the massive vector bosons. It therefore seemed unlikely that $N = 2$ super YangMills would be self-dual, since the perturbative spectrum of the dual theory (i.e., the monopoles) now live in a different supersymmetry multiplet. And in fact, we now know from the results of Seiberg and Witten, that this theory is not self-dual. In this section we will see that this obstacle is overcome in $N = 4$ supersymmetric Yang-Mills theory. The discussion is very similar to that of sections 2.3.3 and 2.3.4, with the important distinction that the short multiplet containing the BPS-monopole and the one containing the massive vector boson are now isomorphic, being the one with spin $s = 0$. This section and the next are based on the work of Osborn Osb79.

The bosonic part of the hamiltonian density corresponding to the $N = 4$ supersymmetric Yang-Mills theory defined by (2.34) and (2.35) is given by:

$$
\begin{aligned}
\mathscr{H}_{\mathrm{b}} = \tfrac{1}{2}\left\|\mathrm{E}_i\right\|^2 + \tfrac{1}{2}\left\|D_0\ \mathrm{S}_I\right\|^2 + \tfrac{1}{2}\left\|D_0\mathrm{P}_J\right\|^2 + \tfrac{1}{2}\left\|\ \mathrm{B}_i\right\|^2 + \tfrac{1}{2}\left\|D_i\ \mathrm{S}_I\right\|^2 + \tfrac{1}{2}\left\|D_i\mathrm{P}_J\right\|^2 \\
+\tfrac{1}{2}e^2\left\|\ \mathrm{S}_I \times \mathrm{P}_J\right\|^2 + \tfrac{1}{4}e^2\left\|\ \mathrm{S}_I \times \mathrm{S}_J\right\|^2 + \tfrac{1}{4}e^2\left\|\mathrm{P}_I \times \mathrm{P}_J\right\|^2
\end{aligned}
$$

Demanding that the energy of a given field configuration be finite doesn't necessarily imply that all the scalars P_I and S_J acquire non-zero vacuum expectation values at spatial infinity. Indeed, looking at the potential terms it is sufficient (for $so(3)$) that they be parallel. This defines the supersymmetric Prasad-Sommerfield limit as in $N = 2$. In more detail, we add a potential term $\lambda\left(\left\|\mathrm{S}_I\right\|^2 + \left\|\mathrm{P}_J\right\|^2 - a^2\right)^2$ to the lagrangian (2.34) to force S_I and P_J to acquire a nonzero vacuum expectation value, but since such a term would break supersymmetry, we take the limit $\lambda \downarrow 0$ while keeping the nonzero vacuum expectation values of S_I and P_J. This restores the supersymmetry provided that $\langle \mathrm{S}_I \rangle$ and $\langle \mathrm{P}_J \rangle$ are parallel. We could choose $\mathrm{S}_I = a_I\phi$ and $\mathrm{P}_J = b_J\phi$ where $\sum_I \left(a_I^2 + b_I^2\right) = 1$, and where $\langle \phi \rangle$ has length a at infinity. Since the potential depends only on the $SO(6)$ invariant combination $\left\|\mathrm{S}_I\right\|^2 + \left\|\mathrm{P}_J\right\|^2$, we could use this symmetry to choose, say, $b_J = a_2 = a_3 = 0, a_1 = 1$ and $\langle \boldsymbol{\phi} \rangle = \boldsymbol{a}$, where $\boldsymbol{a}$ is a fixed vector with $\|\boldsymbol{a}\|^2 = a^2$.

Exercise 2.38 (The perturbative spectrum after higgsing)

We can analyse the spectrum of the model around such a vacuum in exactly the same way as we did in Exercise 1.4. Choosing for example the unitary gauge $\boldsymbol{a} = a\boldsymbol{e}_3$, show that there is now a massless gauge multiplet with helicity $\lambda = -1$, corresponding to the unbroken $U(1)$: $\left(W_\mu^3, \boldsymbol{\psi}^3, S_I^3, P_J^3\right)$; and two massive multiplets $\left(\psi^\pm, W_\mu^\pm, P_I^\pm, S_{1,2}^\pm\right)$ of mass $M_W = \text{ae}\hbar$. Conclude that these massive multiplets are actually short multiplets of spin $s = 0$.

Now let's see to what kind of multiplets the $N = 4$ BPS-monopoles belong. Let $(\mathrm{W}_i, \boldsymbol{\phi})$ be a BPS-monopole and let us set $\mathrm{W}_0 = \boldsymbol{\psi} = 0$, $\mathrm{S}_I = a_I\boldsymbol{\phi}$, and $\mathrm{P}_J = b_J\phi$, where a_I and b_J are real numbers satisfying $\sum_I \left(a_I^2 + b_I^2\right) = 1$. Because the fermions are zero, only the bosonic part of the lagrangian is nonzero. Plugging in these field configurations into (2.34), we find

$$\mathcal{L} = -\frac{1}{4}G_{ij}G_{ij} - \frac{1}{2}\left\|D_i\phi\right\|^2$$

after using that the fields are static and that S_I and P_J are all collinear. But this is precisely the action for static solutions to the bosonic Yang-MillsHiggs theory, hence it is minimised by BPS-monopoles. Therefore the above field configurations minimise the equations of motion of $N = 4$ supersymmetric Yang-Mills. In other words, we have shown that any BPS-monopole can be embedded as a solution of the $N = 4$ supersymmetric Yang-Mills theory. (Compare with Exercise 2.17.)

Now we will prove that such a solution breaks one half of the supersymmetry and hence it lives in a short multiplet. Because the fermions are put to zero, the supersymmetry transformation of the bosonic fields is automatically zero. From the results of Exercise 2.36 we can read off the expression for the supersymmetry transformation of the spinors in this background:

$$\delta\boldsymbol{\psi} = \left(\frac{1}{2}G_{ij}\gamma_{ij} - iD_i\phi\gamma_i\left(a_I\alpha^I + b_J\beta^J\gamma_5\right)\right)\boldsymbol{\epsilon}$$

If we now use equation (2.19), and the Bogomol'nyi equation in the form $G_{ij} = \epsilon_{ijk}D_k\boldsymbol{\phi}$, $\delta\boldsymbol{\psi}$ takes the form:

$$\begin{aligned}\delta\boldsymbol{\psi} &= \gamma_k D_k\boldsymbol{\phi}\left(\gamma_5\gamma_0 - i\left(a_I\alpha^I + b_J\beta^J\gamma_5\right)\right)\boldsymbol{\epsilon}\\ &= \gamma_5\bar{\gamma}_k D_k\boldsymbol{\phi}\left(\mathbb{1} - \bar{\gamma}_5\right)\boldsymbol{\epsilon}\end{aligned}$$

where $\bar{\gamma}_i \equiv \gamma_0\gamma_i, \bar{\gamma}_4 \equiv -i\gamma_0\left(a_I\alpha^I + b_J\beta^J\gamma_5\right)$, and $\bar{\gamma}_5 \equiv \bar{\gamma}_1\bar{\gamma}_2\bar{\gamma}_3\bar{\gamma}_4$. As expected, the $\bar{\gamma}_{\underline{i}}$ generate a euclidean Clifford algebra, and it follows that $\frac{1}{2}\left(\mathbb{1} - \bar{\gamma}_5\right)$ projects out one half of the states: those which have positive chirality with respect to $\bar{\gamma}_5$. Hence we conclude that $N = 4$ BPS-monopoles break one half of the supersymmetry.

From the analysis of $N = 4$ multiplets in section 2.2.3 and, in particular, from Exercise 2.6, we see that a multiplet where states break one half of the supersymmetries are short; and looking at the spectrum, we see that it is the short multiplet with spin $s = 0$, which is the only short multiplet with spins not exceeding 1. Therefore the BPS-monopole and the massive vector boson belong to isomorphic multiplets. This solves the second problem with the Montonen-Olive duality conjecture alluded to in section 1.4.1-for, certainly, if $N = 4$ supersymmetric Yang-Mills is to be self-dual, the BPS-monopole and the massive vector boson should belong to isomorphic multiplets.

Finally, we come to a minor point. The supersymmetry parameter $\boldsymbol{\epsilon}$, just like the fermion $\boldsymbol{\psi}$, is a quartet of Majorana spinors. The additional condition for the parameter to preserve the supersymmetry is that it be chiral with respect to $\bar{\gamma}_5$. One might be tempted to think that there is a problem since in four-dimensions (either with euclidean or lorentzian signature) there are no Majorana-Weyl spinors. However the Majorana condition is a condition in Minkowski spacetime, whereas the chirality condition is a condition relative to the euclidean $\bar{\gamma}_5$. We will see this more explicitly later on when we consider the effective action for the collective coordinates, but for now let us simply state without proof that these two conditions are indeed simultaneously realisable.

8.4.4 2.4.4 The mass bound for $N = 4$ super Yang-Mills

We end this chapter with a derivation of the mass bound for $N = 4$ super Yang-Mills. Keeping in mind the similar calculation for $N = 2$ super YangMills, it should come as no surprise that the mass bound coincides once again with the Bogomol'nyi bound. In order to derive the mass bound, we will first write down the algebra obeyed by the supersymmetry charges in $d = 10$ $N = 1$ super Yang-Mills. After dimensional reduction this will give us an explicit expression for the central charges appearing in the four-dimensional supersymmetry algebra. Naturally one could compute the supersymmetry algebra directly in four dimensions, but we find it simple to dimensionally reduce the algebra in ten dimensions.

The supersymmetry algebra in ten dimensions

The supersymmetry algebra can be derived by varying the supersymmetry current (2.29). Indeed, the supersymmetry algebra will be read off from the space integral of the supersymmetry variation of the timelike (zeroth) component of the supersymmetry current. Explicitly, if $\boldsymbol{\epsilon}$ is a MajoranaWeyl spinor just like Ψ, then

$$\overline{\boldsymbol{\alpha}}\{Q, \bar{Q}\}\boldsymbol{\epsilon} = -i \int_{\text{space}} \delta \bar{J}^0 \boldsymbol{\epsilon}$$

where the integral is over a spacelike hypersurface. We can get an idea of what to expect in the right-hand side of the supersymmetry algebra purely from the fact that Q is an anti-commuting Majorana-Weyl spinor. From Exercise 2.28 we see that in the right-hand side of the supersymmetry algebra, we expect only terms consisting of an odd number of Γ matrices and moreover only those bispinors Γ for which $\boldsymbol{C}\Gamma$ is symmetric, since so is the left-hand side of the supersymmetry algebra. Using equation (2.28), we see that only those terms with 1 and 5Γ matrices survive. We now turn to the computation, which is left as an exercise. We will only be interested in terms which survive in a BPS-monopole monopole background in which the fermions have been put to zero.

Exercise 2.39 (The supersymmetry algebra in ten dimensions)

Prove that up to terms involving the fermions, the variation of the supersymmetry current is given by

$$-i\delta \bar{J}^E \boldsymbol{\epsilon} = -\frac{1}{4}\overline{\boldsymbol{\alpha}}\Gamma^{AB}\Gamma^E\Gamma^{CD}\boldsymbol{\epsilon}\mathrm{G}_{AB} \cdot \mathrm{G}_{CD}$$

Perform the Γ matrix algebra and, taking into account that $\boldsymbol{\alpha}$ and $\boldsymbol{\epsilon}$ are MajoranaWeyl, show that

$$-i\delta \bar{J}^E \boldsymbol{\epsilon} = \overline{\boldsymbol{\alpha}} \left(-\frac{1}{4}\Gamma^{ABCDE}\mathrm{G}_{AB} \cdot \mathrm{G}_{CD} + 2\mathrm{G}^{EA} \cdot \mathrm{G}_{AB}\Gamma^B + \frac{1}{2}\mathrm{G}^{AB} \cdot \mathrm{G}_{AB}\Gamma^E \right) \boldsymbol{\epsilon}$$

Prove the identity

$$\Gamma^{ABCDE} = -\frac{1}{5!}\epsilon^{ABCDEFGHIJ}\Gamma_{FGHIJ}\Gamma_{11}$$

and using the fact that $\Gamma_{11}\boldsymbol{\epsilon} = -\boldsymbol{\epsilon}$, conclude that

$$-i\delta \bar{J}^E \boldsymbol{\epsilon} = 2\overline{\boldsymbol{\alpha}} \left(\frac{1}{8 \cdot 5!} \epsilon^{ABCDEFGHIJ} \mathrm{G}_{AB} \cdot \mathrm{G}_{CD} \Gamma_{FGHIJ} + \left(\mathrm{G}^{EA} \cdot \mathrm{G}_{AB} + \frac{1}{4} \mathrm{G}^{CD} \cdot \mathrm{G}_{CD} \delta^E_B \right) \Gamma^B \right) \boldsymbol{\epsilon}$$

We now define the following tensors

$$T^{AB} = \mathrm{G}^{AC} \cdot \mathrm{G}_C{}^B + \frac{1}{4} \eta^{AB} \mathrm{G}^{CD} \cdot \mathrm{G}_{CD}$$
$$\Theta^{ABCDEF} = \frac{1}{8} \epsilon^{ABCDEFGHIJ} \mathrm{G}_{GH} \cdot \mathrm{G}_{IJ}$$

We recognise T as the bosonic part of the (improved) energy-momentum tensor of the super Yang-Mills theory. The momentum is then given by the space integral of T^{0A}:

$$P^A = \int_{\text{space}} T^{0A}$$

How about Θ? Just as in the case of $N = 2$, it is a topological current.

Exercise 2.40 (Another topological current)

Prove that Θ_{ABCDEF} is gauge invariant and that it is conserved without imposing the equations of motion.

(Hint: Compare with Exercise 2.23.)

We define the topological charge associated to Θ as the space integral of Θ^{0ABCDE}:

$$Z^{ABCDE} = \int_{\text{space}} \Theta^{0ABCDE}$$

In summary, the supersymmetry algebra remains as follows:

$$\{Q, \bar{Q}\} = 2P^A \Gamma_A + \frac{2}{5!} Z^{ABCDE} \Gamma_{ABCDE} \tag{2.36}$$

In 10 dimensions, the 5 -form Z^{ABCDE} can be decomposed into a self-dual and an antiself-dual part. The next exercise asks you to show that only the self-dual part contributes to the algebra.

Exercise 2.41 (A self-dual 5-form)

Using the fact that the supersymmetry charge has negative chirality, show that we can for free project onto the self-dual part of Z^{ABCDE} in the left-hand side of the supersymmetry algebra.

This 5 -form belies the existence of a 5-brane solution of ten dimensional supersymmetric Yang-Mills. Under double dimensional reduction, it gives rise to the string-like solution of six-dimensional supersymmetric Yang-Mills briefly alluded to in section 2.3.4

The supersymmetry algebra in four dimensions

In order to write down the supersymmetry algebra in four dimensions, we need to dimensionally reduce both the momenta and the topological charge appearing in the ten-dimensional algebra (2.36). We will assume from the
start a BPS-monopole background where the fermions are put to zero. We will not demand that the solutions be static, since that is the only way we can generate electric charge. Moreover

we will exploit the internal $SO(6)$ symmetry to choose $\mathrm{P}_J = \mathrm{S}_{2,3} = 0$ and $\mathrm{S}_1 = \phi$. In such a background, the only nonzero components of the field strength G_{AB} are $\mathrm{G}_{\mu\nu}, \mathrm{G}_{\mu 4}$. This limits considerably the nonzero terms of the momentum P^A and the topological charge Z^{ABCDE}, as the next exercise shows.

Exercise 2.42 (Momentum and topological charge in this background)

Prove that in the background chosen above, the only nonzero components of the momentum and topological charge densities are the following: $T^{0\mu}, T^{04}, \Theta^{056789}$. The first term is of course simply the four-momentum density, whereas the other two terms are given by:

$$T^{04} = -\mathrm{G}_{0i} \cdot D_i \boldsymbol{\phi} = -\partial_i \left(\mathrm{G}_{0i} \cdot \boldsymbol{\phi}\right)$$
$$\Theta^{056789} = -\frac{1}{2}\epsilon_{ijk}\mathrm{G}_{ij} \cdot D_k \boldsymbol{\phi} = -\frac{1}{2}\partial_k \left(\epsilon_{ijk}\mathrm{G}_{ij} \cdot \boldsymbol{\phi}\right)$$

(Hint: In order to rewrite the right-hand sides of the equations, use the equations of motion in this background, and the Bianchi identity. (Compare with the discussion following Exercise (2.24))

Taking into account the results of the previous exercise we can rewrite the supersymmetry algebra in four dimensions as follows:

$$\{Q, \bar{Q}\} = 2\Gamma_\mu P^\mu - 2\Gamma_4 \int_{\Sigma_\infty} \mathrm{G}_{0i} \cdot \phi d\Sigma_i - \Gamma_{56789}\epsilon_{ijk} \int_{\Sigma_\infty} \mathrm{G}_{ij} \cdot \phi d\Sigma_k \tag{2.37}$$

But now notice that Γ_μ, Γ_4 and $\Gamma_{56789} = \Gamma_5\Gamma_6\Gamma_7\Gamma_8\Gamma_9$ generate a lorentzian Clifford algebra of signature (1,5), hence the supersymmetry algebra (2.37) is formally identical to one in six-dimensional Minkowski spacetime where the momenta in the extra two dimensions are given by:

$$P_5' = \int_{\Sigma_\infty} \mathrm{G}_{0i} \cdot \phi d\Sigma_i = -aq$$
$$P_6' = \frac{1}{2}\epsilon_{ijk} \int_{\Sigma_\infty} \mathrm{G}_{ij} \cdot \phi d\Sigma_k = ag$$

and where we have used equations (1.24) and (1.25) to rewrite the extra momenta in terms of the electric and magnetic charges. Finally, as in the $N = 2$ case, the mass bound is simply the positivity of the six-dimensional "mass" given by equation (2.20). Plugging in the expression for the extra momenta, we once again recover the Bogomol'nyi bound (1.27).

9 Collective Coordinates for Children by JM Figueroa-O'Farrill

BPS-monopoles are static: any motion, however small, increases their kinetic energy and makes their total energy strictly greater than the Bogomol'nyi bound. Nevertheless, if we keep the velocity small and if the motion starts off tangent to the space of static BPS-monopoles, energy conservation will prevent the motion from taking the monopoles very far away from this space. Much like a point-particle moving slowly near the bottom of a potential well, the motion of slow BPS-monopoles may be approximated by motion on the space of static BPS-monopoles (i.e., along the flat directions of the potential) and small oscillations in the transverse directions. We can trade the limit of velocities going to zero, for a limit in which the potential well becomes infinitely steep. This suppresses the oscillations in the transverse directions (which become increasingly expensive energetically) and motion is effectively constrained to take place along the flat directions, since this motion costs very little energy. Manton [Man82] showed that the motion along the flat directions is geodesic relative to a metric on the moduli space of BPS-monopoles, which is induced naturally from the Yang-Mills-Higgs action functional. Expanding the action functional around a BPS-monopole gives rise to an effective theory in terms of collective coordinates. These are the coordinates on the moduli space of BPS-monopoles and the effective action is nothing but a $(1+0)$-dimensional σ-model with target space the moduli space.

In this chapter we will study the moduli space $\mathcal{M}$ of BPS-monopoles. Our aim is to prove that it is a hyperkähler manifold which, for a given magnetic charge, is finite-dimensional, and to compute its (formal) dimension. For the simplest case of magnetic charge $k=1$, we will also work out the metric explicitly from the field theory and hence the effective action. We quantise the effective action following the review in the introduction of GM86]. This is as much as can be done directly with field-theoretical methods. The only
other case in which the metric on moduli space is known exactly is the $k=2$ monopole sector. This metric was constructed by Atiyah and Hitchin AH85 by indirect methods. We will eventually review their construction as well.

9.1 3.1 The metric on the moduli space

We start by constructing the metric on the true physical configuration space of the Yang-Mills-Higgs theory. This will induce a metric on the moduli space of BPS-monopoles, which is a submanifold.

9.1.1 3.1.1 The physical configuration space

Let $\mathcal{A}'$ denote the space of configurations $(\mathrm{W}_\mu, \boldsymbol{\phi})$ of Yang-Mills-Higgs fields of finite energy in the Prasad-Sommerfield limit. Recall that in this limit, the energy density is given by (1.13) setting $V(\phi)=0$. Configurations which are related by a short-range gauge transformation-that is, gauge transformations which tend to the identity at infinity-are to be thought of as physically indistinguishable. Hence if we let $\mathcal{G}'$ denote the group of shortrange gauge transformations, the true configuration space of the Yang-MillsHiggs system is the quotient

$$\mathcal{C} \cong \mathcal{A}'/\mathcal{G}'$$

It is convenient to fix the gauge partially by setting $\mathrm{W}_0=0$, that is by going to the temporal gauge. This still leaves the freedom of performing time-independent gauge transformations, since these are the gauge transformations which preserve the temporal gauge:

$$e\delta_\epsilon \mathrm{W}_0 = D_0 \boldsymbol{\epsilon} = \dot{\boldsymbol{\epsilon}} = 0$$

We will therefore let $\mathcal{G} \subset \mathcal{G}'$ denote the group of time-independent short-range gauge transformations.

The temporal gauge $\mathrm{W}_0=0$ is preserved by the dynamics provided we impose its equation of motion, that is, Gauss's law:

$$D_i \dot{\mathrm{W}}_i - e\boldsymbol{\phi} \times \dot{\boldsymbol{\phi}} = 0 \tag{3.1}$$

Therefore if we let $\mathcal{A}$ denote the space of finite-energy configurations $(\mathrm{W}_i, \boldsymbol{\phi})$ subject to Gauss's law, then the configuration space $\mathcal{C}$ also admits the description

$$\mathcal{C} \cong \mathcal{A}/\mathcal{G}$$

We can simplify many of the calculations by describing the space $\mathcal{A}$ in a different way. We introduce a fourth spatial coordinate x^4 and interpret the Higgs field $\boldsymbol{\phi}$ as the fourth-component W_4 of the gauge field. Notice however that we must impose that nothing depends on the new coordinate: $\partial_4 \equiv 0$. We will write $\mathrm{W}_{\underline{i}}=(\mathrm{W}_i,\ \mathrm{W}_4=\phi)$, where the underlined indices run from 1 to 4. Notice that the field-strength has components $\mathrm{G}_{\underline{ij}}=(\mathrm{G}_{ij}, \mathrm{G}_{i4}=D_i\phi)$. In this new notation, gauge transformations, Gauss's law and the Bogomol'nyi equation all have natural and simple descriptions.

Exercise 3.1 (The BPS-monopole as an instanton)

Prove that infinitesimal gauge transformations on the Yang-Mills-Higgs system now take the form

$$\delta_\epsilon \mathrm{W}_{\underline{i}}=\frac{1}{e} D_{\underline{i}} \epsilon$$

that Gauss's law (3.1) becomes simply

$$D_{\underline{i}} \dot{\mathrm{W}}_{\underline{i}}=0 \tag{3.2}$$

and that the Bogomol'nyi equation (1.28) is nothing but the (anti)self-duality equation

$$\mathrm{G}_{\underline{ij}}= \pm \frac{1}{2} \epsilon_{ijk\ell} \mathrm{G}_{\underline{k\ell}} \tag{3.3}$$

In summary, this proves that BPS-monopoles in $3+1$ dimensions are in one-to-one correspondence with static instantons in $4+1$ dimensions which are translationally invariant in the fourth spatial direction.

Therefore in this description, the space $\mathcal{A}$ is given as those gauge fields W_i in $4+1$ dimensions, independent of x^4, of finite energy per unit length in the x^4-direction, and whose time-dependence is subject to (3.2).

9.1.2 3.1.2 The metric on the physical configuration space

In the temporal gauge, the Yang-Mills-Higgs lagrangian in the PrasadSommerfield limit is given as a difference of two terms: $L=T-V$, where the kinetic term T is the 3 -space integral of

$$\frac{1}{2}\left\|\dot{\mathrm{W}}_{\underline{i}}\right\|^2=\frac{1}{2}\left\|\dot{\mathrm{W}}_i\right\|^2+\frac{1}{2}\|\dot{\boldsymbol{\phi}}\|^2 \tag{3.4}$$

and the potential term V is the 3 -space integral of

$$\frac{1}{2}\|\mathrm{~B}_i\|^2+\frac{1}{2}\|D_i \phi\|^2=\frac{1}{2}\|\mathrm{~B}_i \mp D_i \phi\|^2 \pm \partial_i\left(\phi \cdot \mathrm{B}_i\right) \tag{3.5}$$

We will now show that the lagrangian is well-defined in the true configuration space $\mathcal{C}$. The kinetic term will induce a metric.

Suppose we would like to compute the value of the potential on some point in $\mathcal{C}$. Points in $\mathcal{C}$ are equivalence classes $[\mathrm{W}_{\underline{i}}]$ of points $\mathrm{W}_{\underline{i}}$ in $\mathcal{A}$: two points in $\mathcal{A}$ belong to the same equivalence class if and only if they are related by a gauge transformation in $\mathcal{G}$; that is, if they lie on the same $\mathcal{G}$-orbit. To define a potential on $\mathcal{C}$ we can simply use the potential term (3.5) on $\mathcal{A}$ as follows: to find out the value of the potential on a point $[\mathrm{W}_i]$ in $\mathcal{C}$, we choose some point W_i in the same equivalence class, and we evaluate the potential (3.5) on it. This will only make sense if the value of the potential doesn't depend on which element in the equivalence class we have chosen; that is, if the potential is gauge-invariant. More formally, a function on $\mathcal{A}$ will induce a function on $\mathcal{C}=\mathcal{A}/\mathcal{G}$ if and only if it is $\mathcal{G}$-invariant. Luckily this is the case, since in the temporal gauge, the potential and kinetic terms are separately invariant under time-independent gauge-transformations.

The kinetic term is trickier since it is not strictly speaking a function on $\mathcal{A}$: it requires not just knowledge of $W_{\underline{i}}$ but also of its time-derivative $\dot{W}_{\underline{i}}$. In other words, it is a function on the tangent bundle $T\mathcal{A}$. The typical fibre at a point $[W_i]$ of the tangent bundle is spanned by the velocities of all smooth curves passing through that point. We may lift such curves to curves in $\mathcal{A}$, but this procedure is not unique. First we have to choose a point $W_{\underline{i}}$ in $\mathcal{A}$ in the equivalence class $[W_i]$. Just as before, this ambiguity is immaterial since the kinetic term is invariant under time-independent gauge transformations. But now we also have to choose a tangent vector $\dot{\mathbf{W}}_{\underline{i}}$. Clearly adding to a tangent vector a vector tangent to the orbits of $\mathcal{G}$ does not change the curve in C since every $\mathcal{G}$-orbit in $\mathcal{A}$ is identified with a single point in C. Hence the kinetic term should be impervious to such a change. Tangent vectors to $\mathcal{G}$ are infinitesimal gauge-transformations whose parameters go to zero at spatial infinity, therefore the kinetic term (3.4) defines a kinetic energy on C provided that $\dot{W}_{\underline{i}}$ and $\dot{W}_{\underline{i}} + D_i\boldsymbol{\epsilon}$ have the same kinetic energy. Integrating by parts we see that this is a consequence of Gauss's law (3.2).

In summary, the Yang-Mills-Higgs lagrangian induces a lagrangian in the true configuration space C, whose energy is of course given by $E = T + V$. The kinetic energy term T defines a metric on C. Motion on C is not "free" of course, since there is also a potential term, but for motion along the flat directions at the bottom of the potential well, this will be a good approximation. We turn to this now.

9.1.3 3.1.3 The metric on the moduli space

Let $\mathcal{M}$ denote the subspace of C where the energy E attains its minimum. From the explicit expression for T and V given in (3.4) and (3.5), we see that the minimum of the energy is given by

$$\left| \int_{\mathbb{R}^3} d^3x \partial_i \left(\boldsymbol{\phi} \cdot \mathrm{B}_i \right) \right| = a|g| = \frac{4\pi a}{e}|k| \tag{3.6}$$

where g is the magnetic charge, and the integer k is the topological or monopole number. Clearly the minimum is attained by those configurations corresponding to static solutions on the Bogomol'nyi equation: BPSmonopoles, and where any two such solutions which are gauge-related are identified. In other words, $\mathcal{M}$ is the moduli space of static BPS-monopoles.

The monopole number labels different connected components of the space $\mathcal{A}$, so that

$$\mathcal{A} = \bigcup_k \mathcal{A}_k$$

but the gauge group $\mathcal{G}$ preserves each component. Therefore we can also decompose the true configuration space C as

$$C = \bigcup_k C_k \quad \text{where} \quad C_k = \mathcal{A}_k / \mathcal{G}$$

Finally, let $\mathcal{M}_k = \mathcal{M} \cap C_k$. This is then the moduli space of static BPSmonopoles of monopole number k, or BPS- k-monopoles, for short.

By definition, the potential is constant on $\mathcal{M}_k$, so that the Yang-MillsHiggs lagrangian is given by

$$L = T - \frac{4\pi a}{e}|k| \tag{3.7}$$

Therefore $\mathcal{M}_k$ corresponds to the manifold of flat directions of the potential. Manton's argument given at the beginning of this chapter, can now be proven. The motion of slow monopoles which start off tangent to $\mathcal{N}_k$ will consists of the superposition of two kinds of motions: motions along the flat directions $\mathcal{M}_k$ and small oscillations in the directions normal to $\mathcal{M}_k$. In the limit of zero velocity, the oscillatory motion is suppressed and we are left with motion on $\mathcal{M}_k$. But this motion is governed by the lagrangian (3.7) which only has a kinetic term - whence the motion is free, or in other words geodesic relative to the metric on $\mathcal{M}_k$ defined by T.

9.1.4 3.1.4 The 1-monopole moduli space

Let us study the moduli space $\mathcal{M}_1$ in the 1-monopole sector. The coordinates for $\mathcal{M}_1$ can be understood as parameters on which the BPS-monopole solution depends. In the 't Hooft-Polyakov Ansatz (1.15), the monopole is centred at the origin in $\mathbb{R}^3$, but the invariance under translations of the Yang-MillsHiggs lagrangian (1.8) means that we can put the centre of the monopole where we please. This introduces three moduli parameters: $\boldsymbol{X}$. The time evolution of these parameters corresponds to the BPS-monopole moving as if it were a particle with mass $4\pi a/e$. The effective lagrangian for these collective coordinates is then

$$L_{\text{eff}} = \frac{2\pi a}{e}\dot{\boldsymbol{X}}^2$$

There is a fourth, more subtle, collective coordinate. Consider a oneparameter family $\mathrm{W}_{\underline{i}}(t)$ of gauge fields, but where the t-dependence is pure gauge:

$$\dot{\mathbf{W}}_{\underline{i}} = \frac{1}{e}D_{\underline{i}}\boldsymbol{\epsilon}(t). \tag{3.8}$$

Since the potential is gauge invariant, this corresponds to a flat direction. But one might think that it is not a physical flat direction since it is tangent to the $\mathcal{G}$-orbits - it is an infinitesimal gauge transformation, after all. But recall that $\mathcal{G}$ is the group of short-range gauge transformations, whence for $D_i\boldsymbol{\epsilon}$ to be tangent to the orbits, $\boldsymbol{\epsilon}$ has to tend to 0 as we approach infinity. Indeed, $D_i\epsilon$ would represent a physical deformation of the BPS-monopole if it would obey Gauss's law (3.2), which implies

$$D^2\boldsymbol{\epsilon} = 0 \tag{3.9}$$

Exercise 3.2 (D^2 has no normalisable zero modes)

Prove that acting on square-integrable functions $D^2 = -D_i^\dagger D_{\underline{i}}$ is a negative-definite operator. Deduce from this that any normalisable zero mode must be a zero mode of D_i for each $\underline{i}$, and deduce from this that the only square-integrable solution to (3.9) is the trivial solution $\boldsymbol{\epsilon} = 0$. In other words, there exist no normalisable solutions.

(Hint: If $D_{\underline{i}}\epsilon = 0$, then $\|\epsilon\|^2$ is constant.)

This discussion suggests that we look for a gauge parameter $\boldsymbol{\epsilon}$ which does not tend to zero asymptotically. For example, let $\boldsymbol{\epsilon}(t) = f(t)\boldsymbol{\phi}$, where $f(t)$ is an arbitrary function. In the 1-monopole sector, ϕ defines a map of degree 1 at infinity, hence it certainly does not go to zero. Moreover, using the Bogomol'nyi equation, it follows at once that $f(t)\phi$ is a (un-normalisable) zero mode of D^2. It is clearly a true moduli parameter because it costs energy to excite it:

$$T = \frac{1}{2e^2}f^2\int_{\mathbb{R}^3}\|D_i\boldsymbol{\phi}\|^2 > 0 \tag{3.10}$$

We can understand this as follows. Let $g = \exp(\chi\boldsymbol{\phi}/a)$ be a timedependent gauge transformation, where all the time-dependence resides in χ. Such a gauge transformation will move us away from the temporal gauge, but assume that at time $t = 0$, say, we start from a configuration $\mathrm{W}_{\underline{i}}$ in the temporal gauge, and suppose that $g(t = 0) = 1$. Then from (1.9),

$$\mathbf{W}_{\underline{i}}(t) = g\mathbf{W}_{\underline{i}}g^{-1} + \frac{1}{e}\partial_{\underline{i}}gg^{-1}$$

whence

$$\dot{\mathrm{W}}_{\underline{i}}(t) = \frac{1}{ae}\dot{\chi}D_{\underline{\underline{x}}}\phi$$

Comparing with (3.8), we see that $f = \dot{\chi}/a$. Using (3.10), the kinetic energy of such a configuration is given by

$$T = \frac{1}{2a^2e^2}\dot{\chi}^2 \int_{\mathbb{R}^3} \|D_i\phi\|^2$$

But notice that since $D_i\phi = \mathrm{B}_i$,

$$\begin{aligned} T &= \frac{1}{2a^2e^2}\dot{\chi}^2 \int_{\mathbb{R}^3} \left(\frac{1}{2}\|D_i\phi\|^2 + \frac{1}{2}\|\,\mathrm{B}_i\|^2\right) \\ &= \frac{1}{2a^2e^2}\dot{\chi}^2 \left(\frac{4\pi a}{e}\right) \\ &= \frac{2\pi}{ae^3}\dot{\chi}^2 \end{aligned}$$

Notice that χ is an angular variable. To see this, let us define $g(\chi) = \exp(\chi\boldsymbol{\phi}/a)$. Then it is easy to see that $g(\chi)$ and $g(\chi + 2\pi)$ are gauge-related in $\mathcal{G}$, that is, via short-range gauge transformations. Indeed, recall that $\|\phi\| \to a$ at infinity in the Prasad-Sommerfield limit, whence $g(2\pi) = \exp(2\pi\boldsymbol{\phi}/a)$ tends to 1 at infinity. Since $g(\chi + 2\pi) = g(\chi)g(2\pi)$, we are done.

Assuming for the moment (we will prove this later) that there are no other collective coordinates in the 1-monopole sector, we have proven that the moduli space of BPS-1-monopoles is given by

$$\mathcal{M}_1 \cong \mathbb{R}^3 \times S^1$$

and the metric can be read off from the expression for the effective action

$$L_{\text{eff}} = \frac{1}{2}g_{ab}\dot{X}^a\dot{X}^b - \frac{4\pi a}{e} \tag{3.11}$$

$$= \frac{2\pi a}{e}\dot{\boldsymbol{X}}^2 + \frac{2\pi}{ae^3}\dot{\chi}^2 - \frac{4\pi a}{e} \tag{3.12}$$

that is

$$g_{ab} = \frac{4\pi a}{e}\begin{pmatrix} \mathbb{1}_3 & 0 \\ 0 & e^{-2} \end{pmatrix}$$

from where we can see that the radius of the circle is inversely proportional to the electric charge.

9.1.5 3.1.5 The quantisation of the effective action

The effective action (3.12) corresponds to a particle moving freely in $\mathbb{R}^3 \times S^1$ with the flat metric. The quantisation of this effective action is straightforward. The canonical momenta $(\boldsymbol{P}, Q)$ given by

$$\boldsymbol{P} = \frac{4\pi a}{e}\dot{\boldsymbol{X}} \quad \text{and} \quad Q = \frac{4\pi}{ae^3}\dot{\chi}$$

are conserved, and the hamiltonian is given by

$$H = \frac{e}{8\pi a}\boldsymbol{P}^2 + \frac{ae^3}{8\pi}Q^2 + \frac{4\pi a}{e} \tag{3.13}$$

Bound states of minimum energy are given by those eigenstates of the hamiltonian for which $\boldsymbol{P} = 0$. Since χ is angular with period 2π, the eigenvalues of Q are quantised in units of λ, whence the spectrum looks like

$$E_n = \frac{4\pi a}{e} + \frac{ae^3}{8\pi}(n\lambda)^2 \tag{3.14}$$

Notice that this energy spectrum has the standard form which corresponds to perturbative states around a nonperturbative vacuum. If we think of e as the coupling constant, then the zero-point energy is not analytic in e, hence it corresponds to a non-perturbative state in the theory: the BPS-monopole in this case. The second term in the energy corresponds to excitations around the monopole, which are clearly perturbative since their energy goes to zero as we let the coupling tend to zero.

Exercise 3.3 (The electric charge)

Prove that the electric field for classical configurations in which $\boldsymbol{P} = 0$ is given by

$$\mathrm{E}_i = -\mathrm{G}_{0i} = -\dot{\mathrm{W}}_i = -e^2 Q\ \mathrm{B}_i/4\pi$$

Conclude that e Q can be interpreted as the electric charge.

Taking the above exercise into consideration, we see that the spectrum of the quantum effective theory corresponds to dyons of magnetic charge $-4\pi/e$ and electric charge $ne\lambda$, for $n \in \mathbb{Z}$. According to the classical BPS formula (1.27), the rest mass of such a dyon would be equal to

$$M_n = a\sqrt{\left(\frac{4\pi}{e}\right)^2 + (ne\lambda)^2} = \frac{4\pi a}{e}\sqrt{1 + \left(\frac{ne^2\lambda}{4\pi}\right)^2}$$

which, if we expand the square root assuming that e is small, becomes

$$= E_n + O\left(e^5\right)$$

where we've used (3.14).

In summary, the energy spectrum obtained from quantising the effective action of the collective coordinates is a small-coupling approximation to the expected BPS energy spectrum. However, even if their energy is only approximately correct, the multiplicity of bound states can be read accurately from the effective action. This is one of the important lessons to be drawn from the collective coordinate expansion.

In principle one can repeat this analysis in the k-monopole sector provided that one knows the form of the metric. But at the present moment this is only the case for $k = 1$ and $k = 2$. We will discuss the effective theory for $k = 2$ later on in the lectures in the context of $N = 4$ supersymmetric Yang-Mills.

9.1.6 3.1.6 Some general properties of the monopole moduli space

Quite a lot is known about the properties of the k-monopole [1] moduli space $\mathcal{N}_k$, even though its metric (and hence the effective action) is known explicitly only for $k = 1, 2$. As we saw in the previous section, the metric on $\mathcal{M}_1$ can be computed directly from the field theory. On the other hand, the metric on $\mathcal{N}_2$ can only be determined via indirect means. This result as well as much else of what is known about $\mathcal{M}_k$ is to be found either explicitly or referenced in the book AH88] by Atiyah and Hitchin (see also [AH85]) to where we refer the reader for details.

We will now state some facts about $\mathcal{N}_k$. Some of them we will be able to prove later with field-theoretical means, but proving some others would take us too far afield. The following properties of $\mathcal{M}_k$ are known AH88: 1. $\mathcal{M}_k$ is a $4k$-dimensional (non-compact) complete riemannian manifold;

The natural metric on $\mathcal{M}_k$ is hyperkähler;

$\mathcal{M}_k \cong \tilde{\mathcal{M}}_k/\mathbb{Z}_k$ where $\tilde{\mathcal{M}}_k \cong (\mathbb{R}^3 \times S^1) \times \tilde{\mathcal{M}}_k^0$ as hyperkähler spaces.

$\tilde{\mathcal{M}}_k^0$ is a $4(k-1)$-dimensional, irreducible, simply-connected, hyperkähler manifold admitting an action of $SO(3)$ by isometries which rotates the three complex structures;

Asymptotically $\tilde{\mathcal{N}}_k \rightarrow \underbrace{\mathcal{N}_1 \times \mathcal{M}_1 \times \cdots \times \mathcal{M}_1}_{k \text{ times}} \equiv \mathcal{M}_1^k$, and $\mathcal{M}_k \rightarrow \mathcal{M}_1^k/\mathbb{Z}_k$. Physically this means that a configuration of well-separated BPS- k monopoles can be considered as k1-monopole configurations. The fact that BPS-monopoles are classically indistinguishable is responsible for the $\mathbb{Z}_k$-quotient.

9.2 3.2 dim$\mathcal{M}_k = 4k$

In this section we compute the dimension of the moduli space of static BPSmonopoles. The strategy is typical of this kind of problems. We fix a reference BPS-monopole and ask in how many directions can we deform the solution infinitesimally and still remain with a BPS-monopole. Most of these directions will be unphysical: corresponding to infinitesimal gauge transformations. Discarding them leaves us with a finite number of physical directions along which to deform the BPS-monopole. In other words, we are computing the dimension of the tangent space at a particular point in the moduli space. If the point is regular (and generic points usually are) then this is the dimension of the moduli space itself. This number is in any case called the formal dimension of the moduli space.

Since the number of all deformations and of infinitesimal gauge transformations are both infinite, it is better to fix the gauge before counting: this eliminates the gauge-redundant deformations and leaves us with only a finite formal dimension. With a little extra argument, the counting can then be done via an index theorem. In the case of BPS-monopoles, the relevant index theorem is that of Callias Cal78 (slightly modified by Weinberg [Wei79]) which is valid for open spaces and for operators with suitable decay properties at infinity. Weinberg's calculation contains steps which from a strictly mathematical point of view may be deemed unjustified. The necessary analytic details have been sorted out by Taubes [Tau83, but we will be following Weinberg's heuristic calculation in any case.

9.2.1 3.2.1 The dimension as an index

First let set up the problem. We want to find out in how many physically different ways can one deform a given BPS-monopole. It will turn out that these are given by zero modes of a differential operator. Asking for the number of zero modes will be the same as asking for the dimension of the tangent space at a given BPS-monopole solution. Hence let $t \rightsquigarrow (\mathrm{W}_i(t), \phi(t))$

*[1] We will only concern ourselves with positive k : $\mathcal{M}_{-k}$ is naturally isomorphic to $\mathcal{M}_k$ by performing a parity transformation on the solutions.

be a family of static BPS-monopoles (here t is an abstract parameter which has nothing to do with time). This means that $\left(\dot{\mathrm{W}}_i, \dot{\boldsymbol{\phi}}\right) \equiv \left(\dot{\mathrm{W}}_i(0), \dot{\boldsymbol{\phi}}(0)\right)$ is a tangent vector to the moduli space at the point $(\mathrm{W}_i, \boldsymbol{\phi}) \equiv (\mathrm{W}_i(0), \phi(0))$. Taking the t-derivative of the Bogomol'nyi equation (1.28), we find that $\left(\dot{\mathrm{W}}_i, \dot{\phi}\right)$ satisfies the linearised Bogomol'nyi equation:

$$D_i\dot{\boldsymbol{\phi}} + e\boldsymbol{\phi} \times \dot{\mathrm{W}}_i = \epsilon_{ijk} D_j \dot{\mathrm{W}}_k \tag{3.15}$$

However, not every solution of the linearised Bogomol'nyi equation need be a physical deformation: it could be an infinitesimal gauge transformation. To make sure that it isn't, it is necessary to impose in addition Gauss's law (3.1). In other words, the dimension of the tangent space of the moduli space of BPS-monopoles is given by the maximum number of linearly independent solutions of both (3.15) and (3.1).

In order to count these solutions it will be convenient to rewrite both of these equations in terms of a single matrix-valued equation. We will define the following 2×2 complex matrix:

$$\Psi = \dot{\phi}\mathbb{1} + i\dot{\mathrm{W}}_j\sigma_j \tag{3.16}$$

and the following linear operator:

$$\mathcal{D} = e\boldsymbol{\phi}\mathbb{1} + iD_j\sigma_j \tag{3.17}$$

where we follow the convention that all fields which appear in operators are in the adjoint representation; that is, ϕ really stands for ad $\phi = \phi \times -$, etc.

Exercise 3.4 (Two equations in one)

Prove that the linearised Bogomol'nyi equation (3.15) and Gauss's law (3.1) together are equivalent to the equation $\mathcal{D}\Psi = 0$.

We want to count the number of linearly independent real normalisable solutions to $\mathcal{D}\Psi = 0$. It is easier to compute the index of the operator $\mathcal{D}$. By definition, the index of $\mathcal{D}$ is difference between the number of its normalisable zero modes and the number of normalisable zero modes of its hermitian adjoint $\mathcal{D}^\dagger$ relative to the inner product:

$$\int d^3x \,\mathrm{tr}\, \Psi^*\Psi = \int d^3x \left(\dot{\boldsymbol{\phi}}^* \cdot \dot{\boldsymbol{\phi}} + \dot{\mathbf{W}}_i^* \cdot \dot{\mathbf{W}}_i\right)$$

where $*$ denotes complex conjugation of the fields and hermitian conjugation on the 2×2 matrices, and where tr denotes the 2×2 matrix trace.

The expression for the index of $\mathcal{D}$

$$\mathrm{ind}\ \mathcal{D} = \dim\ker\mathcal{D} - \dim\ker\mathcal{D}^\dagger$$

can be turned into an inequality

$$\dim\ker\mathcal{D} \geq \ \mathrm{ind}\ \mathcal{D}$$

which saturates precisely when $\mathcal{D}^\dagger$ has no normalisable zero modes. Happily this is the case, as the next exercise asks you to show.

Exercise 3.5 ($\mathcal{D}^\dagger$ has no normalisable zero modes)

Prove that $\mathcal{D}\mathcal{D}^\dagger$ is a positive-definite operator, whence it has no normalisable zero modes. Compare with Exercise 3.2. Also prove that for antimonopoles (i.e., the other sign in the Bogomol'nyi equation, it is $\mathcal{D}^\dagger\mathcal{D}$ that is the positive operator.

(Hint: Use the fact that both ϕ and D_j are antihermitian operators to prove that $\mathcal{D}^\dagger = -e\phi\mathbb{1} + i\sigma_j D_j$, and that $\mathcal{D}\mathcal{D}^\dagger = -e^2\phi^2 - (D_j)^2$. Deduce that this operator is positive-definite.

Therefore the number of normalisable zero modes of $\mathcal{D}$ equals the index of the operator $\mathcal{D}$. In the following sections we will compute the index of $\mathcal{D}$ acting on two-component complex vectors; that is, on functions $\mathbb{R}^3 \to \mathbb{C}^2$. However the (formal) dimension of monopole moduli space is given by the number of normalisable zero modes of $\mathcal{D}$ acting on matrices of the form (3.16). To a deformation $\left(\dot{\phi}, \dot{\mathrm{W}}_i\right)$ there corresponds a matrix

$$\Psi = \begin{pmatrix} \dot{\phi} + i\dot{\mathrm{W}}_3 & \dot{\mathrm{W}}_2 + i\dot{\mathrm{W}}_1 \\ -\dot{\mathrm{W}}_2 + i\dot{\mathrm{W}}_1 & \dot{\phi} - i\dot{\mathrm{W}}_3 \end{pmatrix}$$

Clearly the first column of the above matrix Ψ determines the matrix. Moreover this first column is a normalisable zero mode of $\mathcal{D}$ acting on vectors if and only if Ψ is a normalisable zero mode of $\mathcal{D}$ acting on matrices. This would seems to indicate that there is a one-to-one correspondence between the normalisable zero modes of $\mathcal{D}$ acting on vectors and of $\mathcal{D}$ acting on matrices, but notice that $\mathcal{D}$ is a complex linear operator in a complex vector space, hence the space of its normalisable zero modes is complex, of complex dimension ind $\mathcal{D}$. However the matrices $\boldsymbol{\Psi}$ and $i\boldsymbol{\Psi}$ determine linearly independent tangent vectors to monopole moduli space, hence it is its real dimension which equals the (formal) dimension of monopole moduli space $\boldsymbol{\mathcal{M}}_k$. In other words,

$$\dim \boldsymbol{\mathcal{M}}_k = 2 \text{ ind } \mathcal{D}$$

9.2.2 3.2.2 Computing the index of $\mathcal{D}$

Our purpose is then to compute ind $\mathcal{D}$. To this effect consider the following expression:

$$I\left(M^2\right) = \operatorname{Tr}\left(\frac{M^2}{\mathcal{D}^\dagger\mathcal{D} + M^2}\right) - \operatorname{Tr}\left(\frac{M^2}{\mathcal{D}\mathcal{D}^\dagger + M^2}\right) \tag{3.18}$$

where Tr is the operatorial trace.

Exercise 3.6 (A formula for the index of $\mathcal{D}$)

Prove that the index of $\mathcal{D}$ is given by

$$\begin{aligned} \text{ind } \mathcal{D} &= \dim\ker\mathcal{D}^\dagger\mathcal{D} - \dim\ker\mathcal{D}\mathcal{D}^\dagger \\ &= \lim_{M^2\to 0} I\left(M^2\right) \end{aligned}$$

(Hint: Prove this assuming that there is a gap in the spectrum of these operators. This is not the case, but as argued in Wei79] the conclusion is unaltered.)

In order to manipulate equation (3.18) it is again convenient to use the reformulation of the BPS-monopole as an instanton, in terms of $\mathrm{W}_{\underline{i}} = (\mathrm{W}_i, \phi)$, and to define the following four-dimensional euclidean Dirac matrices:

$$\bar{\gamma}_k = \begin{pmatrix} 0 & -i\sigma_k \\ i\sigma_k & 0 \end{pmatrix} \quad \bar{\gamma}_4 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \quad \bar{\gamma}_5 = \bar{\gamma}_1\bar{\gamma}_2\bar{\gamma}_3\bar{\gamma}_4 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \tag{3.19}$$

obeying

$$\left\{\bar{\gamma}_{\underline{i}}, \bar{\gamma}_{\underline{j}}\right\} = 2\delta_{\underline{i}\underline{j}}$$

Letting $D_{\underline{i}}$ denote the gauge covariant derivative corresponding to W, and remembering that $\partial_4 \equiv 0$, we find that

$$\bar{\gamma} \cdot D \equiv \bar{\gamma}_{\underline{i}} D_{\underline{i}} = \begin{pmatrix} 0 & -\mathcal{D} \\ \mathcal{D}^{\dagger} & 0 \end{pmatrix}$$

whence

$$-(\bar{\gamma} \cdot D)^2 = \begin{pmatrix} \mathcal{D}\mathcal{D}^{\dagger} & 0 \\ 0 & \mathcal{D}^{\dagger}\mathcal{D} \end{pmatrix}$$

Exercise 3.7 (Another formula for $I\left(M^2\right)$)
Prove that

$$I\left(M^2\right) = -\operatorname{Tr} \bar{\gamma}_5 \frac{M^2}{-(\bar{\gamma} \cdot D)^2 + M^2}$$

where Tr now also includes the spinor trace. More generally, if f is any function for which the traces $\operatorname{Tr} f\left(\mathcal{D}^{\dagger}\mathcal{D}\right)$ and $\operatorname{Tr} f\left(\mathcal{D}\mathcal{D}^{\dagger}\right)$ exist, prove that

$$\operatorname{Tr} \bar{\gamma}_5 f\left(-(\bar{\gamma} \cdot D)^2\right) = \operatorname{Tr} f\left(\mathcal{D}\mathcal{D}^{\dagger}\right) - \operatorname{Tr} f\left(\mathcal{D}^{\dagger}\mathcal{D}\right)$$

Let K be any operator acting on square-integrable (matrix-valued) functions $\psi(x).K$ is defined uniquely by its kernel $K(x, y)$:

$$(K\psi)(x) = \int d^3 y K(x, y)\psi(y)$$

If we rewrite this equation using Dirac's "ket" notation, so that $\psi(x) = \langle x \mid \psi\rangle$, then we see that the above equation becomes

$$\langle x|K|\psi\rangle = \int d^3 y \langle x|K|y\rangle\langle y \mid \psi\rangle$$

whence we can think of the kernel $K(x, y)$ as $\langle x|K|y\rangle$. We will often use this abbreviation for the kernel of an operator. In particular, its trace is given by

$$\operatorname{Tr} K = \int d^3 x \operatorname{tr}\langle x|K|x\rangle$$

where tr stands for the matrix trace, if any.
The rest of this section will concern the calculation of the following expression

$$I\left(M^2\right) = -\int d^3 x \operatorname{tr} \bar{\gamma}_5 \left\langle x \left| \frac{M^2}{-(\bar{\gamma} \cdot D)^2 + M^2} \right| x \right\rangle \tag{3.20}$$

where tr now stands for both the spinor and matrix traces. Let's focus on the kernel

$$I(x, y) = -\operatorname{tr} \bar{\gamma}_5 \left\langle x \left| \frac{M^2}{-(\bar{\gamma} \cdot D)^2 + M^2} \right| y \right\rangle$$

Exercise 3.8 (Some properties of kernels)

Let A and B be operators acting on (matrix-valued) square-integrable functions. Let A be a differential operator. Then prove the following identities:

$$A(x) \cdot B(x, y) = (AB)(x, y)$$
$$B(x, y) \cdot \overleftarrow{A}^{\dagger}(y) = (BA)(x, y)$$

where - means action of differential operators, the label on a differential operator denotes on which variable it acts, and the arrow on $A^{\dagger}$ in the second equation means that the derivatives act on B.

Using the results of this exercise and the fact that the trace of an even number of $\bar{\gamma}$-matrices vanishes, we can rewrite $I(x, y)$ slightly. Writing $-(\bar{\gamma}.\ D)^2 + M^2 = (\bar{\gamma} \cdot D + M)(-\bar{\gamma} \cdot D + M)$, we have that

$$I(x, y) = -\operatorname{tr} \bar{\gamma}_5 \left\langle x \left| \frac{M}{\bar{\gamma} \cdot D + M} \right| y \right\rangle = -M \operatorname{tr} \bar{\gamma}_5 \Delta(x, y) \tag{3.21}$$

where we have introduced the propagator

$$\Delta(x, y) = \left\langle x \left| \frac{1}{\bar{\gamma} \cdot D + M} \right| y \right\rangle$$

Using once again the results of Exercise 3.8, one immediately deduces the following identities:

$$\left(\bar{\gamma}_i \frac{\partial}{\partial x^i} - e\bar{\gamma}_{\underline{i}} \mathbf{W}_{\underline{i}}(x) + M\right) \Delta(x, y) = \delta(x - y)$$

$$\Delta(x, y) \left(-\bar{\gamma}_i \frac{\overleftarrow{\partial}}{\partial y^i} - e\bar{\gamma}_{\underline{\underline{i}}}\ \mathrm{W}_{\underline{\underline{i}}}(y) + M\right) = \delta(x - y)$$

and from them:

$$\begin{aligned} I(x, y) = &\frac{1}{2}\left(\frac{\partial}{\partial x^i} + \frac{\partial}{\partial y^i}\right) \operatorname{tr} \bar{\gamma}_5 \bar{\gamma}_i \Delta(x, y) \\ &- \frac{e}{2} \operatorname{tr} \bar{\gamma}_5 \bar{\gamma}_{\underline{i}} \left(\ \mathrm{W}_{\underline{i}}(x) - \mathrm{W}_{\underline{i}}(y)\right) \Delta(x, y) \end{aligned} \tag{3.22}$$

which can be understood as a "conservation law" for the bi-local current

$$J_i(x, y) \equiv \operatorname{tr} \bar{\gamma}_5 \bar{\gamma}_i \Delta(x, y)$$

In order to compute $I\left(M^2\right)$ we have to first take the limit $y \rightarrow x$ of $I(x, y)$. From equation (3.22) we find that

$$I(x, x) = \frac{1}{2} \frac{\partial}{\partial x^i} J_i(x, x) - \lim_{y \rightarrow x} \frac{e}{2} \operatorname{tr} \bar{\gamma}_5 \bar{\gamma}_{\underline{\underline{i}}} \left(\mathbf{W}_{\underline{i}}(x) - \mathrm{W}_{\underline{i}}(y)\right) \Delta(x, y) \tag{3.23}$$

Although the last term has a $\mathbf{W}_{\underline{i}}(x) - \mathrm{W}_{\underline{i}}(y)$ which vanishes as $y \rightarrow x$, the propagator is singular in this limit and we have to pay careful attention to the nature of these singularities in order to conclude that this term does not contribute. Clearly we can admit at most a logarithmic singularity. The purpose of the following (long) exercise is to show that nothing more singular than that occurs.

Exercise 3.9 (Regularity properties of the propagator)

Prove that the following limit has at most a logarithmic singularity:

$$\lim_{y \rightarrow x} \operatorname{tr} \bar{\gamma}_5 \bar{\gamma}_{\underline{i}} \Delta(x, y)$$

where tr now only denotes the spinor trace.

(Hint: Notice that $\Delta(x, y)$ is the propagator of a three-dimensional spinor in the presence of a background gauge field. First let us approximate $\Delta(x, y)$ perturbatively in the coupling constant e:

$$\Delta(x, y)=\sum_{n=0}^{\infty} e^{n} \Delta_{n}(x, y)$$

Imposing the equation

$$(\bar{\gamma} \cdot D+M)(x) \Delta(x, y)=\delta(x-y)$$

order by order in e, we find that

$$\begin{aligned} \Delta(x, y)= & \sum_{n=0}^{\infty} e^{n} \int \prod_{i=1}^{n} d^{3} z_{i} \Delta_{0}\left(x, z_{1}\right) \\ & \times\left[\prod_{i=1}^{n-1}(\bar{\gamma} \cdot \mathrm{W})\left(z_{i}\right) \Delta_{0}\left(z_{i}, z_{i+1}\right)\right](\bar{\gamma} \cdot \mathbf{W})\left(z_{n}\right) \Delta_{0}\left(z_{n}, y\right) \end{aligned}$$

where $\Delta_{0}(x, y)$ is the free propagator:

$$\Delta_{0}(x, y)=\left\langle x\left|\frac{1}{(\bar{\gamma} \cdot \partial)+M}\right| y\right\rangle=\int \frac{d^{3} p}{(2 \pi)^{3}} \frac{e^{i p \cdot(x-y)}}{i(\bar{\gamma} \cdot p)+M}$$

Since we are interested in the behaviour of $\Delta(x, y)$ as $|x-y| \rightarrow 0$, we need to make some estimates. Prove that $\Delta_{n}(x-y) \sim|x-y|^{-2+n}$ in this limit, whence we the potentially singular contributions come from $n=0$ and $n=1$. Prove that $\operatorname{tr} \bar{\gamma}_{5} \bar{\gamma}_{\underline{i}} \Delta_{0}(x, y)=0$ using the facts that $\operatorname{tr} \bar{\gamma}_{5} \bar{\gamma}_{\underline{i}}=\operatorname{tr} \bar{\gamma}_{5} \bar{\gamma}_{\underline{i}} \bar{\gamma}_{j}=0$. These same identities reduce the computation of $\operatorname{tr} \bar{\gamma}_{5} \bar{\gamma}_{i} \Delta_{1}(x, y)$ to

$$-\operatorname{tr} \bar{\gamma}_{5} \bar{\gamma}_{\underline{i}} \bar{\gamma}_{\underline{j}} \bar{\gamma}_{k} \bar{\gamma}_{\ell} \int \frac{d^{3} p}{(2 \pi)^{3}} \frac{d^{3} q}{(2 \pi)^{3}} \frac{p_{k} q_{\ell} e^{i p \cdot x} e^{-i q \cdot y}}{\left(p^{2}+M^{2}\right)\left(q^{2}+M^{2}\right)} \times \int d^{3} z e^{-i(p-q) \cdot z} \mathbf{W}_{\underline{j}}(z)$$

Compute this (introducing Feynman parameters,...) and show that it vanishes.)

From the results of the above exercise, the second term in (3.23) vanishes, and using the expression (3.20) for $I\left(M^{2}\right)$, we find that

$$I\left(M^{2}\right)=\frac{1}{2} \int d^{3} x \partial_{i} J_{i}(x, x)=\frac{1}{2} \int_{\Sigma_{\infty}} d S_{i} J_{i}(x, x)$$

where Σ_{∞} is the sphere at spatial infinity. The (formal) dimension of the moduli space of BPS- k-monopole will then be given by

$$\operatorname{dim} \mathcal{M}_{k}=\lim _{M^{2} \rightarrow 0} \int_{\Sigma_{\infty}} d S_{i} J_{i}(x, x) \tag{3.24}$$

In the remainder of this section, we will compute this integral and show that it is related to the magnetic number of the monopole.

---

*It is instructive to compare this with the calculation of the axial anomaly in four dimensions. The same calculation in four dimensions would have yielded a singularity $\sim|x-y|^{-1}$ in the $n=2$ term of the above calculation. A bit of familiar algebra would then have yielded a multiple of the Pontrjagin density for the second term in equation (3.23). In the four dimensional problem, the gauge fields go to zero at infinity and the integral of $I(x, x)$ would have received contributions only from the Pontrjagin term, since the $\partial_{i} J_{i}(x, x)$ would give a vanishing boundary term. In our case, though, the situation is different. The second term in 3.23 vanishes, whereas the boundary term coming from $\partial_{i} J_{i}(x, x)$ is not zero due to the nontrivial behaviour of the Higgs field at infinity.

9.2.3 3.2.3 Computing the current $J_i(x,x)$

We start by rewriting $J_i(x,x)$. Inserting 1 in the form $(-\bar\gamma\cdot D+M)^{-1}(-\bar\gamma.\ D+M)$ into the definition of $J_i(x,x)$, we get

$$J_i(x,x)=\operatorname{tr}\bar\gamma_5\bar\gamma_i\left\langle x\left|\frac{1}{-(\bar\gamma\cdot D)^2+M^2}(-\bar\gamma\cdot D+M)\right|x\right\rangle$$

Using the fact that the trace of an odd number of $\bar\gamma$-matrices vanishes, we remain with

$$J_i(x,x)=\operatorname{tr}\bar\gamma_5\bar\gamma_i\left\langle x\left|\frac{1}{-(\bar\gamma\cdot D)^2+M^2}\left(-\bar\gamma_jD_j+e\bar\gamma_4\boldsymbol{\phi}\right)\right|x\right\rangle \tag{3.25}$$

where now

$$-(\bar\gamma\cdot D)^2=-\left(D_i\right)^2-e^2\phi^2+\frac{1}{2}e\bar\gamma_{ij}G_{ij}+e\bar\gamma_i\bar\gamma_4D_i\phi$$

We proceed by treating the last terms $\frac{1}{2}e\bar\gamma_{ij}G_{ij}+e\bar\gamma_i\bar\gamma_4D_i\phi$ as a perturbation and expanding

$$\frac{1}{-(\bar\gamma\cdot D)^2+M^2}=\frac{1}{-\left(D_i\right)^2-e^2\phi^2+M^2}$$
$$-\frac{1}{-\left(D_i\right)^2-e^2\phi^2+M^2}\left(\frac{1}{2}e\bar\gamma_{ij}G_{ij}+e\bar\gamma_i\bar\gamma_4D_i\phi\right)\frac{1}{-\left(D_i\right)^2-e^2\phi^2+M^2}+\cdots$$

It is now time to use the fact that W_i corresponds to a monopole background. For such a background $\mathrm{G}_{ij}=O\bar{(}|x|^{-2}\Big)$ asymptotically as $|x|\to\infty$. Because we are integrating $J_i(x,\bar x)$ on Σ_∞, we are free to discard terms which decay faster than $O\left(|x|^{-2}\right)$ at infinity, hence no further terms other than those shown in the above perturbative expansion contribute. Plugging the remaining two terms of the expansion into (3.25), we notice that the first term vanishes due to the trace identity $\operatorname{tr}\bar\gamma_5\bar\gamma_i\bar\gamma_j=0$. Similar identities leave only the following terms:

$$\begin{aligned}J_i(x,x)&=e\operatorname{tr}\bar\gamma_5\bar\gamma_i\bar\gamma_j\bar\gamma_4\bar\gamma_k\left\langle x\left|\frac{1}{K}D_j\phi\frac{1}{K}D_k\right|x\right\rangle\\&-\frac{1}{2}e^2\operatorname{tr}\bar\gamma_5\bar\gamma_i\bar\gamma_{jk}\bar\gamma_4\left\langle x\left|\frac{1}{K}G_{jk}\frac{1}{K}\boldsymbol{\phi}\right|x\right\rangle+O\left(|x|^{-3}\right)\end{aligned}$$

where we have introduced the shorthand $K=-\left(D_i\right)^2-e^2\phi^2+M^2$. Using the trace identity

$$\operatorname{tr}\bar\gamma_5\bar\gamma_i\bar\gamma_j\bar\gamma_k\bar\gamma_4=4\epsilon_{ijk}$$

we can rewrite the above equation as

$$\begin{aligned}J_i(x,x)=&-4e\epsilon_{ijk}\operatorname{tr}\langle x|\frac{1}{K}D_j\phi\frac{1}{K}D_k|x\rangle\\&-2e^2\epsilon_{ijk}\operatorname{tr}\left\langle x\left|\frac{1}{K}\mathrm{G}_{jk}\frac{1}{K}\phi\right|x\right\rangle+O\left(|x|^{-3}\right)\end{aligned} \tag{3.26}$$

where the trace now refers only to the $SO(3)$ adjoint representation.

The propagators K^{-1} are not yet those of a free spinor, thanks to their dependence on $\boldsymbol{\phi}$ and W_i, thus we must treat them perturbatively as well. Since W_i decays at infinity, we can effectively put $\mathrm{W}_i=0$ in the propagators in the above expressions which are already $O\left(|x|^{-2}\right)$. On the other hand, the perturbative treatment of ϕ is a bit more subtle, since it doesn't decay at infinity but rather behaves as a homogeneous function of degree zero; that is, its behaviour on the radius $|x|$ is constant at infinity, but not so its angular dependence, which gives rise to the topological stability of the BPS-monopole. First we notice that in the adjoint representation,

$$\phi^2 \boldsymbol{v} = \boldsymbol{\phi} \times (\boldsymbol{\phi} \times \boldsymbol{v}) = -a^2 \boldsymbol{v} + (\boldsymbol{\phi} \cdot \boldsymbol{v})\boldsymbol{\phi}$$

where we have recalled that $\boldsymbol{\phi} \cdot \boldsymbol{\phi} = a^2$ at infinity. Hence on Σ_∞ we can put $K = Q + e^2\Omega$ where $Q = -\partial^2 + M^2 + a^2e^2$ and Ω is (up to a factor) the projector onto the $\boldsymbol{\phi}$ direction: $\Omega(\boldsymbol{v}) = (\boldsymbol{\phi} \cdot \boldsymbol{v})\boldsymbol{\phi}$. Because $[Q, \Omega] = O\left(|x|^{-1}\right)$ asymptotically, we can effectively treat these two operators as commuting, whence we can write

$$\frac{1}{K} = \frac{1}{Q} + \sum_{n \geq 1} \frac{1}{Q^{n+1}} \left(e^2\Omega\right)^n + O\left(|x|^{-1}\right)$$

Notice moreover that as operators in the adjoint representation of $SO(3)$,

$$(\Omega \circ \boldsymbol{\phi})\boldsymbol{v} = \Omega(\boldsymbol{\phi} \times \boldsymbol{v}) = \boldsymbol{\phi} \cdot (\boldsymbol{\phi} \times \boldsymbol{v})\boldsymbol{\phi} = 0 \tag{3.27}$$

We are now in a position to prove that the first term in equation (3.26) doesn't contribute.
Exercise 3.10 (The first term doesn't contribute)
Prove that

$$\operatorname{tr}\left\langle x \left| \frac{1}{K} D_j \phi \frac{1}{K} D_k \right| x \right\rangle = O\left(|x|^{-3}\right)$$

whence it doesn't contribute to the integral over Σ_∞.
(Hint: First notice that $D_k = \partial_k + O\left(|x|^{-1}\right)$, whence up to $O\left(|x|^{-3}\right)$ we can simply substitute ∂_k for D_k in the above expression. Now use equation (3.27) and the fact that ϕ and $D_j\phi$ are parallel in Σ_∞, to argue that one can substitute K for Q in the above expression (again up to terms of order $O\left(|x|^{-3}\right)$). Then simply take the trace to obtain the result.)
Hence we are left with

$$J_i(x,x) = -2e^2\epsilon_{ijk} \operatorname{tr}\left\langle x \left| \frac{1}{K} \mathrm{G}_{jk} \frac{1}{K} \boldsymbol{\phi} \right| x \right\rangle + O\left(|x|^{-3}\right) \tag{3.28}$$

Since G_{jk} is parallel to ϕ on $\Sigma_\infty, \Omega \circ \mathrm{G}_{jk} = 0$ by (3.27). Thus we are free to substitute the free propagator Q^{-1} for K^{-1} in the above expression, to obtain:

$$J_i(x,x) = -2e^2\epsilon_{ijk} \int d^3x' \operatorname{tr} \boldsymbol{\phi}(x) \mathrm{G}_{jk}\left(x'\right) Q^{-1}\left(x, x'\right) Q^{-1}\left(x', x\right) + O\left(|x|^{-3}\right)$$

where the free propagator is given by

$$Q^{-1}(x,y) = \int \frac{d^3p}{(2\pi)^3} \frac{e^{ip\cdot(x-y)}}{p^2 + M^2 + a^2e^2}$$

Changing variables $x' \mapsto y = x - x'$, and using

$$\operatorname{tr} \boldsymbol{\phi}(x) \mathrm{G}_{jk}(x-y) = \operatorname{tr} \boldsymbol{\phi}(x) \mathrm{G}_{jk}(x) + O\left(|x|^{-3}\right)$$

we remain with

$$\begin{aligned} J_i(x,x) = -2e^2\epsilon_{ijk} \operatorname{tr} \mathrm{G}_{jk}(x)\boldsymbol{\phi}(x) \int d^3y \frac{d^3p}{(2\pi)^3} \frac{d^3q}{(2\pi)^3} e^{i(p-q)\cdot y} \\ \times \frac{1}{\left(p^2 + M^2 + a^2e^2\right)\left(q^2 + M^2 + a^2e^2\right)} + O\left(|x|^{-3}\right) \end{aligned}$$

The y-integral gives $2\pi^3\delta(p-q)$, which gets rid of the q-integral and we remain with

$$J_i(x,x) = -2e^2\epsilon_{ijk}\,\mathrm{tr}\,\phi(x)\mathrm{G}_{jk}(x)\int\frac{d^3p}{(2\pi)^3}\frac{1}{(p^2+M^2+a^2e^2)^2} + O\left(1/|x|^3\right)$$

The p-integral is readily evaluated

$$\int\frac{d^3p}{(2\pi)^3}\frac{1}{(p^2+M^2+a^2e^2)^2} = \frac{1}{8\pi}\frac{1}{\sqrt{M^2+a^2e^2}}$$

whence

$$\int_{\Sigma_\infty} dS_i J_i(x,x) = \frac{1}{2\pi}\frac{e^2}{\sqrt{M^2+a^2e^2}}\int_{\Sigma_\infty} dS_i\epsilon_{ijk}\boldsymbol{\phi}\cdot\mathrm{G}_{jk}$$

where have also used that the trace in the adjoint representation is normalised so that $\mathrm{tr}\,AB = -2A\cdot B$.

Exercise 3.11 (Another expression for the degree of the map ϕ)

Prove that on Σ_∞,

$$\phi\cdot\mathrm{G}_{jk} = \frac{1}{ea^2}\phi\cdot(\partial_j\phi\times\partial_k\phi)$$

and, comparing with equation (1.22), deduce that the degree of the map ϕ from Σ_∞ to the sphere of radius a in $\mathbb{R}^3$ is given by

$$\deg\boldsymbol{\phi} = \frac{e}{8\pi a}\int_{\Sigma_\infty} dS_i\epsilon_{ijk}\boldsymbol{\phi}\cdot\mathrm{G}_{jk}$$

(Hint: Use that $\phi\times D_j\phi = 0$ on Σ_∞, and expand $0 = \phi\cdot(D_j\phi\times D_k\phi)$.)

From the results of this exercise and the fact that for a k-monopole solution, the degree of ϕ is k, we can write

$$\int_{\Sigma_\infty} dS_i J_i(x,x) = \frac{4aek}{\sqrt{M^2+a^2e^2}}$$

whence plugging this into the equation (3.24) for the formal dimension of $\mathcal{N}_k$, we find that

$$\dim\mathcal{M}_k = 4k$$

9.3 3.3 A quick motivation of hyperkähler geometry

In the next section we will prove that the natural metric on $\mathcal{M}_k$ induced by the Yang-Mills-Higgs functional is hyperkähler; but first we will briefly review the necessary notions from riemannian geometry leading to hyperkähler manifolds. The reader familiar with this topic can easily skip this section.

*If you are familiar with the calculation of the number of instanton parameters, you may be surprised by the explicit M^2 dependence in the expression $\int_{\Sigma_\infty} dS_i J_i(x,x)$. This is due to the asymptotic behaviour of ϕ, which prevents the above calculation from being tackled by the methods usually applied to index theorems on compact spaces. For the index of an operator on a compact space, or similarly for fields which decay at infinity, one can prove that the result of the above integral is actually independent of M^2, hence one can compute the integral for already in the limit $M^2\to 0$ (cf. the Witten index). Here we are in fact faced essentially with the calculation of the index of an operator on a manifold with boundary, for which a satisfactory Witten-index treatment is lacking, to the best of my knowledge.

Hyperkähler geometry is probably best understood from the point of view of holonomy groups in riemannian geometry. In this section we review the basic notions. Sadly, the classic treatises on the holonomy approach to riemannian geometry KN63, KN69, Lic76 stop just short of hyperkähler geometry; but two more recent books Bes86, Sal89] on the subject do treat the hyperkähler case, albeit from a slightly different point of view than the one adopted here. We direct the mathematically inclined reader to the classics for the basic results on riemannian and Kähler geometry, which we will only have time to review ever so briefly in these notes; and to the newer references for a more thorough discussion of hyperkähler manifolds. All our manifolds will be assumed differentiable, as will be any geometric object defined on them, unless otherwise stated.

9.3.1 3.3.1 Riemannian geometry

Any manifold M admits a riemannian metric. Fix one such metric g. On the riemannian manifold (M, g) there exists a unique linear connection ∇ which is torsion-free

$$[X, Y] = \nabla_X Y - \nabla_Y X \quad \text{for any vector fields } X, Y \text{ on } M \tag{3.29}$$

and preserves the metric $\nabla g = 0$. It is called the Levi-Civita connection and relative to a local chart x^a, it is defined by the Christoffel symbols ${\Gamma_{ab}}^c$ which in turn are defined by

$$\nabla_a \partial_b = \Gamma^c_{ab} \partial_c$$

where we have used the shorthand $\nabla_a = \nabla_{\partial_a}$. The defining properties of the Levi-Civita connection are sufficient to express the Christoffel symbols in terms of the components g_{ab} of the metric:

$$\Gamma^c_{ab} = \frac{1}{2} g^{cd} \left(\partial_a g_{db} + \partial_b g_{ad} - \partial_d g_{ab} \right) \tag{3.30}$$

which proves the uniqueness of the Levi-Civita connection.

Exercise 3.12 (A coordinate-free expression for ∇)

Using the defining conditions of the Levi-Civita connection ∇, prove that

$$\begin{aligned} 2 \langle Z, \nabla_X Y \rangle = X \langle Y, Z \rangle + Y \langle Z, X \rangle - Z \langle X, Y \rangle \\ - \langle [X, Z], Y \rangle - \langle [Y, Z], X \rangle - \langle [X, Y], Z \rangle \end{aligned} \tag{3.31}$$

where we have used the notation $\langle X, Y \rangle = g(X, Y)$. Equation (3.30) follows after substituting ∂_a, ∂_b and ∂_c for X, Y, and Z respectively.

With ∇ we can give meaning to the notion of parallel transport. Given a curve $t \rightsquigarrow \gamma(t)$ on M with velocity vector $\dot\gamma$, we say that a vector field X is parallel along γ if $\nabla_{\dot\gamma} X = 0$. Relative to a local coordinate chart x^a, we can write this equation as

$$\frac{DX^b}{dt} \equiv \dot\gamma^a \nabla_a X^b = \dot\gamma^a \left(\partial_a X^b + \Gamma^b_{ac} X^c \right) = \dot X^b + {\Gamma_{ac}}^b \dot\gamma^a X^c = 0 \tag{3.32}$$

A curve γ is a geodesic if its velocity vector is self-parallel: $\nabla_{\dot\gamma} \dot\gamma = 0$. In terms of (3.32) we arrive at the celebrated geodesic equation:

$$\ddot\gamma^c + {\Gamma_{ab}}^c \dot\gamma^a \dot\gamma^b = 0 \tag{3.33}$$

This equation follows by extremising the action with lagrangian

$$L(x, \dot{x}) = \frac{1}{2} g_{ab}(x) \dot{x}^a \dot{x}^b$$

whence our claim at the end of section 3.1.3 that free motion on a riemannian manifold is geodesic.

We can integrate equation (3.32) and arrive at the concept of paralleltransport. More concretely, associated with any curve $\gamma : [0,1] \to M$ there is a linear map $\mathbb{P}_\gamma : T_{\gamma(0)}M \to T_{\gamma(1)}M$ taking vectors tangent to M at $\gamma(0)$ to vectors tangent to M at $\gamma(1)$. If $X \in T_pM$ is a tangent vector to M at $p \equiv \gamma(0)$ we define its parallel transport $\mathbb{P}_\gamma(X)$ relative to γ by first extending X to a vector field along γ in a way that solves (3.32), and then simply evaluating the vector field at $\gamma(1)$.

Now fix a point $p \in M$ and let γ be a piecewise differentiable loop based at p, that is, a piecewise differentiable curve which starts and ends at p. Then $\mathbb{P}_\gamma$ is a linear map $T_pM \to T_pM$. We can compose these maps: if γ and γ' are two loops based at p, then $\mathbb{P}_{\gamma'} \circ \mathbb{P}_\gamma$ is the linear map corresponding to parallel transport on the loop based at p obtained by first tracing the path γ and then γ'. (Notice that this new loop may not be differentiable even if γ and γ' are: but it is certainly piecewise differentiable, hence the need to consider such loops from the outset.) Also $\mathbb{P}_\gamma$ is invertible: simply trace the path γ backwards in time. Therefore the transformations $\{\mathbb{P}_\gamma\}$ form a group. If we restrict ourselves to loops which are contractible, the group of linear transformations:

$$H(p) = \{\mathbb{P}_\gamma \mid \gamma \text{ a contractible loop based at } p\}$$

is called the (restricted) holonomy group at p of the connection ∇. It can be shown that to be a Lie group.

One should hasten to add that there is no reason to restrict ourselves to the Levi-Civita connection. We will be mostly interested in the classical
case, where ∇ is the Levi-Civita connection, but these definitions make sense in more generality.

Exercise 3.13 (The holonomy group of a connected manifold)

Prove that if two points p and q in M can be joined by a path in M, their holonomy groups $H(p)$ and $H(q)$ are conjugate and therefore isomorphic.

(Hint: Use parallel transport along the path joining p and q to provide the conjugation.)

Hence it makes sense to speak of the holonomy group of a connected manifold M. From now on all we will only concern ourselves with connected manifolds. A further useful restriction that one can impose on the type of manifolds we consider is that of irreducibility. A manifold is said to be (ir) reducible relative to a linear connection ∇ if the tangent space at any point is an (ir)reducible representation of the holonomy group. Clearly the holonomy group (relative to the Levi-Civita connection) of product manifold $M \times M'$ with the product metric acts reducibly. A famous theorem of de Rham's provides a converse. This theorem states that if a simplyconnected complete riemannian manifold M is reducible relative to the LeviCivita connection, then $M = M' \times M''$ isometrically. We will restrict ourselves in what follows to irreducible manifolds.

For a generic linear connection on an irreducible manifold M, the holonomy group is (isomorphic to) $GL(m)$, where $m = \dim M$. However, the Levi-Civita connection is far from generic as the following exercise shows.

Exercise 3.14 (The holonomy group of a riemannian manifold)

Prove that the holonomy group of an m-dimensional riemannian manifold (relative to the Levi-Civita connection) is actually in $SO(m)$.

(Hint: Show that $\nabla g = 0$ implies that the parallel transport operation $\mathbb{P}_\gamma$ preserves the norm of the vectors, whence the holonomy group is in $O(m)$. Argue that since we consider only contractible loops, the holonomy group is connected and hence it must be in $SO(m)$. By the way, the same would hold for orientable manifolds even if considering non-contractible loops.)

A celebrated theorem of Ambrose and Singer tells us that the Lie algebra of the holonomy group is generated by the Riemann curvature tensor in the following way. Recall that the Riemann curvature tensor is defined as follows. Fix vector fields X and Y on M, and define a linear map from vector fields to vector fields as follows:

$$R(X,Y) = [\nabla_X, \nabla_Y] - \nabla_{[X,Y]}$$

It is easy to prove that this map is actually tensorial in X and Y. Indeed, relative to a coordinate basis, it may be written out as a tensor ${R_{abc}}^d$ defined
by

$$R\left(\partial_a, \partial_b\right)\partial_c = {R_{abc}}^d \partial_d$$

and therefore has components

$${R_{abc}}^d = \partial_a {\Gamma_{bc}}^d + {\Gamma_{bc}}^e {\Gamma_{ae}}^d - \partial_b {\Gamma_{ac}}^d - {\Gamma_{ac}}^e {\Gamma_{be}}^d \tag{3.34}$$

Then the Lie algebra of the holonomy group is the Lie subalgebra of $gl(m)$ spanned by the curvature operators $R_{ab} : \partial_c \mapsto {R_{abc}}^d \partial_d$.

Exercise 3.15 (The holonomy algebra of a riemannian manifold)

Using the Ambrose-Singer theorem this time, prove a second time that the holonomy group of a riemannian manifold lies in $SO(m)$, by showing that its Lie algebra lies in so (m). In other words, prove that each curvature operator R_{ab} (for fixed a and b) is antisymmetric:

$$R_{abcd} = -R_{abdc} \quad \text{where} \quad R_{abcd} = {R_{abc}}^e g_{ed}$$

Adding more structure to a riemannian manifold in a way that is consistent with the metric restricts the holonomy group further. Next we will discuss what happens when we add a complex structure.

9.3.2 3.3.2 Kähler geometry

An almost complex structure is a linear map $I : TM \to TM$ which obeys $I^2 = -\mathbb{1}$. This gives each tangent space T_pM the structure of a complex vector space, since we can multiply a tangent vector X by a complex number $z = x + iy$ simply by $z \cdot X = xX + yI(X)$. In particular, it means that the (real) dimension of each T_pM and hence of M must be even: $2n$, say. We will also assume that the complex structure I is compatible with the metric in the sense that $g(IX, IY) = g(X,Y)$ for all vector fields X and Y. Another way to say this is that the metric g is hermitian relative to the complex structure I.

If we complexify the tangent space, we can diagonalise the complex structure. Clearly the eigenvalues of I are $\pm i$. Complex vector fields Z for which $IZ = iZ$ are said to be of type $(1,0)$, whereas those for which $IZ = -iZ$ are of type $(0,1)$. If we can introduce local complex coordinates $(z^\alpha, \bar{z}^{\bar{\alpha}})$, $\alpha, \bar{\alpha} = 1, \ldots n$, relative to which a basis for the $(1,0)$ (resp. $(0,1)$) vector fields is given by ∂_α (resp. $\partial_{\bar{\alpha}}$) and if when we change charts the local complex coordinates are related by biholomorphic transformations, then we say that I is integrable.

A hard theorem due to Newlander and Nirenberg translates this into a beautiful local condition on the complex structure. According to the

Newlander-Nirenberg theorem, an almost complex structure I is integrable if and only if the Lie bracket of any two vector fields of type $(1,0)$ is again of type $(1,0)$. This in turns translates into the vanishing of a tensor.

Exercise 3.16 (The Nijenhuis tensor)

Using the Newlander-Nirenberg theorem prove that I is integrable if and only if the following tensor vanishes:

$$N_I(X,Y) = I[IX,IY] + [X,IY] + [IX,Y] - I[X,Y]$$

N_I is known as the Nijenhuis tensor of the complex structure I. It is easy to prove that the Nijenhuis tensor N_I vanishes in a complex manifold (do it!)-it is the converse that is hard to prove.

Now suppose that ∇ is a linear connection relative to which I is parallel: $\nabla I = 0$. Let's call this a complex connection.

Exercise 3.17 (The holonomy group of a complex connection)

Let H denote the holonomy group of a complex connection on a complex manifold M. Prove that $H \subseteq GL(n,\mathbb{C}) \subset GL(2n,\mathbb{R})$.

(Hint: It's probably easiest to prove the equivalent statement that the holonomy algebra is a subalgebra of $gl(n,\mathbb{C})$. Choose a basis for $T_pM \cong \mathbb{R}^{2n}$ in which the complex structure I has the form

$$I = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}$$

where $\mathbb{1}$ is the $n \times n$ unit matrix. Argue that the curvature operators R_{ab} commute with I, whence in this basis, they are of the form

$$\begin{pmatrix} A & B \\ -B & A \end{pmatrix}$$

where A and B are arbitrary real $n \times n$ matrices. This then corresponds to the real $2n$-dimensional representation of the matrix $A + iB \in gl(n,\mathbb{C})$.)

If ∇ is the Levi-Civita connection, then the holonomy lies in the intersection $GL(n,\mathbb{C}) \cap SO(2n) \subset GL(2n,\mathbb{R})$.

Exercise 3.18 (The unitary group)

Prove that $GL(n,\mathbb{C}) \cap SO(2n) \subset GL(2n,\mathbb{R})$ is precisely the image of the real $2n$ dimensional representation of the unitary group $U(n)$.

(Hint: Prove the equivalent statement for Lie algebras. In the basis of the previous exercise, prove that a matrix in $so(2n)$ has the form

$$\begin{pmatrix} A & B \\ -B^t & D \end{pmatrix}$$

where $A^t = -A, D^t = -D$ and B are otherwise arbitrary real $n \times n$ matrices. If the matrix is also in $gl(n,\mathbb{C})$, we know that $A = D$ and that $B = B^t$. Thus matrices in $gl(n,\mathbb{C}) \cap so(2n) \subset gl(2n,\mathbb{R})$ are of the form

$$\begin{pmatrix} A & B \\ -B & A \end{pmatrix}$$

where $A^t = -A$ and $B^t = B$, which corresponds to the complex matrix $A + iB \in gl(n,\mathbb{C})$. Prove that this matrix is anti-hermitian, whence in $u(n)$.

If the Levi-Civita connection is complex, so that the holonomy lies in $U(n)$, the manifold (M, g, I) is said to be Kähler. In other words, Kähler geometry is the intersection, so to speak, of riemannian and complex geometries.

There is another perhaps more familiar definition of Kähler manifolds, involving the Kähler form.

Exercise 3.19 (The Kähler form)

Given a complex structure I relative to which g is hermitian, we define a 2-form ω by

$$\omega(X, Y) = g(X, IY) \quad \text{or equivalently} \quad \omega_{ab} = I_a{}^c g_{bc}$$

Prove that $\omega(X, Y) = -\omega(Y, X)$, so that it is in fact a form. Prove that it is actually of type $(1, 1)$.

An equivalent definition of a Kähler manifold is that (M, g, I) is Kähler if and only if ω is closed. These two definitions can of course be reconciled. We review this now.

Exercise 3.20 (Kähler is Kähler is Kähler)

Let (M, g, I) be a complex riemannian manifold, where g is hermitian relative to I. Let ω be the corresponding Kähler form and ∇ the Levi-Civita connection. Prove that the following three conditions are equivalent (and are themselves equivalent to (M, g, I) being a Kähler manifold):

(a) $\nabla I = 0$;

(b) $\nabla\omega = 0$; and

(c) $d\omega = 0$.

(Hint: $(a) \Leftrightarrow (b)$ is obvious. For $(b) \Rightarrow (c)$ simply antisymmetrise and use the fact that ∇ is torsion-less, which implies the symmetry of the Christoffel symbols in the lower two indices. The trickiest calculation is $(c) \Rightarrow (b)$, and we break this up into several steps:

From (3.29), deduce that I is integrable if and only if

$$\nabla_a \omega_{bc} + I_c^d I_b^e \nabla_e \omega_{ad} - (a \leftrightarrow b) = 0$$

From the fact that g is hermitian relative to I, deduce that

$$I_c{}^d \nabla_a \omega_{bd} = -I_b{}^d \nabla_a \omega_{cd}$$

Using the previous two steps, show that

$$\nabla_a \omega_{bc} = -I_b{}^d I_c{}^e \nabla_a \omega_{de}$$

Finally, use these formulae to show that $d\omega = 0 \Rightarrow \nabla\omega = 0$.)

Conversely, one can show that if the holonomy group of a $2n$-dimensional riemannian manifold is contained in $U(n)$, then the manifold is Kähler. The proof is paradigmatic of the more algebraic approach to the study of holonomy, which has begotten some of the more remarkable results in this field. We will therefore allow ourselves a brief digression. We urge the reader to take a look at the books [Bes86, Sal89] for a more thorough treatment.

For simplicity, we start with a torsionless connection ∇. The fundamental elementary observation is that there is a one-to-one correspondence between covariantly constant tensors and singlets of the holonomy group. (Clearly if $\nabla t = 0$, then t is invariant under the holonomy group; conversely, if t is invariant under the holonomy group, taking the derivative of the parallel transport of t along the path is zero, but to first order this is precisely ∇t.) In turn, singlets of the holonomy group determine to a large extent the geometry of M. For example, suppose that M is an m-dimensional irreducible manifold with holonomy group $G \subset GL(m, \mathbb{R})$.

Irreducibility means that the fundamental m-dimensional representation of $GL(m, \mathbb{R})$ remains irreducible under G. Let us call this representation T - the " T " stands for tangent space. Under the action of G, tensors on M will transform according to tensor powers of the representation T. For example, 1-forms will transform according to the dual representation T^*, symmetric rank p tensors will transform as $S^p T^*$, whereas p-forms will transform as $\bigwedge^p T^*$, and so on. We can then break up all these tensorial representations in terms of irreducibles and, in particular, exhibit all the singlets. These singlets will correspond, by the observation made above, in a one-to-one fashion with covariantly constant tensors. Let us run through some examples.

Suppose that $G = SO(m)$. Then T is the fundamental m-dimensional representation of $SO($ m$)$. We know that, in particular, there is a singlet $\bar{g} \in S^2 T^*$ and moreover that the map $T \rightarrow T^*$ defined by this $\bar{g}$ is nondegenerate. Hence by the fundamental observation, there exists a covariantly constant tensor g which can be thought of as a riemannian metric. By the uniqueness of the Levi-Civita connection, it follows that ∇ is the Levi-Civita connection associated to g. In other words, manifolds with $SO(m)$ holonomy
relative to a torsionless connection are simply riemannian manifolds. Well, not just any riemannian manifold. There is another $SO(m)$-invariant tensor $\bar{\Omega} \in \Lambda^m T^*$. The covariantly constant m-form Ω defines an orientation on M. In fact, one can show that there are no other invariant tensors which are algebraic independent from these ones, so that manifolds with $SO(m)$ holonomy (again, relative to a torsionless connection) are precisely orientable riemannian manifolds.

Now suppose that the dimension of M is even: $m = 2n$, say, and that $G = U(n) \subset GL(2n, \mathbb{R})$. Then T is the real $2n$ irreducible representation of $U(n)$, whose complexification breaks up as $T^{\mathbb{C}} = T' \oplus T''$, where T' is the complex n-dimensional (fundamental) representation of $U(n)$, and $T'' = \overline{T'}$ is its conjugate. Since $G \subset SO(2n)$, we know from the previous paragraph that M is riemannian and orientable, and that we can think of ∇ as the LeviCivita connection of this metric. However there is also a singlet $\bar{\omega} \in \bigwedge^2 T^*$. The resulting covariantly constant 2 -form ω is precisely the Kähler form. Hence manifolds with $U(n)$-holonomy are precisely the Kähler manifolds.

9.3.3 3.3.3 Ricci flatness

We can now restrict the holonomy of a Kähler manifold a little bit further by imposing constraints on the curvature: namely that it be Ricci-flat. As we will see, this is equivalent to demanding that the holonomy lie in $SU(n) \subset U(n)$. As Lie groups, $U(n) = U(1) \times SU(n)$. If we think of $U(n)$ as unitary matrices, the $U(1)$ factor is simply the determinant. Hence the manifold will have $SU(n)$ holonomy provided that the determinant of every parallel transport operator $\mathbb{P}_\gamma$ is equal to 1.

Geometrically, the determinant can be understood as follows. Suppose that M is a Kähler manifold and let's look at how forms of type $(n, 0)$ (or $(0, n)$) transform under parallel transport. At a fixed point p in M, the space of such forms is 1-dimensional. Hence if θ is an $(n, 0)$-form, then $\mathbb{P}_\gamma \theta = \lambda_\gamma \theta$ where λ_γ is a complex number of unit norm.

Exercise 3.21 (The determinant of $\mathbb{P}_\gamma$)

Prove that λ_γ is the determinant of the linear map $\mathbb{P}_\gamma : T_p M \rightarrow T_p M$.

Therefore the holonomy lies in $SU(n)$ if and only if $\lambda_\gamma = 1$ for all γ. By our previous discussion, it means that there is a nonzero parallel $(n, 0)$ form θ. Since parallel forms have constant norm, this form if nonzero at some point is nowhere vanishing, hence the bundle of $(n, 0)$-forms is trivial. Equivalently this means that the first Chern class of the manifold vanishes. Such Kähler manifolds are known as Calabi-Yau manifolds, after the celebrated conjecture of Calabi, proven by Yau. Calabi's conjecture stated that
given a fixed Kähler manifold with vanishing first Chern class, there exists a unique Ricci-

flat Kähler metric in the same Kähler class. The Calabi conjecture (now theorem) allows us to construct manifolds admitting Ricci-flat Kähler metrics, by the simpler procedure of constructing Kähler manifolds with vanishing first Chern class. Algebraic geometry provides us with many constructions of such manifolds: as algebraic varieties of complex projective space, for example. The catch is that the Ricci-flat Kähler metric is most definitely not the induced metric. In fact, the form of the metric is very difficult to determine. Even for relatively simple examples like $K3$, the metric is not known.

It follows from this conjecture (now theorem) that an irreducible Kähler manifold has $SU(n)$ holonomy if and only if it is Ricci-flat. We don't need to appeal to the Calabi conjecture to prove this result, though, as we now begin to show.

Let's first recall how the Ricci tensor is defined. If X and Y are vector fields on $M, \operatorname{Ric}(X,Y)$ is defined as the trace of the map $V \mapsto R(V,X)Y$, or relative to a local chart

$$S_{ab} \equiv \operatorname{Ric}\left(\partial_a, \partial_b\right) = R^c_{cab}$$

Exercise 3.22 (The Ricci tensor is symmetric)

Prove that $S_{ab} = S_{ba}$.

In a Kähler manifold and relative to complex coordinates adapted to I, many of the components of the Ricci and Riemann curvature tensors are zero.

Exercise 3.23 (Curvature tensors in Kähler manifolds)

Let $(z^\alpha, \bar{z}^{\bar{\alpha}})$ be complex coordinates adapted to the complex structure I; that is, the corresponding vector fields are of type $(1,0)$ and $(0,1)$ respectively: $I\left(\partial_\alpha\right) = i\partial_\alpha$ and $I\left(\partial_{\bar{\alpha}}\right) = -i\partial_{\bar{\alpha}}$. Prove that the metric has components $g_{\alpha\bar{\beta}}$, and that the Christoffel symbols have components $\Gamma_{\alpha\beta}{}^{\gamma}$ and $\Gamma_{\bar{\alpha}\bar{\beta}}{}^{\bar{\gamma}}$. Prove that the only nonzero components of the Riemann curvature are $R_{\alpha\bar{\beta}\gamma}{}^{\delta} = -R_{\bar{\beta}\alpha\gamma}{}^{\delta}$ and $R_{\alpha\bar{\beta}\bar{\gamma}}{}^{\bar{\delta}} = -R_{\bar{\beta}\alpha\bar{\gamma}}$. Finally deduce that the Ricci tensor has components $S_{\alpha\bar{\beta}}$, so that $S_{\alpha\beta} = S_{\bar{\alpha}\bar{\beta}} = 0$.

The holonomy algebra of a Kähler manifold is $u(n)$, whence for fixed a and b, the curvature operator R_{ab} belongs to $u(n) = u(1) \times su(n)$. How can one extract the $u(1)$-component? For this we need to recall Exercise 3.18. If $A + iB \in u(n), \exp(A+iB) \in U(n)$ and the $U(1)$ component is the determinant: $\det \exp(A+iB) = \exp(\operatorname{tr}(A+iB))$. Since $A^t = -A$, it is traceless, whence $\det \exp(A+iB) = \exp(i \operatorname{tr} B)$. Hence if we let $i\mathbb{1} \in u(n)$ be the generator of the $u(1)$ subalgebra, the $u(1)$ component of a matrix in $u(n)$ is just its trace. In a Kähler manifold, the holonomy representation of
$U(n)$ is real and $2n$-dimensional, which means that a matrix $A + iB \in u(n)$ is represented by a real $2n \times 2n$ matrix

$$Q = \begin{pmatrix} A & B \\ -B & A \end{pmatrix}$$

and therefore its $u(1)$-component, $\operatorname{tr} B$, is simply given by

$$\operatorname{tr} B = -\frac{1}{2} \operatorname{tr}\left[\begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}\begin{pmatrix} A & B \\ -B & A \end{pmatrix}\right] = -\frac{1}{2}\operatorname{tr}(IQ)$$

From this it follows that the $u(1)$-component F_{ab} of the curvature operator is given by $F_{ab} = -\frac{1}{2}\operatorname{tr}\left(I \circ R_{ab}\right) = -\frac{1}{2} R_{abc}{}^{d} I_d{}^{c}$. The next exercise asks you to show that this is essentially the Ricci curvature, from where it follows that Ricci-flat Kähler manifolds have $SU(n)$ holonomy and viceversa.

Exercise 3.24 (An equivalent expression for the Ricci curvature)

Prove that the Ricci curvature on a Kähler manifold can be also be defined by

$$\operatorname{Ric}(X, Y) = \frac{1}{2}\operatorname{tr}(V \mapsto I \circ R(X, IY)V)$$

or equivalently,

$$S_{ac}I_b{}^c = -\frac{1}{2}\operatorname{tr}(I \circ R_{ab}) = F_{ab},$$

which, relative to complex coordinates, becomes

$$S_{\alpha\bar{\beta}} = iF_{\alpha\bar{\beta}}$$

Using the above results, give another proof of the symmetry of the Ricci tensor: $S_{ab} = S_{ba}$. (Compare with Exercise 3.22.

9.3.4 3.3.4 Hyperkähler geometry

Finally we define hyperkähler manifolds. In a hyperkähler manifold we have not one but three parallel almost complex structures I, J, and K which satisfy the quaternion algebra:

$$IJ = K = -JI,\ JK = I = -KJ, \quad KI = J = -IK$$
$$I^2 = J^2 = K^2 = -\mathbb{1},$$

and such that the metric is hermitian relative to all three. Notice that we don't demand that the complex structures be integrable. This is a consequence of the definition.

Exercise 3.25 (Hyperkähler implies integrability)

Let (M, g, I, J, K) be a hyperkähler manifold. Prove that I, J and K are integrable complex structures.

(Hint: Associated to each of the almost complex structures there is a 2-form: ω_I, ω_J and ω_K. Because $\nabla I = 0$, Exercise 3.20 implies that $d\omega_I = 0$, and similarly for J and K. Notice that $\omega_J(X, Y) = g(X, JY) = g(X, KIY) = \omega_K(X, IY)$, whence [2] $\imath_X\omega_J = \imath_{IX}\omega_K$. A complex vector field X is of type $(1, 0)$ with respect to I, if and only if $\imath_X\omega_J = i\imath_X\omega_K$. By the Newlander-Nirenberg theorem, it is sufficient to prove that the Lie bracket of two such complex vector fields also obeys the same relation. But this is a simple computation, where the fact that the forms ω_J and ω_K are closed is used heavily. The same proof holds mutatis mutandis for J and K.)

Just like an almost complex structure I on a manifold allows us to multiply vector fields by complex numbers and hence turn each tangent space into a complex vector space, the three complex structure in a hyperkähler manifold allow us to multiply by quaternions. Concretely, if $q = x + iy + jz + kw \in \mathbb{H}$ is a quaternion, and X is a vector field on M, then we define

$$q \cdot X \equiv xX + yI(X) + zJ(X) + wK(X)$$

This turns each tangent space into a quaternionic vector space (a left $\mathbb{H}$ module, to be precise) and, in particular, this means that hyperkähler manifolds are $4k$-dimensional.

One can prove, just as we did with complex manifolds, that the holonomy group of a hyperkähler manifold lies in $3Sp(2k) \subseteq SU(2k) \subset GL(4k)$. In particular, hyperkähler manifolds are Ricci-flat.

Exercise 3.26 (The holonomy group of a hyperkähler manifold)

Prove that the holonomy group of a hyperkähler manifold is a subgroup of USp $(2k)$. (Hint: Depending on how one looks at this, there may be nothing that needs proving. If we take as definition of $USp(2k) \subset GL(2k, \mathbb{C})$ those matrices which commute with the natural action of the quaternions on $\mathbb{C}^{2k} \cong \mathbb{H}^k$, then the result is immediate since the fact that I, J, and K are parallel means that ∇ and hence the curvature operators commute with multiplication by $\mathbb{H}$. If you have another definition of $USp(2k)$ in mind, then the exercise is to reconcile both definitions.)

Conversely, if the holonomy group of a manifold M is a subgroup of $USp(2k) \subset SO(4k)$, then decomposing tensor powers of the fundamental $4k$ dimensional representation T of $SO(2k)$ into $USp(2k)$-irreducibles, we find that $\Lambda^2 T^*$ possesses three singlets. The resulting covariantly constant 2 forms are of course the three Kähler forms of M. A little bit closer inspection shows that the associated complex structures obey the quaternion algebra, so that M is hyperkähler.

9.4 3.4 $\mathcal{M}_k$ is hyperkähler

In this section we prove that the metric on $\mathcal{M}_k$ defined by the kinetic term in the Yang-Mills-Higgs functional is hyperkähler. We will prove this in two ways. First we will prove that the configuration space $\mathcal{A}$ is hyperkähler and that $\mathcal{M}_k$ is its hyperkähler quotient. In order to do this we will review the notion of a Kähler quotient which should be very familiar as a special case of the symplectic quotient of Marsden-Weinstein which appears in physics whenever we want to reduce a phase space with first-class constraints. In a nutshell, the group $\mathcal{G}$ of gauge transformations acts on $\mathcal{A}$ preserving the hyperkähler structure and the resulting moment mapping is nothing but the Bogomol'nyi equation. We will also give a computationally more involved proof of the hyperkähler nature of $\mathcal{N}_k$, which is independent of the hyperkähler quotient, at least on the face of it.

9.4.1 3.4.1 Symplectic quotients

Let (M, ω) be a symplectic manifold-that is, ω is a nondegenerate closed 2-form-and let G be a Lie group acting on M in a way that preserves ω. Let $\mathfrak{g}$ denote the Lie algebra of G and let $\{e_a\}$ be a basis for $\mathfrak{g}$ which we fix once and for all. To each e_a there is an associated vector field X_a on M. The fact that G preserves ω means that $\mathcal{L}_{X_a}\omega = 0$ for every X_a, where $\mathcal{L}_{X_a}$ is the Lie derivative along X_a. We will often abbreviate $\mathcal{L}_{X_a}$ by $\mathcal{L}_a$.

Exercise 3.27 (The Lie derivative acting on forms)

Prove that if ω is a differential form on M and X is any vector field, then the Lie derivative $\mathcal{L}_X\omega$ is given by

$$\mathcal{L}_X\omega = (d\imath_X + \imath_X d)\,\omega$$

where d is the exterior derivative and $\imath_X$ is the contraction operator characterised uniquely by the following properties:

(a) $\imath_X f = 0$ for all functions f;

(b) $\imath_X\alpha = \alpha(X)$ for all one-forms α; and

(c) $\imath_X(\alpha \wedge \beta) = (\imath_X\alpha) \wedge \beta + (-1)^p \alpha \wedge \imath_X\beta$, for α a p-form and β any form.

Because $d\omega = 0, \mathcal{L}_a\omega = 0$ is equivalent to the one-form $\imath_a\omega$ being closed, where $\imath_a \equiv \imath_{X_a}$. Let us assume that this form is also exact, so that there is a function μ_a such that $\imath_a\omega = d\mu_a$. This would be guaranteed, for example, if M were simply-connected, or if $\mathfrak{g}$ were semi-simple.

*2 The conventions for the interior product $\imath_X$ are summarised in Exercise 3.27.

3 Mathematicians call $Sp(k)$ what we call $USp(2k)$.

More precise conditions on the absence of this obstruction can be written down but we won't need them in what follows. The functions μ_a allow us to define a moment mapping $\mu : M \to \mathfrak{g}^*$ by $\mu(p) = \mu_a(p)e^a$ for every $p \in M$, where $\{e^a\}$ is the canonically dual basis for $\mathfrak{g}^*$. In other words, $\mu(p)(e_a) = \mu_a(p)$.

Exercise 3.28 (The Poisson bracket)

Prove that the Poisson bracket defined by:

$$\{f, g\} = \omega^{ij}\partial_i f \partial_j g$$

where $\omega^{ij}\omega_{jk} = \delta_k{}^i$, is antisymmetric and obeys the Jacobi identity. Using the above definition of the Poisson bracket (or otherwise) prove the following identity:

$$d\{\mu_a, \mu_b\} = f_{ab}{}^c d\mu_c \tag{3.35}$$

where $f_{ab}{}^c$ are the structure constants of $\mathfrak{g}$ in the chosen basis.

(Hint: Prove first that $X_a\mu_b = \{\mu_a, \mu_b\}$ and take d of this expression. You may wish you use the following properties of the contraction: $[\mathcal{L}_a, \imath_b] = f_{ab}{}^c{}_c$.)

Now notice that since μ_a are defined by their gradients, they are defined up to a constant. If these constants can be chosen so that equation (3.35) can be integrated to

$$\{\mu_a, \mu_b\} = f_{ab}{}^c \mu_c \tag{3.36}$$

then the moment mapping μ is equivariant and the action is called Poisson. In other words, the moment mapping intertwines between the action of G on M and the coadjoint action of G on $\mathfrak{g}^*$. Again one can write down precise conditions under which this is the case - conditions which would be met, for example, if $\mathfrak{g}$ were semisimple. We will assume henceforth that the necessary conditions are met and that the moment mapping is equivariant.

The components μ_a of an equivariant moment mapping can be understood as first-class constraints. It is well-known that if the constraints are irreducible, so that their gradients are linearly independent almost everywhere on the constraint submanifold, one can reduce the original symplectic manifold to a smaller symplectic manifold (or, more generally, orbifold). More precisely, the irreducibility condition on the constraints means that their zero locus $\mu^{-1}(0)$ is an embedded submanifold of M. The fact that the constraints are first class means that the vector fields X_a, when restricted to $\mu^{-1}(0)$, are tangent to $\mu^{-1}(0)$.

Equivalently, one may deduce from the equivariance of the moment mapping that G acts on $\mu^{-1}(0)$. Provided that it does so "nicely" (that is, freely and properly discontinuous) the space $\mu^{-1}(0)/G$ of G-orbits is a manifold, and a celebrated theorem of Marsden and Weinstein tells us that it is symplectic. Indeed, if we let $\pi : \mu^{-1}(0) \to \mu^{-1}(0)/G$ denote the natural projection, then the Marsden-Weinstein theorem says that there is a unique symplectic form $\bar{\omega}$ on $\mu^{-1}(0)/G$ such that its pullback $\pi^*\bar{\omega}$ to $\mu^{-1}(0)$ coincides with the restriction to $\mu^{-1}(0)$ of the symplectic form ω on M. The symplectic manifold $(\mu^{-1}(0)/G, \bar{\omega})$ is known as the symplectic quotient of (M, ω) by the action of G. It wouldn't be too difficult to sketch a proof of this theorem, but since we will only need the special case of a Kähler quotient, we will omit it.

9.4.2 3.4.2 Kähler quotients

Now suppose that (M, g, I) is Kähler with Kähler form ω. Then in particular (M, ω) is symplectic. Assume that the action of G on M is not just Poisson, but that G also acts by isometries, that is, preserving g. Because G preserves both g and ω, it also preserves I. On $\mu^{-1}(0)$ we have the induced metric: the restriction to $\mu^{-1}(0)$ of the metric on M. This gives rise to a a metric on $\mu^{-1}(0)/G$ which we will discuss below.

For every $p \in \mu^{-1}(0)$, the tangent space T_pM breaks up as

$$T_pM = T_p\mu^{-1}(0) \oplus N_p\mu^{-1}(0)$$

where $T_p\mu^{-1}(0)$ is the tangent space to $\mu^{-1}(0)$ and the normal space $N_p\mu^{-1}(0)$ is defined as its orthogonal complement $(T_p\mu^{-1}(0))^\perp$. Globally this means that the restriction to $\mu^{-1}(0)$ of the tangent bundle of M decomposes as:

$$TM = T\mu^{-1}(0) \oplus N\mu^{-1}(0) \tag{3.37}$$

where the normal bundle $N\mu^{-1}(0)$ is defined as $(T\mu^{-1}(0))^\perp$. As the next exercise shows, the normal bundle is trivial because $\mu^{-1}(0)$ is defined globally by irreducible constraints.

Exercise 3.29 (Triviality of the normal bundle)

Prove that the normal space $N_p\mu^{-1}(0)$ is spanned by the gradients $\operatorname{grad}_p \mu_a$ of the constraints; or globally, that the gradients of the constraints $\{\operatorname{grad} \mu_a\}$ trivialise the normal bundle.

In fact, the converse is also true. If you feel up to it, prove that the normal bundle to a submanifold is trivial if and only if the submanifold can be described globally as the zero locus of some irreducible "constraints."

(Hint: A vector field X is tangent to $\mu^{-1}(0)$ if and only if it preserves the constraints: $d\mu_a(X) = 0$, but this is precisely $g(\operatorname{grad} \mu_a, X) = 0$, by definition of $\operatorname{grad} \mu_a$.)

Both the metric and the symplectic form restrict to $\mu^{-1}(0)$, but whereas g is nondegenerate on $\mu^{-1}(0)$, the symplectic form isn't. Thus in order to obtain a Kähler manifold it is necessary to perform a quotient. We will describe this quotient locally. To this effect, let us split the tangent space $T_p\mu^{-1}(0)$ further as:

$$T_p\mu^{-1}(0) = H_p \oplus V_p$$

where the vertical vectors V_p are those vectors tangent to the G-orbits and the horizontal vectors $H_p = V_p^\perp$ are their orthogonal complement. The vertical subspace is spanned by the Killing vectors X_a, whereas the horizontal space H_p can be identified with the tangent space to $\mu^{-1}(0)/G$ at the G-orbit of p. Indeed, given any vector field X on $\mu^{-1}(0)/G$ we define its horizontal lift to be the unique horizontal vector field $\tilde{X}$ on $\mu^{-1}(0)$ which projects down to $X : \pi_*\tilde{X} = X$.

Now given any two vector fields on $\mu^{-1}(0)/G$, we define their inner product to be the inner product of their horizontal lifts. This is independent on the point in the orbit to where we lift, because the metric is constant on the orbits. Hence it is a well-defined metric on $\mu^{-1}(0)/G$. In fancier language, this is the unique metric on $\mu^{-1}(0)/G$ which makes the projection π a riemannian submersion. (The reader will surely recognise this construction as the one which in section 3.1.2 yielded the metric on the physical configuration space $\mathcal{C}$ of the Yang-Mills-Higgs system.)

We claim that there is also a symplectic form on $\mu^{-1}(0)/G$ which makes this metric Kähler. We prefer to work with the complex structure.

By definition, if Y is any vector field tangent to M, its inner product with $\operatorname{grad} \mu_a$ is given by

$$g(\operatorname{grad} \mu_a, Y) = d\mu_a(Y) = \omega(X_a, Y) = g(IX_a, Y)$$

whence $\operatorname{grad} \mu_a = IX_a$. Hence if we decompose the restriction of TM to $\mu^{-1}(0)$ as

$$TM = T\mu^{-1}(0) \oplus N\mu^{-1}(0) = H \oplus V \oplus N\mu^{-1}(0)$$

and we choose as bases for V and $N\mu^{-1}(0)$, $\{X_a\}$ and $\{\operatorname{grad}\mu_a\}$ respectively, the complex structure I has the following form:

$$I = \begin{pmatrix} \bar{I} & 0 & 0 \\ 0 & 0 & \mathbb{1} \\ 0 & -\mathbb{1} & 0 \end{pmatrix}$$

whence H is a complex subspace relative to the restriction $\bar{I}$ of I. In other words, the complex structure commutes with the horizontal projection, or a
little bit more precisely, if Y is a vector field on $\mu^{-1}(0)/G$ and $\tilde{Y}$ its horizontal lift, then $I\tilde{Y} = \widetilde{\bar{I}Y}$. The next exercise asks you to prove that this complex structure is integrable, whence $\mu^{-1}(0)/G$ is a complex manifold.

Exercise 3.30 (Integrability of the restricted complex structure)

Use the Newlander-Nirenberg theorem to deduce that $\bar{I}$ is integrable.

(Hint: Relate the Nijenhuis tensor $N_{\bar{I}}$ of $\bar{I}$ to that of I, which vanishes since I is integrable.)

To prove that $\bar{I}$ is parallel, we need to know how the Levi-Civita connection of $\mu^{-1}(0)/G$ is related to the one on M. The next exercise asks you to prove the relevant relation.

Exercise 3.31 (O'Neill's formula)

Let X and Y be vector fields on $\mu^{-1}(0)/G$, and let $\tilde{X}$ and $\tilde{Y}$ be their horizontal lifts. Prove the following formula

$$\nabla_{\tilde{X}}\tilde{Y} = \widetilde{\bar{\nabla}_X Y} + \frac{1}{2}[\tilde{X},\tilde{Y}]^v \tag{3.38}$$

where $\bar{\nabla}$ is the Levi-Civita connection on $\mu^{-1}(0)/G$, and ^{v} denotes the projection onto the vertical subspace. In other words, the horizontal projection of $\nabla_{\tilde{X}}\tilde{Y}$ is precisely the horizontal lift of $\bar{\nabla}_X Y$.

(Hint: Use expressions (3.31) and (3.29) to evaluate the horizontal and vertical components of $\nabla_{\tilde{X}}\tilde{Y}$.)

In other words, formula (3.38) says that if we identify H with the tangent space to $\mu^{-1}(0)/G$, then the Levi-Civita connection on $\mu^{-1}(0)/G$ is given simply by the horizontal projection of the Levi-Civita connection on M. Or, said differently, that the covariant derivative commutes with the horizontal projection. Since the complex structures also commute with the projection, we see that $\nabla I = 0$ on M implies that $\bar{\nabla}\bar{I} = 0$ on $\mu^{-1}(0)/G$. Therefore using Exercise 3.20, $\mu^{-1}(0)/G$ is Kähler.

Notice that the above decomposition (3.37) can be thought of as

$$TM \cong T\left(\mu^{-1}(0)/G\right) \oplus \mathfrak{g}^{\mathbb{C}}$$

where $\mathfrak{g}^{\mathbb{C}}$ is the complexification of the Lie algebra of G. Therefore, morally speaking, it would seem that $\mu^{-1}(0)/G$ is the quotient of M by the action of $G^{\mathbb{C}}$. In some circumstances this is actually an accurate description of the Kähler quotient; for instance, the construction of complex projective space $\mathbb{CP}^n$ as a Kähler quotient of $\mathbb{C}^{n+1}$.

9.4.3 3.4.3 Hyperkähler quotients

Now let (M, g, I, J, K) be a hyperkähler manifold. We have three Kähler forms: $\omega^{(I)}, \omega^{(J)}$, and $\omega^{(K)}$. Suppose that G acts on M via isometries and preserving the three complex structures, hence the three Kähler forms. Assume moreover that the action of G gives rise to three equivariant moment mappings: $\mu^{(I)}, \mu^{(J)}$ and $\mu^{(K)}$; which we can combine into a single map

$$\mu : M \to \mathfrak{g}^* \otimes \mathbb{R}^3$$

Equivariance implies that $\mu^{-1}(0)$ is acted on by G. Assuming that $\mu^{-1}(0)/G$ is a manifold, we claim that it is actually hyperkähler.

Fix one of the complex structures, I, say; and consider the function

$$\nu = \mu^{(J)} + i\mu^{(K)} : M \to \mathfrak{g}^* \otimes \mathbb{C}$$

For each Killing vector field X_a and any vector field Y,

$$\begin{aligned} d\nu_a(Y) &= \omega^{(J)}(X_a, Y) + i\omega^{(K)}(X_a, Y) = g(JX_a, Y) + ig(KX_a, Y) \\ d\nu_a(IY) &= g(JX_a, IY) + ig(KX_a, IY) = -g(KX, Y) + ig(JX, Y) \end{aligned}$$

whence

$$d\nu_a(IY) = i d\nu_a(Y)$$

or in other words, $\bar{\partial}\nu_a = 0$, so that ν is a holomorphic function (relative to I). This means that $\nu^{-1}(0)$ is a complex submanifold of a Kähler manifold and hence its induced metric is Kähler. Now G acts on $\nu^{-1}(0)$ in such a way that it preserves the Kähler structure, and the resulting moment mapping is clearly the restriction of $\mu^{(I)}$ to $\nu^{-1}(0)$. We may therefore perform the Kähler quotient of $\nu^{-1}(0)$ by the action of G, and obtain a manifold:

$$\nu^{-1}(0) \cap \left(\mu^{(I)}\right)^{-1}(0)/G = \mu^{-1}(0)/G$$

whose metric is Kähler relative to (the complex structure induced by) I. To finish the proof that $\mu^{-1}(0)/G$ is hyperkähler, we repeat the above for J and K. This construction is called the hyperkähler quotient, and was described for the first time in [HKLR87.

9.4.4 3.4.4 $\mathcal{M}_k$ as a hyperkähler quotient

Now we will prove that $\mathcal{N}_k$ is a hyperkähler quotient of the configuration space $\mathcal{A}_k$ of fields $\mathrm{W}_{\underline{i}}$ corresponding to finite-energy configurations with monopole number k. We can think of $\mathrm{W}_{\underline{i}}$ as maps $\mathbb{R}^3 \to \mathbb{R}^4 \otimes$ so(3), and $\mathbb{R}^4$
can be thought of as a quaternionic vector space in two inequivalent ways: we first identify $\mathbb{R}^4 = \mathbb{H}$, but then we have to choose whether $\mathbb{H}$ acts by left or right multiplication. Since quaternionic multiplication is not commutative, the two actions are different. Since we will be dealing with monopoles, we choose the right action-left multiplication would correspond to antimonopoles. Let I, J, and K denote the linear maps $\mathbb{R}^4 \to \mathbb{R}^4$ representing right multiplication on $\mathbb{H} \cong \mathbb{R}^4$ by the conjugate quaternion units $-i, -j$, and $-k$ respectively. The next exercise asks you to work out the explicit expressions for I, J, and K relative to a chosen basis.

Exercise 3.32 (Hyperkähler structure of $\mathbb{R}^4$)

Choose a basis $\{1, i, j, k\}$ for $\mathbb{H}$. Then relative to this basis, prove that the linear maps I, J, and K are given by the matrices:

$$I = \begin{pmatrix} i\sigma^2 & 0 \\ 0 & -i\sigma^2 \end{pmatrix} \quad J = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix} \quad K = IJ = \begin{pmatrix} 0 & i\sigma^2 \\ i\sigma^2 & 0 \end{pmatrix}$$

Notice that together with the euclidean metric on $\mathbb{R}^4, I, J, K$ make $\mathbb{R}^4$ into a (linear) hyperkähler manifold.

(Hint: Remember that the matrix associated to a linear transformation is defined by $Ie_{\underline{i}} = e_{\underline{j}}I_{\underline{j}\underline{i}}$. This choice makes composition of linear transformations correspond with matrix multiplication.)

We may now define a hyperkähler structure on $\mathcal{A}_k$ as follows. If $\dot{\mathrm{W}}_{\underline{i}}$ is a vector field on $\mathcal{A}_k$, then we define

$$\left(\hat{I}\dot{\mathrm{W}}_{\underline{i}}(x) = I_{\underline{i}\underline{j}}\dot{\mathrm{W}}_{\underline{j}}(x),\right.$$

and similarly for $\hat{J}$ and $\hat{K}$. Clearly they obey the quaternion algebra $\hat{I}\hat{J} = \hat{K}$, etc because I, J and K do. Moreover since they are constant (and so is the metric) they are certainly parallel relative to the Levi-Civita connection on $\mathcal{A}_k$ with the metric given by the Yang-Mills-Higgs functional. Hence $\mathcal{A}_k$ is an infinite-dimensional (affine) hyperkähler manifold.

Let $\mathcal{G}$ denote the group of finite-range time- and x^4-independent gauge transformations. Since the metric is gauge invariant, $\mathcal{G}$ acts on $\mathcal{A}_k$ via isometries. We also claim that $\mathcal{G}$ preserves the three complex structures and gives rise to three equivariant moment mappings. In fact, we will prove this in one go by constructing the moment mappings from the start.

The Killing vectors of the $\mathcal{G}$ action are just the infinitesimal gauge transformations and they are parametrised by square-integrable functions $\boldsymbol{\epsilon} : \mathbb{R}^3 \to so(3)$. The resulting Killing vector field is $X_{\boldsymbol{\epsilon}} \equiv \delta_{\boldsymbol{\epsilon}}\mathrm{W}_i = D_i\boldsymbol{\epsilon}$. For every such $\boldsymbol{\epsilon}$, define the following function:

$$\mu_{\boldsymbol{\epsilon}}^{(\hat{I})} = \frac{1}{2}\int d^3x I_{\underline{i}\underline{j}}\left(G_{\underline{i}\underline{j}}\cdot\boldsymbol{\epsilon}\right)$$

and the same for $\hat{J}$ and $\hat{K}$.

Now let $\dot{\mathrm{W}}$ be any tangent vector field on $\mathcal{A}_k$. (Here and in the sequel we will suppress the indices $\underline{i}$ whenever they don't play a role in an expression.) Then

$$\begin{aligned}\left(\imath_\epsilon\omega^{(\hat{I})}\right)(\dot{\mathrm{W}}) &= \omega^{(\hat{I})}(D\boldsymbol{\epsilon},\dot{\mathrm{W}})\\ &= g(\hat{I}D\epsilon,\dot{\mathrm{W}})\\ &= \int d^3x I_{\underline{i}\underline{j}}D_{\underline{j}}\boldsymbol{\epsilon}\cdot\dot{\mathrm{W}}_{\underline{i}}\\ &= \int d^3x I_{\underline{i}\underline{j}}D_{\underline{i}}\dot{\mathrm{W}}_{\underline{j}}\cdot\boldsymbol{\epsilon} \qquad \text{(integrating by parts)}\\ &= d\mu_\epsilon^{(\hat{I})}(\dot{\mathrm{W}}).\end{aligned}$$

Hence,

$$\imath_\epsilon\omega^{(\hat{I})} = d\mu_\epsilon^{(\hat{I})}$$

Naturally, the same holds also for $\hat{J}$ and $\hat{K}$. Hence we can construct a moment mapping μ such that $\mu_\epsilon = \left(\mu_\epsilon^{(\hat{I})}, \mu_\epsilon^{(\hat{J})}, \mu_\epsilon^{(\hat{K})}\right)$. The next exercise asks you to prove that it is equivariant.

Exercise 3.33 (Equivariance of the moment mapping)

Prove that the moment mapping $\mu = \left(\mu^{(I)}, \mu^{(J)}, \mu^{(K)}\right)$ is equivariant. In other words, if $\boldsymbol{\epsilon}$ and $\boldsymbol{\eta}$ are gauge parameters, prove that

$$X_\epsilon\mu_\eta^{(I)} = d\mu_\eta^{(I)}\left(X_\epsilon\right) = \mu_{\epsilon\times\eta}^{(I)}$$

and the same for J and K.

Therefore we can apply the preceding discussion about the hyperkähler quotient to deduce that $\mu^{-1}(0)/\mathcal{G}$ is a hyperkähler manifold. But what is $\mu^{-1}(0)$? Configurations $\mathrm{W}_{\underline{i}}$ belonging in

$\mu^{-1}(0)$ are those for which $\mu_{\epsilon}^{(\hat{I})} = 0$ for all $\boldsymbol{\epsilon}$, and the same for $\hat{J}$ and $\hat{K}$. Since $\boldsymbol{\epsilon}$ is arbitrary, this is equivalent to demanding that $I_{ij}G_{ij} = 0$, and the same for $\hat{J}$ and $\hat{K}$. From the explicit expressions for the matrices I, J, and K found in Exercise 3.32, we find that

$$\begin{aligned} I_{ij}G_{ij} &= 0 \Rightarrow G_{12} = G_{34} \\ J_{ij}G_{ij} &= 0 \Rightarrow G_{13} = -G_{24}, \\ K_{ij}G_{ij} &= 0 \Rightarrow G_{14} = G_{23} \end{aligned}$$

But these make up precisely the self-duality condition on G_{ij}, that is, the Bogomol'nyi equation!

Therefore $\mu^{-1}(0)$ is the submanifold of static solutions of the Bogomol'nyi equation with monopole number k (the BPS- k-monopoles) and $\mu^{-1}(0)/\mathcal{G}$ is their moduli space $\mathcal{M}_k$. In summary, $\mathcal{M}_k$ is a $4k$-dimensional hyperkähler manifold, obtained as an infinite-dimensional hyperkähler quotient of $\mathcal{A}_k$ by the action of the gauge group $\mathcal{G}$.

This "proof", although conceptually clear and offering a natural explanation of why $\mathcal{M}_k$ should be a hyperkähler manifold in the first place, relies rather heavily on differential geometry. Therefore a more pedestrian proof might be helpful, and we now turn to one such proof.

9.4.5 3.4.5 Another proof that $\mathcal{M}_k$ is hyperkähler

We start by expanding the Yang-Mills-Higgs action in terms of collective coordinates in order to obtain an expression for the metric. Let X^a, $a = 1, \ldots, 4k$, denote the collective coordinates on the moduli space $\mathcal{M}_k$ of BPSk-monopoles. Let $\mathrm{W}_i(x, X(t))$ be a family of BPS-monopoles whose t-dependence is only through the t-dependence of the collective coordinates; that is,

$$\dot{\mathrm{W}}_{\underline{i}} = \partial_a \, \mathrm{W}_{\underline{i}} \dot{X}^a \tag{3.39}$$

Notice that $\partial_a \, \mathrm{W}_{\underline{i}}$ need not be perpendicular to the gauge orbits. Indeed, generically, we will have a decomposition

$$\partial_a \, \mathrm{W}_{\underline{i}} = \delta_a \, \mathrm{W}_{\underline{i}} + D_{\underline{i}} \epsilon_a \tag{3.40}$$

where $\delta_a \, \mathrm{W}_{\underline{i}}$ is the component perpendicular to the gauge orbits and $D_{\underline{i}}\epsilon_a$ is the component tangent to the gauge orbits, hence an infinitesimal gaugetransformation. The gauge parameters $\boldsymbol{\epsilon}_a$ are determined uniquely by $\partial_a \, \mathrm{W}_i$. Indeed, simply apply $D_{\underline{i}}$ and use the fact that $D^2 \equiv D_{\underline{i}}D_{\underline{i}}$ is negative-definite (hence invertible) to solve for $\boldsymbol{\epsilon}_a$:

$$\boldsymbol{\epsilon}_a = D^{-2} D_{\underline{i}} \partial_a \, \mathrm{W}_{\underline{i}}.$$

Exercise 3.34 (The Yang-Mills-Higgs effective action)

Compute the effective action for such a configuration of BPS-monopoles, and show that provided one sets $\mathrm{W}_0 = \dot{X}^a \epsilon_a$, it is given by

$$L_{\text{eff}} = \frac{1}{2} g_{ab} \dot{X}^a \dot{X}^b - \frac{4\pi a}{e} |k|$$

where the metric on $\mathcal{M}_k$ is given by

$$g_{ab} = \int d^3x \delta_a \, \mathrm{W}_{\underline{i}} \cdot \delta_b \, \mathrm{W}_{\underline{i}}, \tag{3.41}$$

where, by construction, $\delta_a \, \mathrm{W}_i$ are perpendicular to the gauge orbits and satisfy the linearised Bogomol'nyi equation.

One way to systematise the above expansion is in terms of t-derivatives. The zeroth order term is given by the potential, which is a constant since the motion is purely along the flat directions. The first order term vanishes due to our choice for W_0, [$] while the quadratic term, which describes the motion of such monopoles in the limit of zero velocity, corresponds precisely to geodesic motion on $\mathcal{N}_k$ relative to the induced metric - that is, as a particle moving freely on $\mathcal{M}_k$ or, somewhat pedantically, as a $(1+0)$-dimensional σ-model with $\mathcal{M}_k$ as its target space.

It is convenient to think of $\boldsymbol{\epsilon}_a$ as the components of a connection. We define $D_a \equiv \partial_a - e\left(\boldsymbol{\epsilon}_a \times -\right)$, whence we can think of $(\mathrm{W}_{\underline{i}}, \boldsymbol{\epsilon}_a)$ as the components of a connection on $\mathbb{R}^4 \times \mathcal{M}_k$. This allows us to interpret $\delta_a \mathrm{W}_{\underline{i}}$ as the mixed components of the curvature:

$$G_{a\underline{i}} = \partial_a \mathrm{W}_{\underline{i}} - \partial_{\underline{i}}\epsilon_a - e\epsilon_a \times \mathrm{W}_{\underline{i}} = \partial_a \mathrm{W}_{\underline{i}} - D_{\underline{i}}\epsilon_a = \delta_a \mathrm{W}_{\underline{i}}$$

The other components G_{ab} of the curvature may be formally computed from the Bianchi identity:

$$D_{\underline{i}} G_{ab} = -2D_{[a}\delta_{b]}\mathrm{W}_{\underline{i}} = -D_a\delta_b \mathrm{W}_{\underline{i}} + D_b\delta_a \mathrm{W}_{\underline{i}} \tag{3.42}$$

by applying D_i and inverting D^2 as before.

Exercise 3.35 (A somewhat more explicit formula for G_{ab})

Prove that

$$G_{ab} = -2eD^{-2}\left(\delta_a \mathrm{W}_{\underline{i}} \times \delta_b \mathrm{W}_{\underline{i}}\right)$$

(Hint: Apply $D_{\underline{i}}$ to (3.42), and use that $\delta_a \mathrm{W}_{\underline{i}}$ is perpendicular to the gauge orbits.)

Using these formulae it is possible to write a formal expression for the Christoffel symbols of the Levi-Civita connection. Naturally this is left as an exercise.

Exercise 3.36 (The Christoffel symbols)

Prove that

$$\Gamma_{abc} = g_{cd}\Gamma_{ab}{}^{d} = \int d^3x D_a\delta_b \mathrm{W}_{\underline{i}} \cdot \delta_c \mathrm{W}_{\underline{i}} \tag{3.43}$$

Notice that $\Gamma_{abc} = \Gamma_{\mathrm{bac}}$, since $D_{[b}\delta_{c]}\mathrm{W}_{\underline{i}} = -\frac{1}{2}D_{\underline{i}}G_{ab}$ which is orthogonal to $\delta_c \mathrm{W}_{\underline{i}}$.

(Hint: Use the explicit expressions (3.30) and (3.41) and compute.)

Using the explicit expressions found in Exercise 3.32 for the hyperkähler structure in $\mathbb{R}^4$ we now define the following two-forms on $\mathcal{M}_k$:

$$\omega^{(I)}_{ab} = \int d^3x I_{\underline{i}\underline{j}}\delta_a \mathrm{W}_{\underline{i}} \cdot \delta_b \mathrm{W}_{\underline{j}}$$

and similarly for J and K, and their corresponding almost complex structures

$$I_a{}^b = g^{bc}\omega^{(I)}_{ac} \quad J_a{}^b = g^{bc}\omega^{(J)}_{ac} \quad K_a{}^b = g^{bc}\omega^{(K)}_{ac}$$

[*4] Notice that W_0 is not zero for generic choices of $\mathrm{W}_{\underline{i}}(x, X(t))$, but it can be made to vanish after a t-dependent gauge transformation.

Exercise 3.37 (Explicit expressions for the complex structures)

(a) Prove that $I_{ij}\delta_a \, \mathrm{W}_{\underline{j}}$ is orthogonal to the gauge orbits, and the same for J and K.

(b) Then derive the following formula:

$$I_a{}^b \delta_b \, \mathrm{W}_{\underline{i}} = -I_{\underline{ij}}\delta_a \, \mathrm{W}_{\underline{j}} \tag{3.44}$$

and the same for $J_a{}^b$ and $K_a{}^b$.

(c) Using these expressions, prove that $I_a{}^b$, $J_a{}^b$ and $K_a{}^b$ obey the quaternion algebra.

(Hints: (a) This is equivalent to the linearised Bogomol'nyi equation, in the form $I_{ij}G_{ij} = 0$, etc.

(b) Argue that since $I_a{}^b\delta_b \, \mathrm{W}_{\underline{i}}$ is orthogonal to the gauge orbits, $\int d^3x I_a{}^b \delta_b \, \mathrm{W}_{\underline{i}} \cdot \delta_c \, \mathrm{W}_{\underline{i}}$ defines it uniquely. Then just compute the integral and use (a).)

We claim that the forms $\omega^{(I)}, \omega^{(J)}$ and $\omega^{(K)}$ are parallel. Let's see this for one of them $\omega \equiv \omega^{(I)}$, the other cases being identical. By definition,

$$\nabla_a \omega_{bc} = \partial_a \omega_{bc} - \Gamma_{ab}{}^d \omega_{dc} - \Gamma_{ac}{}^d \omega_{bd}$$

We now compute this in steps. First of all we have:

$$\partial_a \omega_{bc} = \int d^3x I_{\underline{ij}} \left(D_a \delta_b \, \mathrm{W}_{\underline{i}} \cdot \delta_c \, \mathrm{W}_{\underline{j}} + \delta_b \, \mathrm{W}_{\underline{i}} \cdot D_a \delta_c \, \mathrm{W}_{\underline{j}} \right) \tag{3.45}$$

Now we notice that $\Gamma_{ab}{}^d \omega_{dc} = -I_c{}^e \Gamma_{abe}$. Using the explicit expression (3.43), we arrive at

$$\Gamma_{ab}{}^d \omega_{dc} = -I_c{}^e \int d^3x D_a \delta_b \, \mathrm{W}_{\underline{i}} \cdot \delta_e \, \mathrm{W}_{\underline{i}} \tag{3.46}$$

and using (3.44) we can rewrite this as

$$\Gamma^d_{ab} \omega_{dc} = \int d^3x I_{\underline{ij}} D_a \delta_b \mathbf{W}_{\underline{i}} \cdot \delta_c \mathbf{W}_{\underline{j}} \tag{3.47}$$

with a similar expression for $\Gamma_{ac}{}^d \omega_{bd} = -\Gamma_{ac}{}^d \omega_{db}$. Adding it all together we find that $\nabla_a \omega_{bc} = 0$. But this means that $I_a{}^b$, $J_a{}^b$, and $K_a{}^b$ are also parallel, whence $\mathcal{M}_k$ is hyperkähler.

Chapter 4

10 The Effective Action for $N = 2$ Supersymmetric Yang-Mills

In this chapter we will perform the collective coordinate expansion of the $N = 2$ supersymmetric $SO(3)$ Yang-Mills theory defined by equation (2.13). We will also discuss the quantisation of the corresponding effective action. As we saw in the discussion in section 3.1.5 on the effective theory for the 1monopole sector, the effective theory offers a qualitatively faithful description of the dyonic spectrum, even though quantitatively it is only an approximation. Of course, in the non-supersymmetric theory there is no reason to expect that the true quantum spectrum should resemble the classical spectrum given by the Bogomol'nyi formula, but as we saw in Chapter 2, supersymmetry protects both the formula for the bound from quantum corrections and also the saturation of the bound. Hence it makes sense to expect that in the supersymmetric theory, the quantisation of the effective action should teach us something about the full quantum theory. As we shall soon discuss, this will have a chance of holding true only in the $N = 4$ theory, but we can already learn something from the $N = 2$ theory we have just studied. We will therefore first discuss the fermionic collective coordinates and then the effective quantum theory in the

k-monopole sector. We will see that there are $2k$ fermionic collective coordinates and that the resulting effective theory is to lowest order a $(0+1)$ supersymmetric σ-model admitting $N=4$ supersymmetry consistent with the fact that $\mathcal{M}_k$ is hyperkähler. The quantisation of this theory then leads to a geometric interpretation of the Hilbert space as the square-integrable $(0,q)$-forms on $\mathcal{N}_k$, and of the hamiltonian as the laplacian. This chapter is based on the work of Gauntlett [Gau94].

10.1 4.1 Fermionic collective coordinates

As we saw in section 3.2, there are $4k$ bosonic collective coordinates in the k-monopole sector. The purpose of this section is to compute the number of fermionic collective coordinates: we will see that there are $2k$ of them. We will prove this in two ways. First we can set up this problem as the computation of the index of an operator, as we did for the bosonic collective coordinates; then essentially the same calculation that was done in section 3.2 yields the answer. Alternatively, and following Zumino [Zum77], we will exhibit a supersymmetry between the bosonic and fermionic collective coordinates which will also allow us to count them.

Suppose that we start with an $N=2$ BPS-monopole obtained, say, in the manner of Exercise 2.17. Fermionic collective coordinates are simply fermionic flat directions of the potential; that is, fermionic configurations which do not change the potential. In order to see what this means, let us first write down the potential for a general field configuration. To this effect let us break up the lagrangian density (2.13) into kinetic minus potential terms $\mathcal{L}=\mathcal{T}-\mathcal{V}$, where

$$\mathcal{T}=\frac{1}{2}\left\|\mathrm{G}_{0i}\right\|^2+\frac{1}{2}\left\|D_0\mathrm{P}\right\|^2+\frac{1}{2}\left\|D_0\ \mathrm{S}\right\|^2+i\overline{\boldsymbol{\psi}}\cdot\gamma^0 D_0\boldsymbol{\psi} \tag{4.1}$$

and

$$\begin{aligned}\mathcal{V}=&\frac{1}{2}\left\|D_i\mathrm{P}\right\|^2+\frac{1}{2}\left\|D_i\ \mathrm{S}\right\|^2+\frac{1}{4}\mathrm{G}_{ij}\cdot\mathrm{G}_{ij}+\frac{1}{2}e^2\|\mathrm{P}\times\mathrm{S}\|^2+i\overline{\boldsymbol{\psi}}\cdot\gamma_i D_i\boldsymbol{\psi}\\ &-ie\bar{\psi}\cdot\gamma_5\mathrm{P}\times\boldsymbol{\psi}-ie\bar{\psi}\cdot\mathrm{S}\times\boldsymbol{\psi}\end{aligned} \tag{4.2}$$

The potential is then the integral $V=\int_{\mathbb{R}^3}\mathcal{V}$. For an $N=2$ BPS-monopole, with $\mathrm{W}_0=0,\ \mathrm{S}=\alpha\boldsymbol{\phi},\mathrm{P}=\beta\boldsymbol{\phi}$ with $\alpha^2+\beta^2=1$, the potential is given by

$$V=\frac{1}{2}\int_{\mathbb{R}^3}\left\|D_i\boldsymbol{\phi}\right\|^2+\frac{1}{4}\int_{\mathbb{R}^3}\mathrm{G}_{ij}\cdot\mathrm{G}_{ij}+i\int_{\mathbb{R}^3}\overline{\boldsymbol{\psi}}\cdot\left(\gamma_i D_i-e\boldsymbol{\phi}\left(\alpha+\beta\gamma_5\right)\right)\boldsymbol{\psi}$$

where in the last term ϕ is in the adjoint representation; that is, $\phi\psi=\phi\times\psi$. The first two terms in the potential already reproduce the potential energy of a nonsupersymmetric BPS-monopole: $\frac{4\pi a}{e}|k|$, for a k-monopole. Therefore turning on the fermions will not change the potential provided that the third term vanishes; in other words, provided that $\boldsymbol{\psi}$ satisfies the Dirac equation in the presence of the BPS-monopole. In other words, fermionic collective coordinates are in one-to-one correspondence with zero modes of the Dirac operator. We will now count the number of zero modes in two ways.

10.1.1 4.1.1 Computing the index

In order to count the zero modes it is again convenient to use the reformulation of the BPS-monopole as an instanton, in terms of $\mathrm{W}_{\underline{i}}=(\mathrm{W}_i,\phi)$, and to define the following four-dimensional euclidean Dirac matrices: $\bar{\gamma}_i=\gamma_0\gamma_i$ and $\bar{\gamma}_4=\gamma_0\left(\alpha+\beta\gamma_5\right)$. In terms of these, the fermion term in the potential is given by $i\int_{\mathbb{R}^3}\boldsymbol{\psi}^{\dagger}\cdot\bar{\gamma}_{\underline{i}}D_{\underline{i}}\psi$, keeping in mind that $\partial_4\equiv 0$.

We want to compute the number of normalisable solutions to the equation $\bar{\gamma}_{\underline{i}}D_{\underline{i}}\boldsymbol{\psi}=0$. Let us choose a Weyl basis in which $\bar{\gamma}_5=\bar{\gamma}_1\bar{\gamma}_2\bar{\gamma}_3\bar{\gamma}_4$ is diagonal. In such a basis, a convenient representation of the euclidean Dirac matrices is the one given in equation (3.19). In that representation the euclidean Dirac equation becomes:

$$\begin{pmatrix} 0 & -i\sigma_i D_i - e\boldsymbol{\phi}\mathbb{1} \\ i\sigma_i D_i - e\mathbf{1}\mathbb{1} & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\psi}_+ \\ \boldsymbol{\psi}_- \end{pmatrix} = 0$$

But notice that we have seen these very operators before, in the computation of the number of bosonic collective coordinates in section 3.2. In fact, in terms of the operator $\mathcal{D}$ defined in equation (3.17), the above Dirac equation breaks up into two equations, one for each chirality:

$$\mathcal{D}\boldsymbol{\psi}_- = 0 \quad \text{and} \quad \mathcal{D}^\dagger\boldsymbol{\psi}_+ = 0$$

But now in Exercise 3.5 you showed that the operator $\mathcal{D}\mathcal{D}^\dagger$ is positive, whence it has no normalisable zero modes, hence neither does $\mathcal{D}^\dagger$. Therefore we notice that fermionic zero modes in the presence of a BPS-monopole necessarily have negative chirality. (For antimonopoles, it would have been $\mathcal{D}$ which has no normalisable zero modes, and fermion zero modes would have positive chirality.)

We can arrive at the same result in a different way which doesn't use the explicit realisation of the $\bar{\gamma}_i$-matrices. In fact, it is an intrinsic property of fermions coupled to instantons (in four-dimensions). The next exercise takes you through it.

Exercise 4.1 (Fermion zero modes are chiral)

Consider solutions of the four-dimensional euclidean Dirac equation $\bar{\gamma}_i D_i \psi = 0$ in the presence of an (anti)self-dual gauge field. Prove that if the gauge field is self-dual (respectively, antiselfdual), then fermion zero modes have negative (respectively, positive) chirality.

(Hint: Compute the Dirac laplacian $\bar{\gamma}_i\bar{\gamma}_{\underline{j}} D_{\underline{i}} D_{\underline{j}}\psi$ and use the fact that $D^2 = D_{\underline{i}} D_{\underline{i}}$ is negative-definite and has not normalisable zero modes.)

Finally, just as in section 3.2.1, the number of normalisable zero modes of $\mathcal{D}$ is given by its index, which was computed in section 3.2.3 to be $2k$, where k is the monopole number.

10.1.2 4.1.2 Using supersymmetry

We can reproduce this result in a different, but more useful way by exhibiting a supersymmetry between the bosonic and fermionic zero modes. This is based on work by Zumino [Zum77].

Let $\delta \mathrm{W}_{\underline{i}}$ be a bosonic zero mode; that is, $\delta \mathrm{W}_{\underline{i}}$ satisfies the linearised Bogomol'nyi equation (3.15) and Gauss's law (3.1). Let η_+be a constant, commuting spinor of positive chirality, normalised to $\eta_+^\dagger \eta_+ = 1$. Define

$$\psi \equiv \delta \mathrm{W}_{\underline{i}} \bar{\gamma}_{\underline{i}} \eta_+ \tag{4.3}$$

It is clear that $\boldsymbol{\psi}$ has negative chirality and, as the next exercise asks you to show, ψ satisfies the Dirac equation.

Exercise 4.2 (From bosonic to fermionic zero modes)

Let $\delta \mathrm{W}_i$ be a bosonic zero mode as above. With $\boldsymbol{\psi}$ defined as above, prove that $\bar{\gamma} \cdot D\psi = 0$.

(Hint: Use Exercise 3.4.)

Conversely, suppose that $\boldsymbol{\psi}$ is a fermionic zero mode with negative chirality; that is, $\bar{\gamma}_5\boldsymbol{\psi} = -\boldsymbol{\psi}$ and $\bar{\gamma} \cdot D\boldsymbol{\psi} = 0$. Then define

$$\delta \mathrm{W}_{\underline{i}} \equiv i\eta_+^\dagger \bar{\gamma}_{\underline{\gamma}} \psi - i\boldsymbol{\psi}^\dagger \bar{\gamma}_{\underline{i}} \eta_+ .$$

The next exercise asks you to prove that $\delta \mathrm{W}_{\underline{i}}$ is a bosonic zero mode.

Exercise 4.3 (... and back)

With δW_i defined as above, prove that it satisfies the linearised Bogomol'nyi equation (3.15) and Gauss's law (3.1).

The above result seems to suggest that there is a one-to-one correspondence between the bosonic and fermionic zero modes, but this is fictitious, since not all the fermionic zero modes obtained in this fashion are independent. Indeed, as we will now see, they are related by the complex structure.

Let $\delta_a W_{\underline{i}}$ for $a = 1, \ldots, 4k$ denote the bosonic zero modes, and let $\boldsymbol{\psi}_a = \delta_a W_{\underline{i}} \bar{\gamma}_i \eta_+$ be the corresponding fermionic zero modes. We will prove that $I_a{}^b \boldsymbol{\psi}_b = i\boldsymbol{\psi}_a$, where I is one of the complex structures of $\mathcal{M}_k$. Hence comparing with the discussion at the end of section 3.2.1, we see that unlike bosonic zero modes, $\boldsymbol{\psi}_a$ and $i\boldsymbol{\psi}_a$ are not linearly independent.

Let η_+ be a commuting spinor of positive chirality normalised to $\eta_+^\dagger \eta_+ = 1$. Define a 4×4 matrix A with entries

$$A_{i\underline{j}} \equiv \eta_+^\dagger \bar{\gamma}_{i\underline{j}} \eta_+ \tag{4.4}$$

We start by listing some properties of this matrix.

Exercise 4.4 (A complex structure)

Let A be the 4×4 matrix with entries A_{ij} given by (4.4). Prove that A satisfies the following properties:

(1) A is antisymmetric;
(2) iA is real;
(3) A is antiselfdual: $A_{\underline{ij}} = -\frac{1}{2}\epsilon_{\underline{ij\ell\ell}} A_{\underline{k\ell}}$;
(4) $A^2 = \mathbb{1}$, so that iA is a complex structure; and
(5) $A_{i\underline{j}} \bar{\gamma}_j \eta_+ = -\bar{\gamma}_{\underline{i}} \eta_+$.
(Hint: This requires the Fierz identity:

$$\eta_+ \eta_+^\dagger = \frac{1}{4}\left(\mathbb{1} + \bar{\gamma}_5\right) - \frac{1}{8} A_{i\underline{j}} \bar{\gamma}_{ij}$$

which you should prove.)

We will now prove that we can choose η_+ in such a way that that iA agrees with any one of the complex structures on $\mathbb{R}^4$ defined in Exercise 3.32, and that such an η_+ is unique up to a phase. We start by noticing that the 4×4 matrix iA defined above is real and antisymmetric, hence it belongs to $so(4)$. As a Lie algebra, so(4) is isomorphic to $so(3) \times so(3)$ (see Exercise 2.32). The fact that iA is antiselfdual, means that iA belongs to one of these so(3)'s. In fact, it is the $so(3)$ spanned by the complex structures I, J, and K of Exercise 3.32. (Check that they are antiselfdual!) In fact, $\mathbb{R}^4$ has a two-sphere worth of complex structures: $\{aI + bJ + cK \mid a^2 + b^2 + c^2 = 1\}$, and from the above exercise, we see that iA defines a point in this two-sphere. In the next exercise we see this explicitly.

Exercise 4.5 (iA lives on the sphere)

Compute the matrix i A explicitly in the representation of the Dirac matrices given by equation (3.19), and show that it is given by

$$iA = \begin{pmatrix} 0 & q_3 & -q_2 & -q_1 \\ -q_3 & 0 & q_1 & -q_2 \\ q_2 & -q_1 & 0 & -q_3 \\ q_1 & q_2 & q_3 & 0 \end{pmatrix} = -q_1 K - q_2 J + q_3 I \tag{4.5}$$

where I, J, and K are the complex structures in $\mathbb{R}^4$ defined in Exercise 3.32 and $q_i = \eta^\dagger \sigma_i \eta$, where η is a complex Weyl spinor normalised to $\eta^\dagger \eta = 1$. (In the Weyl basis above $\eta_+ = \binom{\eta}{0}$.) Prove that q_i are real and that they satisfy $\sum_i q_i^2 = 1$, hence iA defines a point in the unit two-sphere in $\mathbb{R}^3$.

Now, in the Weyl basis introduced above, $\eta_+ = \binom{\eta}{0}$. Any other normalised positive chirality spinor η'_+ will have the same form with η' replacing η. This new Weyl spinor will be related to η by an element of $U(2) : \eta' = U\eta$. The matrix iA' obtained from η' has the form given by (4.5) but with q_i replaced by $q'_i \equiv \eta'^\dagger \sigma_i \eta' = \eta^\dagger U^{-1} \sigma_i U \eta$.

Exercise 4.6 (Adjoint transformation)

In the notation above, prove that $q'_i = U_{ij} q_j$, where U_{ij} is the three-dimensional adjoint representation of $U(2)$. (Notice that because the $U(1)$ subgroup corresponding to the scalar matrices act trivially, only the $SU(2)$ subgroup acts effectively in this representation.)

Therefore the action of $U(2)$ on η induces the adjoint action on the (q_i). This action is transitive on the unit sphere $\sum_i q_i^2 = 1$, hence any two points (q_i) and (q'_i) are related by an element of $U(2)$. Notice that $U(2) = SU(2)\times\ U(1)$ and that the $SU(2)$ subgroup acts freely, whereas the $U(1)$ acts trivially. Hence once a complex structure has been chosen, η is unique up to the action of $U(1)$; that is, a phase.

Let us then exercise our right to choose η_+. We do so in such a way that

$$\eta_+^\dagger \bar{\gamma}_{ij} \eta_+ = i I_{i\underline{j}}$$

where I_{ij} is given by Exercise 3.32. Then using equation (3.44), we have

$$\begin{aligned} I_a{}^b \boldsymbol{\psi}_b &= I_a{}^b \delta_b \ \mathrm{W}_{\underline{i}} \bar{\gamma}_{\underline{i}} \eta_+ = -I_{ij} \delta_a \ \mathrm{W}_{\underline{j}} \bar{\gamma}_{\underline{i}} \eta_+ \\ &= i\delta_a \ \mathrm{W}_{\underline{j}} \bar{\gamma}_j \eta_+ \\ &= i\boldsymbol{\psi}_a \end{aligned} \tag{4.6}$$

where in the next to last line we have used (5) in Exercise4.5. Therefore there are only half as many linearly independent fermionic zero modes as there are bosonic ones, in agreement with the index calculation in the previous section.

10.1.3 4.2 The effective action

In this section we will write down the effective action governing the dynamics of the collective coordinates to lowest order. Let us introduce bosonic collective coordinates X^a, for $a = 1, \ldots, 4k$. These coordinates parametrise the moduli space $\boldsymbol{\mathcal{M}}_k$ of BPS-monopoles with topological number k. In addition there will be fermionic collective coordinates λ^a, for $a = 1, \ldots, 4k$ satisfying the condition $\lambda^a I_a{}^b = i\lambda^b$. We now expand the supersymmetric Yang-Mills action (2.13) in terms of the collective coordinates $\{X^a, \lambda^a\}$ and keep only
the lowest nontrivial order. In order to count the order of an expression we take the conventions that λ^a has order $\frac{1}{2}, X^a$ has order 0, but time derivatives have order 1. These conventions are such that a free theory of bosons X^a and fermions λ^a is of quadratic order.

We start by performing an $SO(2)$ transformation which puts $\mathrm{S} = \phi$ and $\mathrm{P} = 0$, and choosing an appropriate parametrisation for the fields $\mathrm{W}_{\underline{i}}, \mathrm{P}, \mathrm{W}_0$, and $\boldsymbol{\psi}$ in terms of the collective coordinates. As in the nonsupersymmetric theory, we leave W_0 and, in this case also P, undetermined for the moment. We choose to parametrise W_i as $\mathrm{W}_i(x, X(t))$, where all the time dependence comes from the collective coordinates. For this reason, equation (3.39) still holds where, as before, $\partial_a \ \mathrm{W}_i$ need not be perpendicular to the gauge orbits. Nevertheless we can project out a part which does: $\delta_a\, \mathrm{W}_{\underline{i}}$ as in (3.40). Because $\delta_a\, \mathrm{W}_{\underline{i}}$ is perpendicular to the gauge orbits, we know that $\boldsymbol{\psi}_a$ given by (4.3) is a fermion zero mode. We therefore parametrise $\psi = \psi_a \lambda^a$. Notice that since ψ_a is commuting (because so is η_+), λ^a is anticommuting as expected.

There is no reason, in principle, to expect W_i not to depend also on the fermionic collective coordinates. In fact, using that these are odd, we can expand $\mathrm{W}_{\underline{i}}$ as follows:

$$\mathbf{W}_{\underline{i}}(x, X, \lambda) = \mathbf{W}_{\underline{i}}(x, X) + \lambda^a \mathbf{W}_{\underline{i},a}(x, X) + \lambda^a \lambda^b\ \mathrm{W}_{\underline{i},ab}(x, X) + \cdots ;$$

but it is not hard to see that all terms but the first in the above expansion contribute only higher order terms to the effective action.

Because our choice of ψ is a zero mode of the Dirac equation in the presence of a BPS-monopole, the discussion of section 4.1 applies and provided that $\mathrm{P} = 0$, the potential remains at its minimum value: $\frac{4\pi a}{e} k$. However, P need not remain zero for this to be case. It can evolve along a flat direction as we now show. The P-dependent terms in the potential (4.2) can be written as follows:

$$\frac{1}{2} \int_{\mathbb{R}^3} \|D_{\underline{i}} \mathrm{P}\|^2 + ie \int_{\mathbb{R}^3} \boldsymbol{\psi}^\dagger \cdot (\mathrm{P} \times \boldsymbol{\psi})$$

where we have used that $\bar{\gamma}_5 \boldsymbol{\psi} = \gamma_0 \gamma_5 \boldsymbol{\psi} = -\boldsymbol{\psi}$. Integrating the first term by parts, and using the invariance of the inner product in the second term, the above expression becomes:

$$-\frac{1}{2} \int_{\mathbb{R}^3} \mathrm{P} \cdot \left(D^2 \mathrm{P} + 2ie \boldsymbol{\psi}^\dagger \times \boldsymbol{\psi} \right)$$

where $D^2 = D_{\underline{i}} D_{\underline{i}}$.

Exercise 4.7 (Computing $\psi^\dagger \times \psi$)

Prove the following identity:

$$\boldsymbol{\psi}^\dagger \times \boldsymbol{\psi} = -\frac{1}{e} \bar{\lambda}^a \lambda^b D^2 G_{ab}$$

where G_{ab} are some of the components of the curvature of the connection $(\mathrm{W}_{\underline{i}}, \boldsymbol{\epsilon}_a)$ on $\mathbb{R}^4 \times \boldsymbol{\mathcal{M}}_k$, which were computed in Exercise 3.35.

(Hint: You might want to use the identity:

$$\eta_+^\dagger \bar{\gamma}_{\underline{i}} \bar{\gamma}_{\underline{j}} \eta_+ = \delta_{\underline{ij}} + i I_{\underline{ij}} \tag{4.7}$$

which is (up to a factor) the projector onto the I-antiholomorphic subspace of the complexification $\mathbb{C}^4$ of $\mathbb{R}^4$.)

Therefore we see that the condition that the potential remains constant demands that we either set P to zero or else

$$\mathrm{P} = 2i \bar{\lambda}^a \lambda^b G_{ab}$$

Next we tackle the kinetic terms. Notice that either of the choices for P allow us to discard P from the kinetic terms. Indeed, if P is nonzero, then the above expression shows that it is already of order 1 and hence its contribution to the kinetic term (4.1) will be of order higher than quadratic. Having discarded P from the kinetic terms, we remain with

$$\frac{1}{2} \int_{\mathbb{R}^3} \|G_0\|^2 + i \int_{\mathbb{R}^3} \boldsymbol{\psi}^\dagger \cdot D_0 \boldsymbol{\psi}$$

The first term is computed in the following exercise.

Exercise 4.8 (The first kinetic term)

Prove that the first term in the kinetic energy above is given by

$$\frac{1}{2}\int_{\mathbb{R}^3}\left\|\mathrm{G}_{0\underline{e}}\right\|^2 = \frac{1}{2}g_{ab}\dot{X}^a\dot{X}^b + \frac{1}{2}\int_{\mathbb{R}^3}\left\|D_{\underline{i}}\left(\boldsymbol{\epsilon}_a\dot{X}^a - \mathrm{W}_0\right)\right\|^2$$

where g_{ab} was defined in (3.41).

(Hint: Use the fact that $\delta_a\ \mathrm{W}_{\underline{i}}$ is perpendicular to the gauge orbits!)

Finally we come to the second kinetic term. Plugging in the expression for $\boldsymbol{\psi}$ and using equations (4.7), (3.44), and (4.6) we can rewrite the second term as

$$i\int_{\mathbb{R}^3}\boldsymbol{\psi}^\dagger\cdot D_0\boldsymbol{\psi} = 2ig_{ab}\bar{\lambda}^a\dot{\lambda}^b + 2i\bar{\lambda}^a\lambda^b\dot{X}^c\int_{\mathbb{R}^3}\delta_a\ \mathrm{W}_{\underline{i}}\cdot\partial_c\delta_b\ \mathrm{W}_{\underline{i}}$$
$$+ 2ie\bar{\lambda}^a\lambda^b\int_{\mathbb{R}^3}\ \mathrm{W}_0\cdot\left(\delta_a\ \mathrm{W}_{\underline{i}}\times\delta_b\ \mathrm{W}_{\underline{j}}\right)$$

The next exercise finishes off the calculation.

Exercise 4.9 (... and the second kinetic term)

Prove that the second kinetic term can be written as

$$i\int_{\mathbb{R}^3}\psi^\dagger\cdot D_0\psi = 2ig_{ab}\bar{\lambda}^a\left(\dot{\lambda}^b + \Gamma_{cd}{}^b\dot{X}^c\lambda^d\right) - i\bar{\lambda}^a\lambda^b\int_{\mathbb{R}^3}\left(\ \mathrm{W}_0 - \epsilon_c\dot{X}^c\right)\cdot D^2G_{ab}$$

where the Christoffel symbols $\Gamma_{cd}{}^b$ were defined in equation (3.43).

Putting the results of Exercises 4.8 and 4.9, we find that the kinetic terms of the action are (to lowest order) given by:

$$\frac{1}{2}g_{ab}\dot{X}^a\dot{X}^b + 2ig_{ab}\bar{\lambda}^a\dot{\lambda}^b + 2ig_{ab}\Gamma^b_{cd}\bar{\lambda}^a\dot{X}^c\lambda^d$$
$$- i\bar{\lambda}^a\lambda^b\int_{\mathbb{R}^3}\left(\ \mathrm{W}_0 - \boldsymbol{\epsilon}_c\dot{X}^c\right)\cdot D^2G_{ab} + \frac{1}{2}\int_{\mathbb{R}^3}\left\|D_{\underline{i}}\left(\boldsymbol{\epsilon}_a\dot{X}^a - \mathrm{W}_0\right)\right\|^2$$

We see that we can cancel the last two terms provided that we set

$$\mathbf{W}_0 = \boldsymbol{\epsilon}_a\dot{X}^a - 2i\bar{\lambda}^a\lambda^bG_{ab}$$

With this choice, and to lowest order, the effective action then becomes:

$$L_{\text{eff}} = \frac{1}{2}g_{ab}\dot{X}^a\dot{X}^b + 2ig_{ab}\bar{\lambda}^a\left(\dot{\lambda}^b + \Gamma_{cd}{}^b\dot{X}^c\lambda^d\right) - \frac{4\pi a}{e}k \tag{4.8}$$

Ignoring the constant term, this action describes a $(0+1$-dimensional supersymmetric (as we shall see shortly) σ-model with target $\boldsymbol{\mathcal{M}}_k$.

10.2 4.3 $N = 4$ supersymmetry of the effective action

In general, symmetries of the theory under consideration play important roles in the effective action. Broken symmetries give rise to collective coordinates, whereas unbroken symmetries remain symmetries of the effective action. As we have seen in section 2.3.4, $N = 2$ BPS-monopoles preserve one half of the four-dimensional $N = 2$ supersymmetry. This supersymmetry must be present in the effective action. In $0+1$ dimensions, supersymmetry charges are one-component Majorana spinors, hence one supersymmetry charge in four dimensions gives rise to four supersymmetry charges in $0+1$. In this section we will prove that the effective action given by (4.8) admits $N = 4$ supersymmetry. From the proof it follows that the same is true for any supersymmetric σ-model with hyperkähler target manifold-which is, of course, a well-known fact.

We start by discarding the constant term in the action (4.8) and rewriting the remaining terms in terms of complex coordinates adapted to the complex structure I of $\mathcal{M}_k$. To this end we define complex coordinates $\left(Z^\alpha, \bar{Z}^{\bar{\alpha}}\right)$ which diagonalise the complex structure; that is, such that $I_\alpha{}^\beta = i\mathbb{1}_\alpha{}^\beta$ and $I_{\bar{\alpha}}{}^{\bar{\beta}} = -i\mathbb{1}_{\bar{\alpha}}{}^{\bar{\beta}}$. As for the fermions, equation (4.6) implies that $\lambda^a I_a{}^b = i\lambda^b$, whence $\lambda^{\bar{\alpha}} = 0$. Similarly, $\bar{\lambda}^\alpha = 0$. This prompts us to define new fermions ζ such that $\zeta^\alpha = \sqrt{2}\lambda^\alpha$ and $\zeta^{\bar{\alpha}} = \sqrt{2}\bar{\lambda}^{\bar{\alpha}}$. In terms of these new variables, the effective action remains

$$L_{\text{eff}} = g_{\alpha\bar{\beta}}\dot{Z}^\alpha\dot{\bar{Z}}^{\bar{\beta}} + ig_{\bar{\alpha}\beta}\zeta^{\bar{\alpha}}\left(\dot{\zeta}^\beta + \Gamma_{\gamma\delta}{}^\beta\dot{Z}^\gamma\zeta^\delta\right) \tag{4.9}$$

where we have used that for a Kähler metric in complex coordinates, the only nonzero components of the metric and the Christoffel symbols are $g_{\alpha\bar{\beta}} = g_{\bar{\alpha}\beta}$ and $\Gamma_{\alpha\beta}{}^\gamma$ and $\Gamma_{\bar{\alpha}\bar{\beta}}\bar{\gamma}$, as was proven in Exercise 3.23.

10.2.1 4.3.1 $N = 4$ supersymmetry in $\mathbb{R}^4$: a toy model

In order to understand the supersymmetry of the action (4.9), let us first discuss the case of $\mathbb{R}^4$ with the standard euclidean flat metric. We can think of $\mathbb{R}^4$ as $\mathbb{C}^2$ and introduce complex coordinates $Z^\alpha, \bar{Z}^{\bar{\alpha}}$ where $\alpha = 1, 2$. The analogous action to (4.9) in this case is given simply by

$$L_{\text{eff}} = \sum_\alpha \left(\dot{Z}^\alpha\dot{\bar{Z}}^{\bar{\alpha}} + i\zeta^{\bar{\alpha}}\dot{\zeta}^\alpha\right) \tag{4.10}$$

where ζ^α and $\zeta^{\bar{\alpha}}$ are the accompanying fermions. This toy action has four real supersymmetries. Two of them are manifest, as the next exercise asks you to show.

Exercise 4.10 ($N = 2$ supersymmetry in flat space)

Let $\delta_{\mathbb{1}}$ and δ_I be the supersymmetries defined as follows:

$$\begin{array}{llll} \delta_{\mathbb{1}} Z^\alpha = \zeta^\alpha & \delta_{\mathbb{1}}\zeta^\alpha = i\dot{Z}^\alpha & \delta_I Z^\alpha = i\zeta^\alpha & \delta_I\zeta^\alpha = \dot{Z}^\alpha \\ \delta_{\mathbb{1}}\bar{Z}^{\bar{\alpha}} = \zeta^{\bar{\alpha}} & \delta_{\mathbb{1}}\zeta^{\bar{\alpha}} = i\dot{Z}^{\bar{\alpha}} & \delta_I\bar{Z}^{\bar{\alpha}} = -i\zeta^{\bar{\alpha}} & \delta_I\zeta^{\bar{\alpha}} = -\dot{Z}^{\bar{\alpha}} \end{array}$$

(The names chosen for these transformations will appear more natural below.) Prove that they are invariances of the toy action (4.10), and that they satisfy the following algebra:

$$\delta_{\mathbb{1}}^2 = \delta_I^2 = i\frac{d}{dt} \quad \text{and} \quad \delta_{\mathbb{1}}\delta_I = -\delta_I\delta_{\mathbb{1}}$$

We can rewrite the second of these supersymmetries in a way that makes its generalisation obvious. If we let I denote the complex structure in $\mathbb{R}^4 = \mathbb{C}^2$ which is diagonalised by our choice of complex coordinates, then the second supersymmetry δ_I can be rewritten as follows:

$$\begin{array}{ll} \delta_I Z^\alpha = I^\alpha{}_\beta\zeta^\beta & \delta_I\zeta^\alpha = -iI^\alpha{}_\beta\dot{Z}^\beta \\ \delta_I\bar{Z}^{\bar{\alpha}} = I^{\bar{\alpha}}{}_{\bar{\beta}}\zeta^{\bar{\beta}} & \delta_I\zeta^{\bar{\alpha}} = -iI^{\bar{\alpha}}{}_{\bar{\beta}}\dot{Z}^{\bar{\beta}} \end{array}$$

which explains the notation. It now doesn't take much imagination to write down the remaining two supersymmetries. We simply replace I in turn by each of the other two complex structures J and K of Exercise 3.32. The fact that I, J, and K satisfy the quaternion algebra is instrumental in showing that these transformations obey the right supersymmetry algebra.

Exercise 4.11 (The $N = 4$ supersymmetry algebra)

Let $\delta_{\mathbb{1}}$ and δ_I be the supersymmetries given in Exercise 4.10, Define δ_J and δ_K in the obvious way. Let δ be any of these supersymmetries and $\delta' \neq \delta$ be a second of these supersymmetries. Prove that the following algebra is obeyed:

$$\delta^2 = i\frac{d}{dt} \quad \text{and} \quad \delta\delta' + \delta'\delta = 0$$

This is the $N = 4$ supersymmetry algebra.

10.2.2 4.3.2 $N = 4$ supersymmetry in hyperkähler manifolds

We now abandon our toy model and return to the action L_{eff} given by (4.9). We expect that the supersymmetry $\delta_{\mathbb{1}}$ defined in Exercise 4.10 should also be an invariance of L_{eff} and, given our choice of coordinates, that so should δI. This is because I is diagonal and constant on the chosen basis. In fact, leaving aside for the moment the issue of the invariance of L_{eff} under these transformations, Exercise 4.10 shows that they obey the right supersymmetry algebra. On the other hand, the other two complex structures J and K will not be constant in this basis, and hence the transformations δ_J and δ_K defined above will not obey the supersymmetry algebra. We will have to modify them appropriately.

To see this we will investigate the supersymmetry transformations associated to a a covariantly constant complex structure I. Let us not work on a complex basis adapted to I, but rather on some arbitrary basis (X^a, ζ^a). We will attempt to write down a supersymmetry transformation δ using I. Because the δ has order $\frac{1}{2}$ (being essentially a "square root" of d/dt), δX^a is determined up to an inconsequential overall constant:

$$\delta X^a = I_b{}^a \zeta^b. \tag{4.11}$$

Computing δ^2 we find

$$\delta^2 X^a = \partial_c I_b{}^a I_d{}^c \zeta^d \zeta^b + I_b{}^a \delta\zeta^b$$

If we now use the fact that I is covariantly constant, so that

$$\partial_c I_b{}^a = \Gamma_{cb}{}^d I_d{}^a - \Gamma_{cd}{}^a I_b{}^d$$

we can solve for $\delta\zeta^b$ by demanding that $\delta^2 X^a = i\dot{X}^a$:

$$\delta\zeta^a = -iI_b^a \dot{X}^b - \Gamma_{bc}{}^a I_d^b \zeta^d \zeta^c \tag{4.12}$$

where we have discarded a term $-\frac{1}{2} I_b{}^a T_{cd}{}^b I_e{}^c I_f{}^d \zeta^c \zeta^f$ where $T_{cd}{}^b \equiv \Gamma_{cd}{}^b - \Gamma_{dc}{}^b$ is the torsion of the connection, which in our case is zero. In order to show that $\delta^2 = id/dt$ on ζ^a, two approaches present themselves. One can use the fact that $\zeta^a = -I_b{}^a \delta X^b$ and use the fact that on (any function of) X, $\delta^2 = id/dt$:

$$\begin{aligned}
\delta^2\zeta^a &= -\delta^2 \left(I_b{}^a \delta X^b \right) \\
&= -i\partial_c I_b{}^a \dot{X}^c \delta X^b - I_b{}^a \delta^2 \delta X^b \\
&= -i\partial_c I_b{}^a \dot{X}^c \delta X^b - I_b{}^a \delta\delta^2 X^b && \text{since } \delta^2\delta = \delta^3 = \delta\delta^2 \\
&= -i\partial_c I_b{}^a \dot{X}^c \delta X^b - iI_b{}^a \delta\dot{X}^b \\
&= -i\partial_c I_b{}^a \dot{X}^c \delta X^b - iI_b{}^a \left(I_c{}^b \zeta^c \right)^{\cdot} \\
&= -i\partial_c I_b{}^a I_d{}^b \dot{X}^c \zeta^d - iI_b{}^a \partial_d I_c{}^b \dot{X}^d \zeta^c + i\dot{\zeta}^a \\
&= i\dot{\zeta}^a
\end{aligned}$$

where in the last line we have used an identity resulting from taking the derivative of $I^2 = -\mathbb{1}$. Alternatively, one can compute $\delta^2\zeta^a$ directly. This is naturally left as an exercise.

Exercise 4.12 (Another proof that $\delta^2\zeta^a = i\dot{\zeta}^a$)

By taking δ of $\delta\zeta^a$, prove that $\delta^2\zeta^a = i\dot{\zeta}^a$.

(Hint: You might find it necessary to use two properties of the Riemann curvature tensor: ${R_{abc}}^d + {R_{bca}}^d + {R_{cab}}^d = 0$; and ${R_{abc}}^d {I_d}^e = {R_{abd}}^e {I_c}^d$.

You are encouraged to prove these identities. The first one is the (first) Bianchi identity, the other one follows from the fact that I is covariantly constant, and hence commutes with the curvature operator.)

Let δ_I denote the supersymmetry transformation associated to the complex structure I. If we define $\delta_{\mathbb{1}}$ as above: $\delta_{\mathbb{1}} X^a = \zeta^a$ and $\delta_{\mathbb{1}}\zeta^a = i\dot{X}^a$, then just as before $\delta_{\mathbb{1}}\delta_I = -\delta_I\delta_{\mathbb{1}}$. In other words, $\delta_{\mathbb{1}}$ and δ_I generate an $N = 2$ supersymmetry algebra. Therefore this result holds for any Kähler manifold, and not just for $\mathbb{R}^4$ as in the previous section.

Now let J be a second covariantly constant complex structure. It will give rise to its own supersymmetry transformation given by equations (4.11) and (4.12), but with J replacing I. Let us call this supersymmetry transformation δ_J. When will δ_I and δ_J (anti)commute? The next exercise provides the answer.

Exercise 4.13 (Commuting supersymmetries)

Let δ_I and δ_J denote the supersymmetries generated by covariantly constant complex structure I and J. Prove that

$$\delta_I\delta_J X^a = -i{I_c}^b{J_b}^a\dot{X}^c - {\Gamma_{bc}}^a{I_d}^b{J_e}^c\zeta^d\zeta^e.$$

Conclude that $(\delta_I\delta_J + \delta_J\delta_I)\, X^a = 0$ if and only if $IJ = -JI$. Prove that if this is the case, then $(\delta_I\delta_J + \delta_J\delta_I)\,\zeta^a = 0$ as well, so that the two supersymmetries (anti)commute.

(Hint: To compute $\delta_I\delta_J + \delta_J\delta_I$ on ζ^a, you might find it easier to exhibit $\zeta^a = -{I_b}^a\delta_I X^b$, say, and then use that $\delta_I\delta_J + \delta_J\delta_I$ is zero on (functions of) X^a.)

This means that if $\{X^a\}$ denote the coordinates of a hyperkähler manifold and $\{\zeta^a\}$ are the accompanying fermions, then the four supersymmetries $\delta_{\mathbb{1}}$, δ_I, δ_J and δ_K satisfy an $N = 4$ supersymmetry algebra.

10.2.3 4.3.3 $N = 4$ supersymmetry of L_{eff}

It remains to show that the four supersymmetries defined above, are indeed symmetries of the effective action L_{eff} . This will be easier to ascertain if we first rewrite the supersymmetry transformations (4.11) and (4.12) in complex coordinates adapted to one of the complex structures: I, say.

Let us therefore choose coordinates $\left(Z^\alpha, \bar{Z}^{\bar{\alpha}}, \zeta^\alpha, \zeta^{\bar{\alpha}}\right)$ adapted to the complex structure I. Because the the metric is hermitian relative to this complex structure (in fact, relative to all three), we can rewrite the equations (4.11) and (4.12) using the results of Exercise 3.23, Because I is constant relative to these basis, δ_I precisely agrees with δ_I in Exercise 4.10.

Now consider the second complex structure J. Because $IJ = -JI, J$ maps vectors of type $(0, 1)$ relative to I to vectors of type $(1, 0)$ and viceversa. In other words, relative to the above basis adapted to I, J has components ${J_\alpha}^{\bar{\beta}}$ and ${J_{\bar{\alpha}}}^\beta$. Therefore J generates the following supersymmetry
transformation:

$$\delta_J Z^\alpha = i{J_{\bar{\beta}}}^\alpha\zeta^{\bar{\beta}} \;\; \delta_J\zeta^\alpha = -i{J_{\bar{\beta}}}^\alpha\dot{\bar{Z}}^{\bar{\beta}} - {\Gamma_{\beta\gamma}}^\alpha {J_{\bar{\delta}}}^\beta\zeta^{\bar{\delta}}\zeta^\gamma$$
$$\delta_J\bar{Z}^{\bar{\alpha}} = {J_\beta}^{\bar{\alpha}}\zeta^\beta \;\; \delta_J\zeta^{\bar{\alpha}} = -i{J_\beta}^{\bar{\alpha}}\dot{Z}^\beta - {\Gamma_{\bar{\beta}\bar{\gamma}}}^{\bar{\alpha}} J_\delta^{\bar{\beta}}\zeta^\delta\zeta^{\bar{\gamma}}$$

where we have used (see Exercise 3.23) that $\Gamma_{\alpha\beta}{}^{\gamma}$ and $\Gamma_{\bar\alpha\bar\beta}\bar\gamma^2$ are the only nonzero components of the Christoffel symbols. Similar formulas hold for δ_K.

As they stand, these supersymmetries are fermionic transformations. We can make them bosonic by introducing an anticommuting parameter ε and defining the following transformations:

$$\begin{aligned}
\delta_{\mathbb{1}}^{\varepsilon} Z^{\alpha} &= \varepsilon\zeta^{\alpha} \\
\delta_{\mathbb{1}}^{\varepsilon} \zeta^{\alpha} &= i\varepsilon\dot Z^{\alpha} \\
\delta_{\mathbb{1}}^{\varepsilon} \bar Z^{\bar\alpha} &= \varepsilon\zeta^{\bar\alpha} \\
\delta_{\mathbb{1}}^{\varepsilon} \zeta^{\bar\alpha} &= i\varepsilon\dot{\bar Z}^{\bar\alpha}
\end{aligned} \tag{4.13}$$

$$\begin{aligned}
\delta_{I}^{\varepsilon} Z^{\alpha} &= i\varepsilon\zeta^{\alpha} \\
\delta_{I}^{\varepsilon} \zeta^{\alpha} &= \dot Z^{\alpha} \\
\delta_{I}^{\varepsilon} \bar Z^{\bar\alpha} &= -i\varepsilon\zeta^{\bar\alpha} \\
\delta_{I}^{\varepsilon} \zeta^{\bar\alpha} &= -\varepsilon\dot{\bar Z}^{\bar\alpha}
\end{aligned} \tag{4.14}$$

$$\begin{aligned}
\delta_{J}^{\varepsilon} Z^{\alpha} &= i\varepsilon J_{\bar\beta}{}^{\alpha}\zeta^{\bar\beta} \\
\delta_{J}^{\varepsilon} \zeta^{\alpha} &= -i\varepsilon J_{\bar\beta}{}^{\alpha}\dot{\bar Z}^{\bar\beta} - \varepsilon\Gamma_{\beta\gamma}{}^{\alpha}J_{\bar\delta}{}^{\beta}\zeta^{\bar\delta}\zeta^{\gamma} \\
\delta_{J}^{\varepsilon} \bar Z^{\bar\alpha} &= \varepsilon J_{\beta}{}^{\bar\alpha}\zeta^{\beta} \\
\delta_{J}^{\varepsilon} \zeta^{\bar\alpha} &= -i\varepsilon J_{\beta}{}^{\bar\alpha}\dot Z^{\beta} - \varepsilon\Gamma_{\bar\beta\bar\gamma}{}^{\bar\alpha}J_{\delta}{}^{\bar\beta}\zeta^{\delta}\zeta^{\bar\gamma}
\end{aligned} \tag{4.15}$$

$$\begin{aligned}
\delta_{K}^{\varepsilon} Z^{\alpha} &= i\varepsilon K_{\bar\beta}{}^{\alpha}\zeta^{\bar\beta} \\
\delta_{K}^{\varepsilon} \zeta^{\alpha} &= -i\varepsilon K_{\bar\beta}{}^{\alpha}\dot{\bar Z}^{\bar\beta} - \varepsilon\Gamma_{\beta\gamma}{}^{\alpha}K_{\bar\delta}{}^{\beta}\zeta^{\bar\delta}\zeta^{\gamma} \\
\delta_{K}^{\varepsilon} \bar Z^{\bar\alpha} &= \varepsilon K_{\beta}{}^{\bar\alpha}\zeta^{\beta} \\
\delta_{K}^{\varepsilon} \zeta^{\bar\alpha} &= -i\varepsilon K_{\beta}{}^{\bar\alpha}\dot Z^{\beta} - \varepsilon\Gamma_{\bar\beta\bar\gamma}{}^{\bar\alpha}K_{\delta}{}^{\bar\beta}\zeta^{\delta}\zeta^{\bar\gamma}
\end{aligned} \tag{4.16}$$

The task ahead is now straightforward-albeit a little time consuming. Taking each of these supersymmetries in turn, and letting ϵ depend on time, we vary the action L_{eff} . Invariance of the action implies that

$$\delta^{\varepsilon} L_{\text{eff}} = \dot\varepsilon Q + \dot X$$

where X is arbitrary, and Q is the charge generating the supersymmetry. The next exercise summarises the results of this calculation.

Exercise 4.14 (The supersymmetry charges)

Prove that L_{eff} is invariant under the supersymmetries given by equations (4.13)(4.16), with the following associated supersymmetry charges:

$$\begin{aligned}
Q_{\mathbb{1}} &= g_{\alpha\bar\beta}\zeta^{\alpha}\dot Z^{\bar\beta} + g_{\alpha\bar\beta}\zeta^{\bar\beta}\dot Z^{\alpha} \\
Q_{I} &= ig_{\alpha\bar\beta}\zeta^{\alpha}\dot Z^{\bar\beta} - ig_{\alpha\bar\beta}\zeta^{\bar\beta}\dot Z^{\alpha} \\
Q_{J} &= J_{\alpha\beta}\zeta^{\alpha}\dot Z^{\beta} + J_{\bar\alpha\bar\beta}\zeta^{\bar\alpha}\dot Z^{\bar\beta} \\
Q_{K} &= K_{\alpha\beta}\zeta^{\alpha}\dot Z^{\beta} + K_{\bar\alpha\bar\beta}\zeta^{\bar\alpha}\dot Z^{\bar\beta}
\end{aligned}$$

where $J_{\alpha\beta} = J_{\alpha}\bar\gamma_{\beta\bar\gamma}$ and $J_{\bar\alpha\bar\beta} = J_{\bar\alpha}{}^{\gamma}g_{\bar\beta\gamma}$, and similarly for K.

(Hint: The calculation uses the two fundamental identities described in the hint to Exercise 4.12. In complex coordinates, and using the results of Exercise 3.23, they now look as follows:

$$R_{\alpha\bar{\beta}\gamma}{}^{\delta} = R_{\gamma\bar{\beta}\alpha}{}^{\delta} \qquad R_{\alpha\bar{\beta}\bar{\gamma}}{}^{\bar{\delta}} = R_{\alpha\bar{\gamma}\bar{\beta}}{}^{\bar{\delta}}$$
$$R_{\alpha\bar{\beta}\gamma}{}^{\epsilon} J_{\epsilon}{}^{\bar{\delta}} = R_{\alpha\bar{\beta}\bar{\epsilon}}{}^{\bar{\delta}} J_{\gamma}{}^{\bar{\epsilon}} \quad R_{\alpha\bar{\beta}\bar{\gamma}}{}^{\epsilon} J_{\bar{\epsilon}}{}^{\delta} = R_{\alpha\bar{\beta}\epsilon}{}^{\delta} J_{\bar{\gamma}}{}^{\epsilon}$$

and similarly for K.)

Having established the $N = 4$ supersymmetry of the effective action L_{eff} it is now time to quantise the theory. It turns out that supersymmetry will provide a geometric description of the Hilbert space and of the hamiltonian. This will require some basic concepts of harmonic theory on Kähler manifolds. The purpose of the next section is to provide a brief review for those who are not familiar with this topic.

10.3 4.4 A brief review of harmonic theory

This section contains a brief scholium on the harmonic theory of orientable riemannian manifolds and in particular of Kähler manifolds. The reader familiar with these results can easily skip this section. Other readers are encouraged to read on. We will of necessity be brief: details can be found in many fine books on the subject [Gol62, GH78, War83, Wel80].

10.3.1 4.4.1 Harmonic theory for riemannian manifolds

Let M be a smooth manifold. We will let $\mathcal{E} = \bigoplus_p \mathcal{E}^p$ denote the algebra of differential forms on M. The de Rham operator $d : \mathcal{E}^p \to \mathcal{E}^{p+1}$ obeys $d^2 = 0$ and hence one can define its cohomology (the de Rham cohomology of M) as follows:

$$H^p_{\text{dR}}(M) = \frac{\ker d : \mathcal{E}^p \to \mathcal{E}^{p+1}}{\operatorname{im} d : \mathcal{E}^{p-1} \to \mathcal{E}^p}$$

In other words, the p-th de Rham cohomology is a vector space whose elements are equivalence classes of closed p-forms ($d\omega = 0$) —two closed p-forms ω_1 and ω_2 being equivalent if their difference is exact: $\omega_1 - \omega_2 = d\theta$, for some $(p-1)$-form θ. The crown jewel of harmonic theory is the decomposition theorem of Hodge, which states that if M is a compact orientable manifold there exists a privileged representative for each de Rham cohomology class. This representative is obtained by introducing more structure on M-namely a riemannian metric. From the above definition, it is clear that
the de Rham cohomology does not depend on any geometric properties of the manifold. It is precisely this reason why the Hodge theorem is of fundamental importance: because it establishes a link between the topological and the geometric properties of riemannian manifolds. (Actually, the fact that the de Rham cohomology is a topological invariant of M is not obvious. It is called the de Rham theorem and it is proven in War83, BT81.)

We therefore let (M, g) be an m-dimensional orientable riemannian manifold. Let $\{e^i\}$ for $i = 1, \ldots, m$ be a local orthonormal basis for the 1-forms. Orthonormality means that the line element is locally $ds^2 = \sum_i e^i \otimes e^i$. In general such a basis will of course not exist globally, but will transform under a local $O(m)$ transformation when we change coordinate charts. In this basis the volume form is given by $\text{vol} = e^1 \wedge e^2 \wedge \cdots \wedge e^m$. This volume form defines a local orientation in M. Orientability simply means that, unlike the 1-forms $\{e^i\}$, the volume form-and hence the orientation-does exist globally. It also means that upon changing charts, the $\{e^i\}$ will change by a local $SO(m)$ transformation. (Prove this!)

A local basis for the differential forms $\mathcal{E}$ on M is given by wedge products of these 1-forms. It is convenient to introduce multi-indices $I = (i_1, i_2, \ldots, i_p)$ where $1 \leq i_1 < i_2 < \cdots < i_p \leq m$. We say that I has length p or that $|I| = p$. We then define $e^I \equiv e^{i_1} \wedge e^{i_2} \wedge \cdots \wedge e^{i_p}$. In this notation, $\left\{e^I | I | I\ |, = p\right\}$ is a local basis for $\mathcal{E}^p$; that is, any p-form ω on M can be written locally like $\sum_{|I|=p} \omega_I e^I$, where the coefficients ω_I are smooth functions. If $I = (i_1, i_2, \ldots, i_p)$ is a multi-index of length p, we let $I^c = (i_{p+1}, i_{p+2}, \ldots, i_m)$ denote the multi-index of length $m-p$ uniquely defined by the fact that $\{i_1, \ldots, i_p\} \cup \{i_{p+1}, \ldots, i_m\} = \{1, 2, \ldots, m\}$.

We can now define the Hodge $\star$-operation. This is a linear map $\star : \mathcal{E}^p \to \mathcal{E}^{m-p}$ defined by

$$\star e^I = \operatorname{sign} \sigma e^{I^c}$$

where if $I = (i_1, \ldots, i_p)$ and $I^c = (i_{p+1}, \ldots, i_m)$, then sign σ is the sign of the permutation

$$\sigma = \begin{pmatrix} i_1 & i_2 & \cdots & i_m \\ 1 & 2 & \cdots & m \end{pmatrix}$$

In particular $\star 1 = \mathrm{vol}$. The following result is important for calculations.

Exercise 4.15 (The square of the Hodge $\star$)

Prove that acting on $\mathcal{E}^p, \star^2 = (-)^{(m-p)p}$.

The Hodge $\star$-operator allows us to define a pointwise metric $\langle -, - \rangle$ on forms as follows:

$$\alpha \wedge \star\beta = \langle \alpha, \beta \rangle \,\mathrm{vol} \tag{4.17}$$

The properties of this pointwise metric are summarised in the following exercise.

Exercise 4.16 (The pointwise metric on forms)

Prove that the basis $\left\{e^I\right\}$ is orthonormal relative to the pointwise metric defined in (4.17) and therefore that it agrees on 1-forms with the one induced by the riemannian metric on M. Conclude that the pointwise metric is positive-definite.

If, in addition, M is compact we can define an honest metric (called the Hodge metric) on forms by integrating the pointwise metric over the manifold relative to the volume form:

$$(\alpha, \beta) = \int_M \langle \alpha, \beta \rangle \mathrm{vol} = \int_M \alpha \wedge \star\beta$$

If M is not compact, then we can restrict ourselves to compactly supported forms or to forms α for which the Hodge norm $\|\alpha\|^2 \equiv (\alpha, \alpha)$ is finite. Such forms are often called square-integrable.

The Hodge metric allows us to define the adjoint d^* to the de Rham operator, with which the following exercise concerns itself.

Exercise 4.17 (The adjoint de Rham operator)

Define the adjoint d^* of the de Rham operator by

$$(d\alpha, \beta) = (\alpha, d^*\beta)$$

for all forms $\alpha, \beta \in \mathcal{E}$. Prove that d^* satisfies the following properties:
(1) $d^* : \mathcal{E}^p \to \mathcal{E}^{p+1}$;
(2) $(d^*)^2 = 0$; and
(3) $d^* = (-)^{mp+m+1} \star d \star$ acting on $\mathcal{E}^p$.

Now let us define the Hodge laplacian $\triangle : \mathcal{E}^p \to \mathcal{E}^p$ by $\triangle \equiv dd^* + d^*d$. We say that a p-form is harmonic if $\triangle\alpha = 0$.

Exercise 4.18 (Harmonic forms)

Prove that a form α is harmonic if and only if $d\alpha = d^*\alpha = 0$. Prove that harmonic forms have minimal Hodge norm in their cohomology class. That is, if α is harmonic, then prove that the Hodge norm of $\alpha + d\beta$ is strictly greater than that of α.

The Hodge decomposition theorem states that in a compact orientable manifold each de Rham cohomology class has a unique harmonic representative; that is, that there is a vector space isomorphism

$$H^p_{\mathrm{dR}}(M) \cong \text{ harmonic } p\text{-forms}$$

The proof of this theorem is rather technical. The idea is to use the normminimising property to define the harmonic representative; but one then has to prove that this form is smooth. This calls for the use of regularity theorems which are beyond the scope of these notes. A proof can be found, for example, in War83.

The Hodge decomposition theorem has a very important corollary, which the following exercise asks you to prove.

Exercise 4.19 (Poincaré duality)

Prove that the Hodge *-operator commutes with the Hodge laplacian. Use the Hodge decomposition theorem to conclude that for M an m-dimensional compact orientable manifold, there is an isomorphism

$$H^p_{\mathrm{dR}}(M) \cong H^{m-p}_{\mathrm{dR}}(M)$$

This isomorphism is known as Poincaré duality.

10.3.2 4.4.2 Harmonic theory for Kähler manifolds

Now suppose that M is a complex manifold of complex dimension n. As explained in section 3.3.2, on a complex manifold one has local coordinates $\left(z^\alpha, \bar{z}^{\bar{\beta}}\right)$, where $\alpha, \beta = 1, 2, \ldots, n$. This allows us to refine the grading of the complex differential forms. We say that a complex differential form ω is of type (p,q) if it can be written in local complex coordinates as

$$\omega = \omega_{\alpha_1\cdots\alpha_p\bar{\beta}_1\cdots\bar{\beta}_q}(z,\bar{z})dz^{\alpha_1}\wedge\cdots\wedge dz^{\alpha_p}\wedge d\bar{z}^{\bar{\beta}_1}\wedge\cdots\wedge d\bar{z}^{\bar{\beta}_q}$$

where $\omega_{\alpha_1\cdots\alpha_p\bar{\beta}_1\cdots\bar{\beta}_q}(z,\bar{z})$ are smooth functions. The algebra of complex differential forms is then bigraded as follows:

$$\mathcal{E} = \bigoplus_{0\le p,q\le n} \mathcal{E}^{p,q} \tag{4.18}$$

where $\mathcal{E}^{p,q}$ is the space of (p,q)-forms.

The de Rham operator d also breaks up into a type (1,0) piece and a type $(0,1)$ piece:

$$\begin{aligned} &d = \partial + \bar{\partial} \quad \text{where} \\ &\partial : \mathcal{E}^{p,q} \to \mathcal{E}^{p+1,q} \\ &\bar{\partial} : \mathcal{E}^{p,q} \to \mathcal{E}^{p,q+1} \end{aligned}$$

Breaking $d^2 = 0$ into types we find that $\partial^2 = \bar{\partial}^2 = \partial\bar{\partial} + \bar{\partial}\partial = 0$. We call $\bar{\partial}$ the Dolbeault operator, and its cohomology

$$H^{p,q}_{\bar{\partial}}(M) = \frac{\ker\bar{\partial} : \mathcal{E}^{p,q} \to \mathcal{E}^{p,q+1}}{\operatorname{im}\bar{\partial} : \mathcal{E}^{p,q-1} \to \mathcal{E}^{p,q}}$$

the Dolbeault cohomology.

Now suppose that we give M a hermitian metric h; that is, a riemannian metric compatible with the complex structure: $h(IX, IY) = h(X, Y)$. Such metrics always exist: one simply takes any riemannian metric g, say, and averages it over the finite group generated by $I : h(X, Y) \equiv \frac{1}{2}g(X, Y) + \frac{1}{2}g(IX, IY)$. If we forget the complex structure for a moment, M is an orientable riemannian manifold of (real) dimension $2n$. Therefore we have a Hodge $\star$-operator defined as in the previous section. The next exercise asks you to show how $\star$ interacts with the complex structure.

Exercise 4.20 (The Hodge $\star$ and the complex structure)

Prove that the Hodge $\star$-operator maps (p, q)-forms to $(n-q, n-p)$-forms:

$$\star : \mathcal{E}^{p,q} \to \mathcal{E}^{n-q,n-p}$$

and that acting on (p, q)-forms, $\star^2 = (-)^{p+q}$.

(Hint: The first part is computationally quite involved, but the idea is easy. We can always find a local basis $\{\theta^i\}$ for the (1,0)-forms on M such that the line element (relative to the hermitian metric) has the form

$$ds^2 = \sum_{i=1}^{n} \left(\theta^i \otimes \bar{\theta}^i + \bar{\theta}^i \otimes \theta^i\right)$$

where $\left\{\bar{\theta}^i\right\}$ are the complex conjugate (0,1)-forms. We can decompose these forms into their real and imaginary parts as follows: $\theta^j = \frac{1}{\sqrt{2}}\left(e^{2j-1} + ie^{2j}\right)$ and $\bar{\theta}^j = \frac{1}{\sqrt{2}}\left(e^{2j-1} - ie^{2j}\right)$. In terms of these real 1-forms, the line element becomes $ds^2 = \sum_{j=1}^{2n} e^j \otimes e^j$; in other words, they form an orthonormal basis. Therefore we know the action of the Hodge $\star$-operator on the $\left\{e^I\right\}$. Your mission, should you decide to accept it, is to find out what it is in terms of the $\theta^I \wedge \bar{\theta}^J$. The second part simply uses Exercise 4.15,

Another operation that we have on a complex manifold is complex conjugation, which exchanges (p, q)-forms with (q, p)-forms. Using the Hodge *-operator and complex conjugation we can define a pointwise hermitian metric for the complex forms, also denoted $\langle -, - \rangle$ as in the real case treated in the previous section. This metric is defined by

$$\alpha \wedge \star\bar{\beta} = \langle \alpha, \beta \rangle \mathrm{vol}$$

Notice that relative to this metric, the decomposition in equation (4.18) is orthogonal: if β is a (p, q)-form, then $\bar{\beta}$ is a (q, p)-form, and $\star\bar{\beta}$ is a $(n-p, n-q)$-form. The only way one can obtain the volume form, which is an (n, n)-form, is to wedge with another (p, q)-form.

Exercise 4.21 (The pointwise hermitian metric)

Prove that the basis $\left\{\theta^I \wedge \bar{\theta}^J\right\}$ is orthonormal relative to the pointwise hermitian metric, and conclude that it is positive-definite.

If M is compact, we can then integrate this pointwise metric relative to the volume form and define an honest hermitian metric on complex forms:

$$(\alpha, \beta) = \int_M \langle \alpha, \beta \rangle \mathrm{vol} = \int_M \alpha \wedge \star\bar{\beta} \tag{4.19}$$

This metric is again called the Hodge metric. As in the real case, if M is not compact, then we can still make sense of this provided we restrict our attention to square-integrable forms.

It follows from Exercise 4.17 that on a complex manifold, $d^* = - \star d \star$ regardless on which forms it is acting. Breaking d^* into types we find

$$d^* = \partial^* + \bar{\partial}^* \quad \text{where} \quad \begin{array}{l} \partial^* : \mathcal{E}^{p,q} \to \mathcal{E}^{p-1,q} \\ \bar{\partial}^* : \mathcal{E}^{p,q} \to \mathcal{E}^{p,q-1} \end{array}$$

On the other hand, breaking $- \star d \star$ into types, and comparing we see that

$$\partial^* = - \star \bar{\partial} \star \quad \text{and} \quad \bar{\partial}^* = - \star \partial \star$$

We can therefore define two laplacian operators:

$$\Box = \partial\partial^* + \partial^*\partial \quad \text{and} \quad \bar{\Box} = \bar{\partial}\bar{\partial}^* + \bar{\partial}^*\bar{\partial}$$

both of which map $\mathcal{E}^{p,q} \to \mathcal{E}^{p,q}$.

Just as for de Rham cohomology, there is a Hodge decomposition theorem for Dolbeault cohomology, which says that the $\bar{\partial}$-cohomology on (p,q)-forms is isomorphic to the space of $\bar{\Box}$-harmonic (p,q)-forms:

$$H^{p,q}_{\bar{\partial}} \cong \bar{\Box}\text{-harmonic } (p,q)\text{-forms.}$$

In a generic complex manifold there is no reason to expect any relation between the Dolbeault laplacians and the Hodge laplacian $\triangle = d^*d + dd^*$; but the magic of Kähler geometry is that if M is Kähler, then

$$\triangle = 2\Box = 2\Box. \tag{4.20}$$

This is not a hard result to obtain, but it requires quite a bit of formalism that we will not need in the remainder of this course, hence we leave it unproven and refer the interested reader to the literature Gol62, GH78, Wel80.

As an immediate corollary of equation (4.20) and of the Hodge decomposition theorems for de Rham and Dolbeault cohomologies, we have

$$H^r_{\mathrm{dR}}(M) \cong \bigoplus_{p+q=r} H^{p,q}_{\bar{\partial}}(M)$$

and the following exercise describes another immediate corollary of equation (4.20).

Exercise 4.22 (Serre duality)

Prove that both the Hodge *-operator and complex conjugation commute with the laplacian. Use this to conclude that for M a compact Kähler manifold of complex dimension n, there exist isomorphisms:

$$H^{p,q}_{\bar{\partial}}(M) \cong H^{n-q,n-p}_{\bar{\partial}}(M) \cong H^{n-p,n-q}_{\bar{\partial}}(M)$$

These isomorphisms are known collectively as Serre duality.

Finally, a curiosity. If we define the r-th Betti number b_r of a manifold as the real dimension of the r-th de Rham cohomology, we have as a consequence of Serre duality that for a compact Kähler manifold all the odd Betti numbers are even.

10.3.3 4.4.3 Explicit formulas for $\bar{\partial}$ and $\bar{\partial}^*$

The purpose of this section is simply to derive some explicit expressions for the differential operators $\bar{\partial}$ and $\bar{\partial}^*$. These are the expressions by which we will be able to recognise them when we quantise the effective action. Throughout this section M shall be a Kähler manifold of complex dimension n.

Let us first start by deriving formulas for d and d^*. For this we can forget momentarily the complex structure and think of M simply as an orientable riemannian manifold of dimension $2n$. Let $\{e_i\}$ denote a local orthonormal basis for the vector fields, and let $\{e^i\}$ denote the canonical dual basis for the 1-forms. They are also orthonormal relative to the induced metric. Let ∇ denote the Levi-Civita connection. We claim that d can be written as

$$d = \sum_{i=1}^{2n} e^i \wedge \nabla_{e_i} \tag{4.21}$$

Proving this will be the purpose of the following exercise.

Exercise 4.23 (An explicit expression for d)

Let $d?$ denote the right-hand side of equation (4.21).

Prove that $d_?$ is independent of the orthonormal basis chosen so that it is well-defined.

Let $e^i = e^i_a dx^a$ and $e_i = e^a_i \partial_a$. Prove that $d_? = \sum_a dx^a \wedge \partial_a$, and conclude that $d_? = d$.

(Hint: Use the fact that the Levi-Civita connection is torsionless.)

With this result we can now describe a similar formula for d^*. Letting $\{e_i\}$ and $\{e^i\}$ be as above, we will prove that

$$d^* = -\sum_{i=1}^{2n} \iota\left(e^i\right) \nabla_{e_i} \tag{4.22}$$

where $\iota\left(e^i\right)$ is the contraction operation, defined by:

$\iota\left(e^i\right) f = 0$ for f a function; $\iota\left(e^i\right) e^j = \delta_{ij}$; and $\iota\left(e^i\right)\left(e^j \wedge \omega\right) = \delta_{ij}\omega - e^j \wedge \iota\left(e^i\right)\omega$.

The next exercise asks you to prove equation (4.22).

Exercise 4.24 (An explicit expression for d^*)

Let $d^*_?$ stand for the right-hand side of equation (4.22). Prove that $d^*_? = -\star d \star$, whence it agrees with d^*. (We are using (3) in Exercise 4.17, with $m = 2n$.)

(Hint: Prove first that $d^*_?$ is well-defined; that is, it is independent of the choice of orthonormal frame. Because of this and by linearity, conclude that it is sufficient to compare $d^*_?$ and $-\star d \star$ on a p-form of the form $f e^1 \wedge e^2 \wedge \cdots \wedge e^p$. Moreover argue that it is sufficient to compute this at a point where $\nabla_{e_i} e^j = 0$. Then do it.)

As a corollary of the previous exercise, it follows that relative to a coordinate basis, we can write

$$d^* = -\sum_a \iota\left(dx^a\right) \partial_a \tag{4.23}$$

where now $\iota\left(dx^a\right) dx^b = g^{ab}$.

We can now re-introduce the complex structure. Let $\left\{\theta_i, \bar{\theta}_i\right\}$ be a complex basis for the complex vector fields, and let $\left\{\theta^i, \bar{\theta}^i\right\}$ be the canonical dual basis for the complex 1-forms. In terms of the above basis $\left\{e^i\right\}, \theta^i$ is given as in Exercise 4.20. The canonical dual basis for the vector fields are related by

$$\theta_i = \frac{1}{\sqrt{2}}\left(e_{2i-1} - i e_{2i}\right) \quad \bar{\theta}_i = \frac{1}{\sqrt{2}}\left(e_{2i-1} + i e_{2i}\right)$$

Inverting this change of basis, and using equations (4.21) and (4.22), we find

$$d = \sum_{i=1}^{n} \left(\theta^i \wedge \nabla_{\theta_i} + \bar{\theta}^i \wedge \nabla_{\bar{\theta}_i} \right)$$

and

$$d^* = -\sum_{i=1}^{n} \left(\iota\left(\theta^i\right) \nabla_{\theta_i} + \iota\left(\bar{\theta}^i\right) \nabla_{\bar{\theta}_i} \right)$$

Breaking up into types, one concludes that

$$\bar{\partial} = \sum_{i=1}^{n} \bar{\theta}^i \wedge \nabla_{\bar{\theta}_i} \quad \text{and} \quad \bar{\partial}^* = -\sum_{i=1}^{n} \iota\left(\theta^i\right) \nabla_{\theta_i}$$

Or in a coordinate basis,

$$\bar{\partial} = \sum_{\bar{\alpha}=1}^{n} d\bar{z}^{\bar{\alpha}} \wedge \partial_{\bar{\alpha}} \quad \text{and} \quad \bar{\partial}^* = -\sum_{\alpha=1}^{n} \iota\left(dz^\alpha\right) \partial_\alpha \tag{4.24}$$

These equations will be important in the sequel.

10.3.4 4.5 Quantisation of the effective action

In this section we discuss the canonical quantisation of the effective action (4.8). We will be able to identify the Hilbert space with the square-integrable $(0, q)$-forms on the moduli space $\mathcal{M}_k$. We will exploit the supersymmetry to write the hamiltonian as the anticommutator of supersymmetry charges which, under the aforementioned isomorphism, will be identified as the Dolbeault operator $\bar{\partial}$ and its adjoint under the Hodge metric. This will then allow us to identify the ground states of the effective quantum theory as the harmonic $(0, q)$-forms on the moduli space.

10.3.5 4.5.1 Canonical analysis

The first step in this direction is to find the expression for the canonical momenta. Then we write the hamiltonian and the supersymmetry charges in terms of the momenta. We write down the Poisson brackets and make sure that the classical algebra is indeed the $N = 4$ supersymmetry algebra. Most of these calculations are routine, and are therefore left as exercises.

The first exercise starts you in this path by asking you to compute the canonical momenta.

Exercise 4.25 (The canonical momenta)

Prove that the canonical momenta defined by L_{eff} take the following form:

$$\begin{aligned}
P_\alpha &= \frac{\partial L_{\text{eff}}}{\partial \dot{Z}^\alpha} = g_{\alpha\bar{\beta}} \dot{\bar{Z}}^{\bar{\beta}} + i \Gamma_{\alpha\beta\bar{\gamma}} \zeta^{\bar{\gamma}} \zeta^\beta \\
\bar{P}_{\bar{\alpha}} &= \frac{\partial L_{\text{eff}}}{\partial \dot{Z}^{\bar{\alpha}}_{\bar{\alpha}}} = g_{\bar{\alpha}\beta} \dot{Z}^\beta \\
\pi_\alpha &= \frac{\partial L_{\text{eff}}}{\partial \dot{\zeta}^\alpha} = -i g_{\alpha\bar{\beta}} \zeta^{\bar{\beta}} \\
\pi_{\bar{\alpha}} &= \frac{\partial L_{\text{eff}}}{\partial \dot{\zeta}^{\bar{\alpha}}} = 0
\end{aligned}$$

where $\Gamma_{\alpha\beta\bar{\gamma}} = \Gamma_{\alpha\beta}{}^\delta g_{\delta\bar{\gamma}}$.

The fact that $\pi_{\bar{\alpha}} = 0$ is not very important. It is simply a consequence of the fact that the fermionic part of the effective lagrangian is already in first order form, so that morally speaking $\{\zeta^{\bar{\alpha}}\}$ play the role of momenta while $\{\zeta^{\alpha}\}$ are coordinates.

The effective hamiltonian H_{eff} is defined as usual by:

$$H_{\text{eff}} = \dot{Z}^{\alpha} P_{\alpha} + \dot{\bar{Z}}^{\bar{\alpha}} \bar{P}_{\bar{\alpha}} + \dot{\zeta}^{\alpha} \pi_{\alpha} - L_{\text{eff}}$$

The next exercise asks you to compute it.

Exercise 4.26 (The effective hamiltonian)

Prove that the effective hamiltonian is given by

$$H_{\text{eff}} = g^{\alpha\bar{\beta}} P_{\alpha} \bar{P}_{\bar{\beta}} + g^{\alpha\bar{\beta}} \Gamma_{\alpha\gamma}{}^{\delta} \bar{P}_{\bar{\beta}} \pi_{\delta} \zeta^{\gamma}$$

Next we write the supersymmetry charges obtained in Exercise 4.14 in terms of momenta. This is another easy exercise.

Exercise 4.27 (The supersymmetry charges revisited)

Prove that the supersymmetry charges obtained in Exercise 4.14 have the following form:

$$\begin{aligned}
Q_{\mathbb{1}} &= \zeta^{\alpha} P_{\alpha} + \zeta^{\bar{\alpha}} \bar{P}_{\bar{\alpha}} = \zeta^{\alpha} P_{\alpha} + i g^{\alpha\bar{\beta}} \pi_{\alpha} \bar{P}_{\bar{\beta}} \\
Q_{I} &= i\zeta^{\alpha} P_{\alpha} - i\zeta^{\bar{\alpha}} \bar{P}_{\bar{\alpha}} = i\zeta^{\alpha} P_{\alpha} + g^{\alpha\bar{\beta}} \pi_{\alpha} \bar{P}_{\bar{\beta}} \\
Q_{J} &= J_{\alpha}\bar{\alpha}^{\bar{\alpha}} \bar{P}_{\bar{\alpha}} + J_{\bar{\alpha}}{}^{\alpha} \zeta^{\bar{\alpha}} P_{\alpha} - i J_{\bar{\alpha}}^{\alpha} \Gamma_{\alpha\beta\bar{\gamma}} \zeta^{\bar{\alpha}} \zeta^{\bar{\gamma}} \zeta^{\beta} \\
Q_{K} &= K_{\alpha}{}^{\bar{\alpha}} \zeta^{\alpha} \bar{P}_{\bar{\alpha}} + K_{\bar{\alpha}}{}^{\alpha} \zeta^{\bar{\alpha}} P_{\alpha} - i K_{\bar{\alpha}}{}^{\alpha} \Gamma_{\alpha\beta\bar{\gamma}} \zeta^{\bar{\alpha}} \zeta^{\bar{\gamma}} \zeta^{\beta}
\end{aligned}$$

The canonical Poisson brackets are defined to be the following:

$$\left\{P_{\alpha}, Z^{\beta}\right\} = \delta_{\alpha}^{\beta} \quad \left\{\bar{P}_{\bar{\alpha}}, \bar{Z}^{\bar{\beta}}\right\} = \delta_{\bar{\alpha}}^{\bar{\beta}} \quad \left\{\pi_{\alpha}, \zeta^{\beta}\right\} = \delta_{\alpha}^{\beta}$$

Exercise 4.28 (Some checks)

As a check on our calculations, show that the supersymmetry transformations given in equations (4.13)-(4.16) are indeed generated via Poisson brackets by the supersymmetry charges computed in the previous exercise.

Finally, we are ready to verify that we have a classical realisation of the $N = 4$ supersymmetry algebra. Let $q = q_1 i + q_2 j + q_3 k + q_4 \in \mathbb{H}$ be a quaternion. Let $Q_q \equiv q_1 Q_I + q_2 Q_J + q_3 Q_K + q_4 Q_{\mathbb{1}}$. The next exercise asks you to prove that the supersymmetry charges obey the $N = 4$ supersymmetry algebra.

Exercise 4.29 (Classical $N = 4$ supersymmetry algebra)

Let $q, q' \in \mathbb{H}$ be quaternions. Then prove that under Poisson bracket:

$$\{Q_q, Q_{q'}\} = i\,(\bar{q}q')\, H_{\text{eff}}$$

where $\bar{q} = -q_1 i - q_2 j - q_3 k + q_4$ is the conjugate quaternion and $\bar{q}q' = \sum_i q_i q_i'$ is the quaternionic product.

Because the supersymmetry charges generate under Poisson bracket the supersymmetry transformations, the above exercise implies that the effective hamiltonian indeed generates time translation. If you feel up to it you can check this directly from the expression of the hamiltonian.

10.3.6 4.5.2 The quantisation of the effective hamiltonian

To quantise the effective hamiltonian we first need to identify the Hilbert space. Let us quickly quantise the bosons. We choose to realise Z^α and $\bar{Z}^{\bar{\alpha}}$ as multiplication operators and hence P_α and $\bar{P}_{\bar{\alpha}}$ will be realised as derivatives:

$$P_\alpha \mapsto -i\frac{\partial}{\partial Z^\alpha} \quad \text{and} \quad \bar{P}_{\bar{\alpha}} \mapsto -i\frac{\partial}{\bar{Z}^{\bar{\alpha}}}$$

For the fermions, we notice that the canonical Poisson brackets can be rewritten in terms of ζ^α and $\zeta^{\bar{\alpha}}$ as follows:

$$\left\{\zeta^\alpha, \zeta^{\bar{\beta}}\right\} = ig^{\alpha\bar{\beta}}$$

Upon quantisation this gives rise to the following anticommutation relations

$$\zeta^\alpha\zeta^{\bar{\beta}} + \zeta^{\bar{\beta}}\zeta^\alpha = g^{\alpha\bar{\beta}}$$

with all other anticommutators vanishing. Of course, $g^{\alpha\bar{\beta}}$ is a function of $Z^\alpha, \bar{Z}^{\bar{\alpha}}$; but for each point $\left(Z^\alpha, \bar{Z}^{\bar{\alpha}}\right)$ in $\mathcal{M}_k$, the above anticommutation relations define a Clifford algebra. In other words, this defines a Clifford bundle on $\mathcal{M}_k$. Fixing a point in $\mathcal{M}_k$, we have a standard Clifford algebra of
the type studied in section 2.2.2. It has a unique irreducible representation constructed as follows. We choose a Clifford vacuum $|\Omega\rangle$, defined by the condition

$$\zeta^\alpha|\Omega\rangle = 0 \quad \text{for all } \alpha$$

The representation is then built on $|\Omega\rangle$ by acting with the $\zeta^{\bar{\alpha}}$.

We now tensor together the representations of the bosons and the fermions and what we have is linear combinations of objects of the form

$$f(Z,\bar{Z})\zeta^{\bar{\alpha}}\zeta^{\bar{\beta}}\cdots\zeta^{\bar{\gamma}}|\Omega\rangle$$

If we take the $f(Z,\bar{Z})$ smooth, this space is clearly isomorphic to the space $\oplus_{0\leq p\leq 2k}\mathcal{E}^{0,p}$ of differential forms of type $(0,p)$ on $\mathcal{M}_k$:

$$f(Z,\bar{Z})\zeta^{\bar{\alpha}}\zeta^{\bar{\beta}}\cdots\zeta^{\bar{\gamma}}|\Omega\rangle \leftrightarrow f(Z,\bar{Z})d\bar{Z}^{\bar{\alpha}} \wedge d\bar{Z}^{\bar{\beta}} \wedge \cdots \wedge d\bar{Z}^{\bar{\gamma}}$$

Of course the Hilbert space will consist of (the completion of) the subspace formed by those forms which are square integrable relative to a suitable inner product. As we saw in section 4.4.2, the natural inner product to consider is the Hodge metric given by (4.19). Therefore we have the following geometric interpretation of the Hilbert space $\mathcal{H}$ of the quantum effective theory:

$$\mathcal{H} \cong \bigoplus_{0\leq p\leq 2k} \mathcal{E}^{0,p}_{L^2} \tag{4.25}$$

where $\mathcal{E}^{0,p}_{L^2}$ denotes the space of $(0,p)$-forms on $\mathcal{M}_k$ with finite Hodge norm; that is, square-integrable.

In order to identify the hamiltonian we will use supersymmetry. The expressions for the supersymmetry charges and the hamiltonian, being polynomial, suffer from ordering ambiguities. One way to get around this problem is to define the quantisation in a way that the $N = 4$ supersymmetry algebra is realised quantum-mechanically, and in such a way that we can identify the resulting operators geometrically. The hamiltonian can be defined as the square of any of the supersymmetry charges, but we find it more convenient to take complex linear combinations of $Q_{\mathbb{1}}$ and Q_I. Indeed, let us define

$$Q = i\frac{1}{2}(Q_{\mathbb{1}} + iQ_I) \quad \text{and} \quad Q^* = -i\frac{1}{2}(Q_{\mathbb{1}} - iQ_I)$$

The classical expressions for these charges are very simple

$$Q = i\zeta^{\bar\alpha}\bar{P}_{\bar\alpha} \quad \text{and} \quad Q^* = -i\zeta^{\alpha}P_{\alpha}$$

and they obey the following algebra

$$\{Q, Q^*\} = i\frac{1}{2}H_{\text{eff}} \tag{4.26}$$

The quantisation is now clear. Quantise the charges Q and Q^* as follows:

$$Q \mapsto \zeta^{\bar\alpha}\frac{\partial}{\partial \bar{Z}^{\bar\alpha}} \quad \text{and} \quad Q^* \mapsto -\zeta^{\alpha}\frac{\partial}{\partial Z^{\alpha}} \tag{4.27}$$

But notice that we have seen these operators before. Indeed, acting on $f \equiv f_{\bar\alpha\bar\beta\cdots\bar\gamma}\zeta^{\bar\alpha}\zeta^{\bar\beta}\cdots\zeta_{\bar\gamma}|\Omega\rangle$, we find that

$$Qf = \partial_{\bar\delta}f_{\bar\alpha\bar\beta\cdots\bar\gamma}\zeta^{\bar\delta}\zeta^{\bar\alpha}\zeta^{\bar\beta}\cdots\zeta^{\bar\gamma}|\Omega\rangle$$

Under the isomorphism (4.25), this corresponds to the form $\bar\partial f$. In other words, $Q \mapsto \bar\partial$.

How about Q^*? Acting on a $(0,0)$-form, Q^* is zero, since ζ^{α} annihilates the Clifford vacuum. Acting on a $(0,1)$-form $f_{\bar\alpha}\bar\alpha_{\bar\alpha}|\Omega\rangle$, we find

$$Q^* f_{\bar\alpha}\zeta^{\bar\alpha}|\Omega\rangle = -\partial_{\beta}f_{\bar\alpha}g^{\bar\alpha\beta}|\Omega\rangle$$

In other words, up to a sign, it is given by the divergence. This fact persists to higher $(0,p)$-forms. Indeed, the next exercise asks you to show that $Q^* = \bar\partial^*$, the adjoint of $\bar\partial$ under the Hodge metric.

Exercise 4.30 (Q^* **is** $\bar\partial^*$)

Show that under the isomorphism (4.25), the quantisation of Q^* given by (4.27) agrees with $\bar\partial^* = -\star\partial\star$, the adjoint of $\bar\partial$ under the Hodge metric.

(Hint: Compare with equation (4.24).)

Finally, we quantise the hamiltonian by demanding that the $N = 4$ supersymmetry algebra be preserved quantum-mechanically. In other words, and taking into account equation (4.26), we quantise the hamiltonian as follows:

$$H_{\text{eff}} \mapsto 2(QQ^* + Q^*Q)$$

Under the identification $Q \leftrightarrow \bar\partial$ and $Q^* \leftrightarrow \bar\partial^*$, the quantum effective hamiltonian agrees with twice the Dolbeault laplacian $\bar\Box$ or-since $\mathcal{M}_k$ is (hyper)Kähler-with the Hodge laplacian $\triangle$.

This result doesn't just provide a beautiful geometric interpretation of the effective quantum theory, but also allows us to use geometric information to derive physical results. For example, the ground states of the theory will be in one-to-one correspondence with (square-integrable) harmonic $(0,p)$-forms. This sort of reasoning will play a crucial role in the test of Montonen-Olive duality in $N = 4$ supersymmetric Yang-Mills theory, which shall be the focus of the next chapter.

Chapter 5

11 The Effective Action for $N = 4$ Supersymmetric Yang-Mills

In the previous chapter we found that the low energy effective action for the collective coordinates of $N = 2$ supersymmetric Yang-Mills was given by supersymmetric quantum mechanics on the moduli space of BPS-monopoles. In this chapter we will do the same for $N = 4$ super Yang-Mills. As we saw when we discussed that theory in Chapter 2, N=4 super Yang-Mills is a prime candidate to exhibit Montonen-Olive duality: not just are the masses and the structure of the multiplets protected by supersymmetry, but the massive vector bosons and the BPS-monopole belong to isomorphic multiplets. Therefore it would be possible for this theory to afford two inequivalent descriptions: one the standard one and a dual description where the perturbative fields are those in the multiplet containing the BPS-monopole. The structure of this chapter is therefore very similar to that of the previous chapter. We will first count the number of fermionic collective coordinates and will perform the collective coordinate expansion of the action up to second order. The resulting theory is again a $(0 + 1)$ supersymmetric σ-model, this time admitting $N = 8$ supersymmetry due to the fact that there are twice as many fermionic collective coordinates as in the $N = 2$ case. The quantisation of the effective action will proceed along lines similar to the previous chapter: this time the Hilbert space will be isomorphic to square integrable forms on the monopole moduli space, and the hamiltonian will once again be given by the laplacian. This chapter is based on the work of Blum [Blu94].

11.0.1 5.1 Fermionic collective coordinates

We saw in section 3.2 that there are $4k$ bosonic coordinates in the k-monopole sector; and, as we saw in section 4.1, $N = 2$ supersymmetry contributed $2k$ fermionic collective coordinates. In this section we will show that for $N = 4$ supersymmetric Yang-Mills the number of fermionic collective coordinates will double. We can understand this heuristically in a very simple matter. It follows from the discussion in section 4.1, that fermionic collective coordinates are in one-to-one correspondence with zero modes of the Dirac equation in the monopole background. The Dirac operator is the same in both the $N = 2$ and the $N = 4$ theories, but it acts on different types of fermions. In the $N = 2$ theory, ψ was an unconstrained Dirac spinor (it came from a Weyl spinor in six dimensions); whereas in the $N = 4$ theory, it acts on a quartet of Majorana fermions (the dimensional reduction from ten-dimensions of a Majorana-Weyl fermion). But now four Majorana spinors have twice the number of degrees of freedom that an unconstrained Dirac spinor does: $4 \times 4 = 16$ real components to only 4 complex.

To make this argument precise, we need to look in detail at how the Dirac operator breaks up. We start with a monopole background like the one in 2.4.3. Namely, we choose $\mathrm{W}_0 = 0$, $\mathrm{S}_I = a_I \phi$, and $\mathrm{P}_J = b_J \phi$, where a_I and b_J are real numbers satisfying $\sum_I (a_I^2 + b_I^2) = 1$, and where $(\mathrm{W}_i, \boldsymbol{\phi})$ define a k-monopole. Because the scalar fields are collinear, the potential remains at the minimum provided that the fermions satisfy the Dirac equation:

$$\bar{\gamma}_{\underline{i}} D_{\underline{i}} \psi = 0$$

where $\bar{\gamma}_i = \gamma_0 \gamma_i$, and $\bar{\gamma}_4 = -i\gamma_0 \left(a_I \alpha^I + b_J \beta^J \gamma_5\right)$.

Exercise 5.1 (Euclidean Clifford algebra)

Prove that the matrices $\bar{\gamma}_i$ defined above satisfy a euclidean Clifford algebra in four-dimensions.

From Exercise 4.1 we know that the normalisable zero modes of the Dirac operator $\bar{\gamma}_i D_{\underline{i}}$ will have negative chirality with respect to $\bar{\gamma}_5$. But remember that $\boldsymbol{\psi}$ is also Majorana. We now check what chirality with respect to $\bar{\gamma}_5$ and the Majorana condition imply on a spinor.

Recall that our choice (2.32) of ten-dimensional Γ-matrices is such that the Majorana condition in ten-dimensions translates directly into the Majorana condition in four-dimensions. In four-dimensional Minkowski spacetime, there cannot be Majorana-Weyl spinors, but the euclidean $\bar{\gamma}$-matrices preserve the Majorana condition as the next exercise asks you to show.

Exercise 5.2 ($\bar{\gamma}_{\underline{i}}$ and the Majorana condition)

Let $\boldsymbol{\psi}$ be a quartet of Majorana spinors. Prove that $\bar{\gamma}_{\underline{i}}\boldsymbol{\psi}$ is again Majorana. Deduce that one can simultaneously impose the Majorana and $\bar{\gamma}_5$-chirality conditions.

Let us now start by choosing a explicit realisation for the γ-matrices:

$$\gamma_0 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \quad \gamma_i = \begin{pmatrix} i\sigma_i & 0 \\ 0 & -i\sigma_i \end{pmatrix} \quad \gamma_5 = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \tag{5.1}$$

The next exercise asks you to compute the charge conjugation matrix in this realisation.

Exercise 5.3 (The charge conjugation matrix explicitly)

Prove that the charge conjugation matrix C in the above realisation can be chosen to be

$$C = \begin{pmatrix} i\sigma_2 & 0 \\ 0 & -i\sigma_2 \end{pmatrix}$$

(Hint: Using that $C^t = -C$ and that $C\gamma_\mu = -\gamma_\mu^t C$ determine C up to a constant multiple. A possible choice for this multiple is then one for which $C^\dagger C = \mathbb{1}$. That is the choice exhibited above.)

Now let $\boldsymbol{\psi}$ denote a quartet of Majorana spinors which in addition obey $\bar{\gamma}_5\psi = -\boldsymbol{\psi}$. Because $\bar{\gamma}_5 = -\gamma_5\bar{\gamma}_4$ (prove it!), the chirality condition on $\boldsymbol{\psi}$ means that $\bar{\gamma}_4\boldsymbol{\psi} = -\gamma_5\boldsymbol{\psi}$. This means that the euclidean Dirac equation $\bar{\gamma}_i D_i\boldsymbol{\psi} = 0$ becomes $(\bar{\gamma}_i D_i - \gamma_5 D_4)\,\boldsymbol{\psi} = 0$. For the explicit realisation (5.1), this has the virtue that the Dirac operator doesn't see the internal $SU(4)$ indices. Indeed, the Dirac operator is given by:

$$\bar{\gamma}_{\underline{i}} D_{\underline{i}} = \begin{pmatrix} 0 & -\mathcal{D} \\ \mathcal{D} & 0 \end{pmatrix} \otimes \mathbb{1}_4 = (-i\sigma_2 \otimes \mathcal{D}) \otimes \mathbb{1}_4) \tag{5.2}$$

where $\mathbb{1}_4$ is the identity matrix in the internal $SU(4)$ space, and $\mathcal{D} = iD_i\sigma_i + e\phi\mathbb{1}$ is the operator introduced in (3.17).

We are now ready to count the zero modes of the euclidean Dirac operator, by relating them to zero modes of $\mathcal{D}$, which we have already calculated to be $2k$. We first choose the explicit realisation for the α^I and β^J matrices found in Exercise 2.33: $\alpha^I = e_I^+$and $\beta^J = e_J^-$. Next we exploit the internal $SU(4)$ invariance to fix $a_1 = 1$ and all the other a_I and b_J to zero. This means that $\bar{\gamma}_5 = -i\gamma_0\gamma_5 \otimes \alpha^1 = (\sigma_3 \otimes \mathbb{1}) \otimes (\sigma_2 \otimes \mathbb{1})$. From Exercise 5.3 we know that the charge conjugation matrix is given by $C = (\sigma_3 \otimes i\sigma_2) \otimes \mathbb{1}_4$. The next exercise asks you to write down the typical quartet of Majorana spinors $\boldsymbol{\psi}$ which in addition are chiral with respected to $\bar{\gamma}_5$.

*Notice that if $\boldsymbol{\psi}$ had the opposite chirality with respect to $\bar{\gamma}_5$, then it would have been $\mathcal{D}^\dagger$ which would have appeared. This is as expected from the results of Exercise 3.5 and Exercise 4.1

Exercise 5.4 ("Majorana-Weyl" spinors)

Prove that every quartet of Majorana spinors $\boldsymbol{\psi}$ obeying $\bar{\gamma}_5\boldsymbol{\psi} = \pm\boldsymbol{\psi}$ is of the form:

$$\left(\begin{pmatrix}\eta\\ -i\sigma_2\eta^*\end{pmatrix}\begin{pmatrix}\zeta\\ -i\sigma_2\zeta^*\end{pmatrix}\begin{pmatrix}\mp i\eta\\ \pm\sigma_2\eta^*\end{pmatrix}\begin{pmatrix}\mp i\zeta\\ \pm\sigma_2\eta^*\end{pmatrix}\right)$$

where η and ζ are complex two-component spinors.

Finally we count the zero mode of the euclidean Dirac operator $\bar{\gamma}_{\underline{i}}D_{\underline{i}}$.

Exercise 5.5 (Counting zero modes)

Show that ψ is a "Majorana-Weyl" zero mode of the euclidean Dirac operator, in the sense of the previous exercise, if and only if η and ζ are zero modes of $\mathcal{D}$. Therefore if η_a for $a = 1, \ldots, 2k$ is a basis for the normalisable zero modes of $\mathcal{D}$, then the $4k$ spinors

$$\left(\begin{pmatrix}\eta_a\\ -i\sigma_2\eta_a^*\end{pmatrix}\begin{pmatrix}0\\ 0\end{pmatrix}\begin{pmatrix}\mp i\eta_a\\ \pm\sigma_2\eta_a^*\end{pmatrix}\begin{pmatrix}0\\ 0\end{pmatrix}\right) \quad \left(\begin{pmatrix}0\\ 0\end{pmatrix}\begin{pmatrix}\eta_a\\ -i\sigma_2\eta_a^*\end{pmatrix}\begin{pmatrix}0\\ 0\end{pmatrix}\begin{pmatrix}\mp i\eta_a\\ \pm\sigma_2\eta_a^*\end{pmatrix}\right)$$

form a basis for the normalisable zero modes of $\bar{\gamma}_{\underline{i}}D_{\underline{i}}$.

In summary, there are $4k$ fermionic collective coordinates for $N = 4$ supersymmetric Yang-Mills with gauge group $SO(3)$.

12 Monopoles for Arbitrary Gauge Groups for Children by JM Figueroa-O'Farrill

In this chapter we start the study of electromagnetic duality in supersymmetric gauge theories with an arbitrary gauge group. We will be interested in this part of the notes only on $N = 4$ super Yang-Mills. Our principal aim is to frame an analogue of the Montonen-Olive duality conjecture for these theories, to develop testable predictions and then to test them. This will occupy several chapters, but in this one we will start with the analysis of the kind of monopole solutions that can exist in a Yang-Mills-Higgs theory with gauge group G, taken to be a compact, connected Lie group, and a Higgs field with values in the adjoint representation. We will cover the homotopy classification of topologically stable solutions and the generalised Dirac quantisation condition. This chapters borrows quite a lot from the magnificent lectures of Coleman Col77, and from the paper of Goddard, Nuyts and Olive GNO77.

12.1 6.1 Topologically stable solutions

Let G be a compact connected Lie group, and Φ a scalar field taking values in some finite-dimensional representation $\vee$ of G. We will assume that there is a G-invariant potential $V(\Phi)$ which is positive semi-definite and also that $\mathbb{V}$ admits a G-invariant metric. This is necessary in order to write down the kinetic term for Φ in the action. We will let $\mathfrak{g}$ denote the Lie algebra of G. We will fix once and for all an invariant metric on $\mathfrak{g}$. As the next exercise shows, such a metrics always exists.

Exercise

6.1 (Invariant metrics exist)

Prove that there exists a G-invariant metric in the Lie algebra of a compact Lie group.

(Hint: Start with any metric and average over the group with respect to the Haar measure. Does this argument work for any representation?)

We will denote both metrics on $\mathfrak{g}$ and on $\mathbb{V}$ by $(-,-)$, hoping that no confusion will arise. The lagrangian density of the Yang-Mills-Higgs system is given by

$$\mathcal{L} = -\frac{1}{4}\left(\mathrm{G}_{\mu\nu}, \mathrm{G}^{\mu\nu}\right) + \frac{1}{2}\left(D_\mu \Phi, D^\mu \Phi\right) - V(\Phi) \tag{6.1}$$

where

$D_\mu = \partial_\mu \Phi - e\ \mathrm{W}_\mu \cdot \Phi\ \mathrm{G}_{\mu\nu} = \partial_\mu \mathrm{W}_\nu - \partial_\nu \mathrm{W}_\mu - e\left[\,\mathrm{W}_\mu, \mathrm{W}_\nu\right]$,

where W_μ are the $\mathfrak{g}$-valued gauge potentials, and by $\cdot$ we mean the action of $\mathfrak{g}$ on the representation $\mathbb{V}$.

Let M_0 denote the manifold of vacua: those values of Φ for which $V(\Phi) = 0$. Because V is G-invariant, G will map M_0 to M_0; in other words, G stabilises M_0.

We now choose the temporal gauge $\mathrm{W}_0 = 0$. As the next exercise shows, this can always be done and leaves intact the freedom of performing timeindependent gauge transformations.

Exercise 6.2 (The temporal gauge)

Prove that the temporal gauge exists by exhibiting a gauge transformation which makes $\mathrm{W}_0 = 0$. Prove that this gauge is preserved by time-independent gauge transformations.

(Hint: Use path-ordered exponentials.)

In this gauge, the energy density corresponding to the lagrangian density (6.1) is given by

$$\mathcal{H} = \frac{1}{2}\left(\dot{\mathrm{W}}_i, \dot{\mathrm{W}}_i\right) + \frac{1}{2}(\dot{\Phi}, \dot{\Phi}) + \frac{1}{4}\left(\mathrm{G}_{ij}, \mathrm{G}_{ij}\right) + \frac{1}{2}\left(D_i \Phi, D_i \Phi\right) + V(\Phi)$$

where a dot indicates the time derivative, and where repeated indices are summed. The energy is of course the integral over space $\mathbb{R}^3$ of the energy density $\mathcal{H}$, and hence finite-energy configurations must obey the following asymptotic conditions as $|\overrightarrow{\boldsymbol{r}}| \to \infty$:

$\dot{\mathrm{W}}_i = 0$ and $\dot{\Phi} = 0$, whence fields are asymptotically static; $\mathrm{G}_{ij} = 0$ faster than $O(1/r)$; $D_i \Phi = 0$ faster than $O(1/r)$; and $V(\Phi) = 0$.

In particular this last condition says that Φ defines a map from the asymptotic 2 -sphere $S_\infty^2 \subset \mathbb{R}^3$ to the manifold of vacua M_0. Because M_0 is stabilised by G, it will be foliated by orbits of G. For example, in $G = SO(3)$ and the Higgs is in the adjoint, the leaves of the foliation of $SO(3)$ in $\mathbb{R}^3$ are the round spheres centred at the origin, with a "singular" orbit corresponding to the sphere of zero size. A priori there is no reason to expect that the mapping $S_\infty^2 \to M_0$ defined by the asymptotics of the Higgs field should lie on only one of the orbits, but because $D_i \Phi = 0$ in this limit, this is actually the case. The proof is left as an exercise.

Exercise

6.3 ($\Phi\left(S_\infty^2\right) \subset M_0$ lies in a single orbit)

Prove that the image of S_∞^2 lies in a single orbit in M_0.

(Hint: Integrate the equation $D_i \Phi = 0$ on S_∞^2.)

A sufficient-but as shown by Coleman [Col77]) not necessary-condition for a finite-energy configuration to be non-dissipative is that it should be topologically stable. As explained in section 1.2.2, a way to guarantee the topological stability of a field configuration is for the map $\Phi : S_\infty^2 \to M_0$ to belong to a nontrivial homotopy class. It therefore behoves us to study the homotopy classes of maps from the asymptotic two-sphere S_∞^2 to the manifold of vacua M_0 or, more precisely, to the orbit to which $\Phi\left(S_\infty^2\right)$ belongs. To this effect we find it useful to set down some basic notions about homotopy groups. Readers familiar with this material can easily skip the next section.

12.1.1 6.1.1 Some elements of homotopy

This section contains a brief review of homotopy theory. Homotopy theory is the study of continuous change, and is particularly concerned with the determination of quantities which are impervious to such changes. Roughly speaking a homotopy is a continuous deformation parametrised by the unit interval $I=[0,1]$. We will have to be a little bit more precise than this in what follows, but we will avoid getting too technical. In particular, we will not give many proofs. Luckily for us, the aspects of homotopy theory that we will need in these notes can be understood quite intuitively. Proofs can be given but they rely quite a bit on point-set topology. Since that is not the main point of these notes, we simply point the reader who wishes to look at the proofs of the statements made in this section to the old but still excellent book by Steenrod Ste51.

The useful objects in homotopy theory are not just topological spaces, but spaces with a privileged point called the basepoint. A map between two
such spaces is understood to be a continuous function which sends basepoint to basepoint. Let X and Y be two topological spaces with basepoints x_0 and y_0 respectively, and let f_0 and f_1 be two continuous functions $X \rightarrow Y$ taking x_0 to y_0. We say that these two functions are homotopic if there exists a family of functions parametrised by the interval which interpolates continuously between them. More precisely, f_0 is homotopic to f_1 (written $f_0 \simeq f_1$) if there exists a continuous function $F: X \times I \rightarrow Y$, such that for all $x \in X, F(x, 0)=f_0(x)$ and $F(x, 1)=f_1(x)$ and such that for all $t \in I$, $F\left(x_0, t\right)=y_0$. This last condition says that the homotopy is relative to the basepoint.

The fundamental group

A good example with which to visualise these definitions is to take X to be the circle. We can think of the circle as the unit interval with endpoints identified. Then a map from the circle to Y as a map $f: I \rightarrow Y$ with $f(0)=f(1)=y_0$. That is, a continuous loop based at y_0. Then two such loops are homotopic if they can be continuously deformed to each other through loops which are based at y_0.

The set of homotopy equivalence classes of maps $f: X \rightarrow Y$ with $f\left(x_0\right)=y_0$ is written $\left[X, x_0 ; Y, y_0\right]$. In the special case above where X is the circle, the set of homotopy equivalence classes is written $\pi_1\left(Y, y_0\right)$. But $\pi_1\left(Y, y_0\right)$ is more than just a set: indeed, based loops can be composed. Given two loops f_1 and f_2 based at y_0, we can form a third loop $f_1 * f_2$ by simply going first along f_1 and then along f_2 at twice the speed. In other words,

$$\left(f_1 * f_2\right)(t)= \begin{cases}f_1(2 t) & \text { for } t \in\left[0, \frac{1}{2}\right] \\ f_2(2 t-1) & \text { for } t \in\left[\frac{1}{2}, 1\right]\end{cases}$$

Notice that composition is not just defined for loops but also for paths, provided that the first path ends where the second begins. As the following exercise shows, composition of based loops induces a well-defined operation on homotopy classes, which makes $\pi_1\left(Y, y_0\right)$ into a group.

Exercise 6.4 ($\pi_1\left(Y, y_0\right)$ **is a group**)

In this exercise we prove that $\pi_1\left(Y, y_0\right)$ is a group, with group multiplication given by composition of loops. The proof consists of several steps which are all very easy. The idea is to first prove that $*$ makes sense in $\pi_1\left(Y, y_0\right)$ and then that $*$ on loops satisfies all the properties of a group up to homotopy. This means that in $\pi_1\left(Y, y_0\right)$ they are satisfied exactly.

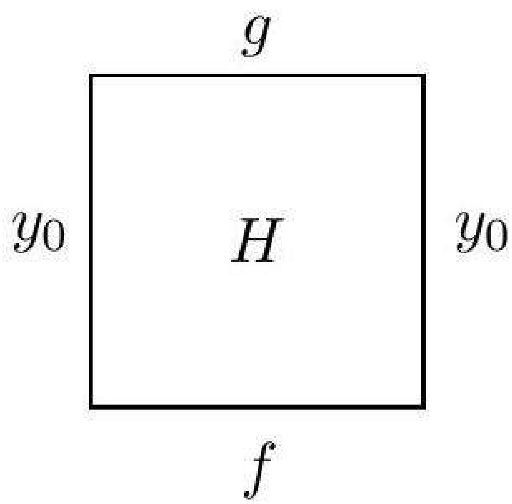
is well-defined in homotopy. Prove that if $f_0 \simeq f_1$ and $g_0 \simeq g_1$ are loops, then $f_0 * g_0 \simeq f_1 * g_1$. (This allows us to work with loops, knowing that up
to homotopy it doesn't really matter which loop we choose to represent its homotopy class.)

is associative in $\pi_1(Y, y_0)$. Prove that if f_1, f_2, and f_3 are loops then $(f_1 * f_2) * f_3 \simeq f_1 * (f_2 * f_3)$. In other words, $*$ is associative up to homotopy.

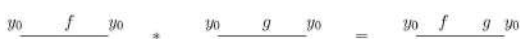
$\pi_1(Y, y_0)$ has an identity. Prove that the constant loop sending all the circle to y_0 is an identity for $*$ up to homotopy; that is, if k denotes the constant loop, then $k * f \simeq f * k \simeq f$ for any loop f. Inverses exist. Let f be a loop, and let $\bar{f}$ denote the loop obtained by following f backwards in time: $\bar{f}(t) = f(1-t)$. Prove that $f * \bar{f} \simeq \bar{f} * f \simeq k$, where k is the constant loop.

(Hint: It may be convenient to devise a pictorial way to denote loops and homotopies. For instance a loop f based at y_0 can be depicted as a unit interval whose endpoints are marked y_0:

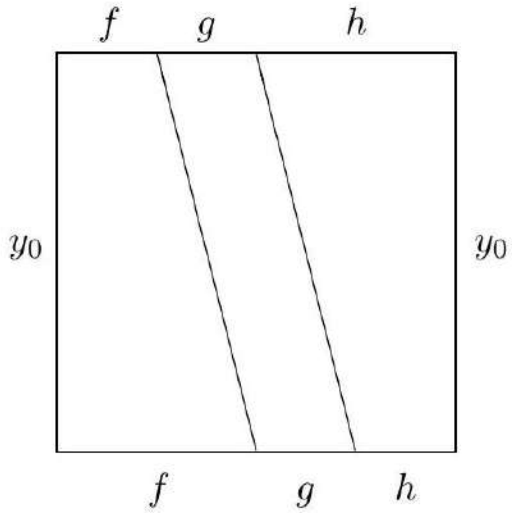
Similarly if f and g are two such loops, a homotopy H between them can be depicted as a square whose left and right edges are marked y_0 and whose top and bottom edges correspond to f and g:



Composition of loops can then be depicted simply as pasting the intervals together by their endpoints and contracting (reparametrising time) so that the resulting interval has again unit length. The following picture illustrates this:



In this language, the group properties become almost self-evident. For example, the associativity property of $*$ simply becomes



and similarly for the other axioms.)

The group $\pi_1(Y, y_0)$ is known as the first homotopy group of the pointed space (Y, y_0). If Y is path-connected, so that any two points in Y can be joined by a continuous path, then the first homotopy group does not depend (up to isomorphism) on the basepoint. This fact has a simple proof which we leave to the next exercise. Incidentally, the condition of connectedness and path-connectedness are not equivalent, but they do agree for manifolds, and hence for all the spaces we will be considering in these notes.

Exercise 6.5 ($\pi_1(Y, y_0) \cong \pi_1(Y, y_1)$)

Let Y be path-connected and y_0 and y_1 be two points in Y. Fix a path $\gamma : I \to Y$ with $\gamma(0) = y_0$ and $\gamma(1) = y_1$. Because Y is path-connected, γ exists. We can use this path to turn any loop f_1 based at y_1 into a loop based on y_0: one simply composes $\bar{\gamma} * f_1 * \gamma$, where $\bar{\gamma}(t) = \gamma(1-t)$. Prove that this defines a group isomorphism $\pi_1(Y, y_1) \cong \pi_1(Y, y_0)$.

Therefore when Y is connected, it makes sense to talk about $\pi_1(Y)$ without reference to a basepoint. This group is called the fundamental group of Y. If this group is trivial, so that all loops are homotopic to the constant map, then Y is said to be simply-connected.

Notice that the isomorphism in Exercise 6.5 depends on the choice of path γ joining the two basepoints. How does the isomorphism depend on γ? It is easy to show (do it!) that if γ' is any other path which is homotopic to γ with endpoints fixed, then the isomorphisms induced by γ and γ' agree. On the other hand, paths in different homotopy classes generally define different isomorphisms. This can be formulated in a way that shows an action of the fundamental group of the space on itself by conjugation. This will play a role later and we will discuss this further we study the addition of topological charges for largely separated monopoles.

The fundamental groups of some manifolds are well known. Here are some examples.

$\mathbb{R}^n$ is simply-connected for any n. The punctured plane $\mathbb{R}^2\backslash\{0\}$ is no longer simply-connected; in fact, $\pi_1(\mathbb{R}^2\backslash\{0\}) \cong \mathbb{Z}$. The isomorphism is given by the following well-known integral formula from complex analysis. To see notice that $\mathbb{R}^2\backslash\{0\} = \mathbb{C}^\times$is the punctured complex plane. Let γ be a loop in $\mathbb{C}^\times$, and compute the contour integral

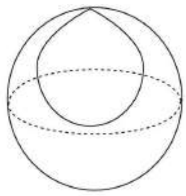
$$\oint_\gamma \frac{1}{2\pi i}\frac{dz}{z}$$

It is well known that this is an integer and is a homotopy invariant of the loop.

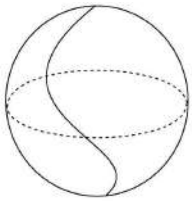
$\pi_1(S^1) \cong \mathbb{Z}$. This is just the above example in disguise. We can think of S^1 as the unit circle in the complex plane $S^1 \subset \mathbb{C}^\times$. Any loop (or homotopy for that matter) in $\mathbb{C}^\times$can be projected onto the unit circle by $t \mapsto \gamma(t) \mapsto \gamma(t)/|\gamma(t)|$. The isomorphism $\pi_1(S^1) \cong \mathbb{Z}$ is known as the degree of the map. It basically counts the number of times one circle winds around another. Puncturing $\mathbb{R}^n$, for $n > 2$, does not alter the fundamental group. In fact, any loop in $\mathbb{R}^n\backslash\{0\}$

is homotopic to a loop on its unit sphere $S^{n-1} \subset \mathbb{R}^n$, again by projecting. But for $n > 2$, it is intuitively clear that any loop on S^{n-1} is homotopic to a constant. For $n = 3$ it is the well-known principle that "you cannot lasso an orange." You cannot lasso higher-dimensional oranges either. You can make a non-simply-connected space out of $\mathbb{R}^3$ by removing a circle (or a knot), say, or an infinite line.

However it is not only by making holes in a space that we can generate nontrivial loops. We can also identify points. For example, if we take the sphere S^n, for $n > 1$, and identify antipodal points, we describe a space which is not simply-connected. The space in question is the space of lines through the origin in $\mathbb{R}^{n+1}$: since every line through the origin will intersect the unit sphere in two antipodal points. We call this space the real projective space $\mathbb{RP}^n$. We can lift a loop in $\mathbb{RP}^n$ up to S^n. This procedure is locally well-defined once we choose a starting point in S^n. There is no further choice and when we are done with the lift, we are either at the starting point or at its antipodal point, since both points map down to the same point in $\mathbb{RP}^n$.



(a)



(b)

Figure 6.1: The two possible lifts to S^n of a loop in $\mathbb{RP}^n$.

In the former case, the loop has lifted to an honest loop in S^n, which is depicted by (a) in Figure 6.1. Since S^n is simply-connected, we can project the homotopy to $\mathbb{RP}^n$ and this gives a homotopy for the original loop. On the other hand, if the loop ends at the antipodal point, as shown in (b) in Figure 6.1, there is clearly no way to deform it to the constant map while keeping endpoints fixed, so it defines a nontrivial loop in $\mathbb{RP}^n$. However notice that the loop obtained by going twice around the loop lifts to an honest loop in the sphere, and is hence trivial. This shows that $\pi_1(\mathbb{RP}^n) \cong \mathbb{Z}_2$. The two-to-one map $\rho : S^n \to \mathbb{RP}^n$ is a covering map, in that it is a local homeomorphism and every point p in $\mathbb{RP}^n$ has a neighbourhood U such that its inverse image by the covering map $\rho^{-1}(U) \subset S^n$ consists of two disconnected neighbourhoods. Because S^n is simply-connected, we say that S^n is the universal covering space of $\mathbb{RP}^n$. All reasonable spaces X (certainly all manifolds and hence all spaces considered in these notes) possess a universal covering space $\tilde{X}$. This space is simply-connected and is such that it admits a free action of the fundamental group of X. In the case of S^n, it admits an action of $\mathbb{Z}_2$, sending a point on the sphere to its antipodal point. The case $n = 3$ is particularly interesting, because it is intimately related with two of our favourite Lie groups: $SU(2)$ and $SO(3)$.

Exercise 6.6 ($SU(2)$ and $SO(3)$)

Prove that the Lie group $SU(2)$ of 2×2 special unitary matrices is parametrised by a three-sphere S^3 and that the group $SO(3)$ of 3×3 special orthogonal matrices is parametrised by the real projective space $\mathbb{RP}^3$. Prove that there is a group homomorphism $SU(2) \to SO(3)$ which sends both $\mathbb{1}$ and $-\mathbb{1}$ in $SU(2)$ to $\mathbb{1}$ in $SO(3)$. Notice that $\mathbb{1}$ and $-\mathbb{1}$ generate the centre of $SU(2)$, which is isomorphic to Z_2. Hence $SU(2)$ is the universal covering group of $SO(3)$.

This situation persists for other Lie groups. Every semisimple compact

Lie group G has a universal covering group $\tilde{G}$, sharing the same Lie algebra g. The fundamental group $\pi_1(G)$ is naturally identified with a subgroup of the centre of $\tilde{G}$. We will be able to compute $\pi_1(G)$ by comparing the finite-dimensional irreducible representations of G with those of $\tilde{G}$, which are those of $\mathfrak{g}$. For example, not every irreducible representation of $SU(2)$ is a representation of $SO(3)$: only those with integer spin. Since representations of $SU(2)$ can have integer or half-integer spin, this means that $SU(2)$ has twice as many irreducible representations as $SO(3)$-which is precisely the order of $\pi_1(SO(3)) \cong \mathbb{Z}_2$. This is no accident, as we will see later on.

Before we abandon the subject of the fundamental group, we mention one last fact. Notice that all the fundamental groups that we have discussed so far are abelian. This is not always the case. In fact, the fundamental group of any compact Riemann surface of genus $g > 1$ is non-abelian. However, there is an important class of manifolds for which the fundamental group is abelian.

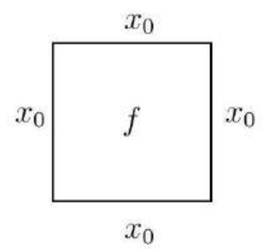
Exercise 6.7 ($\pi_1(G)$ **is abelian)**

Let G be a connected Lie group. Prove that $\pi_1(G)$ is abelian.

(Hint: In a Lie group there are two ways to compose loops. We can use the loop composition * defined above, or we can use pointwise group multiplication, provided that the loops are based at the identity. Indeed, if f and g are loops in G based at the identity, one can define $(f \bullet g)(t) = f(t)g(t)$. Prove that $f * g \simeq f \bullet g$, so that we can use group multiplication to define the multiplication in the fundamental group. Use this to write down a homotopy between $f * g$ and $g * f$, for any two loops f and g in G.)

Higher homotopy groups

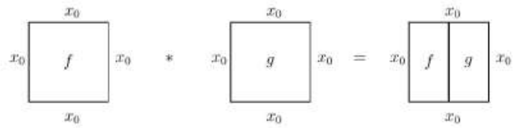
The fundamental group has higher dimensional analogues obtained by substituting the circle by a sphere. Just like we could think of the circle as the interval I with edges identified, we can think of the n-sphere as the multiinterval I^n with its boundary ∂I^n identified to one point. There is a rich theory for all n, but we will only need $n = 2$ in these notes, so we will concentrate mainly on this case. Any map $S^2 \to X$ can be thought of as a map $I^2 \to X$ which sends the boundary ∂I^2 to the basepoint $x_0 \in X$. Just as in Exercise 6.4, we choose to depict such a map f as a rectangle I^2 with the basepoint x_0 along the edges to remind us that x_0 is where ∂I^2 gets mapped to:



We will denote the homotopy classes of such maps by $\pi_2(X, x_0)$. Just like for π_1, we can turn this space into a group. We first discuss composition. Two maps $S^2 \to X$ can be composed by adjoining the squares, just like we did for loops. However in this case there seems to be an ambiguity: we can adjoin the squares horizontally or vertically. We will see that there is indeed no such ambiguity, but for the present we choose to resolve it by composing them horizontally. In other words, if f and g are two maps $I^2 \to X$, we define the composition $f * g$ by

$$(f * g)(t_1, t_2) = \begin{cases} f(2t_1, t_2) & \text{for } t_1 \in \left[0, \frac{1}{2}\right] \\ g(2t_1 - 1, t_2) & \text{for } t_1 \in \left[\frac{1}{2}, 1\right] \end{cases}$$

Clearly the resulting map is continuous, since the boundary conditions agree: $f(1, t_2) = g(0, t_2) = x_0$ for all t_2. Pictorially this composition corresponds to the following diagram:



Just like we did for $\pi_1(X, x_0)$, we can show that $\pi_2(X, x_0)$ is a group.

Exercise 6.8 ($\pi_2(X, x_0)$ **is a group**)

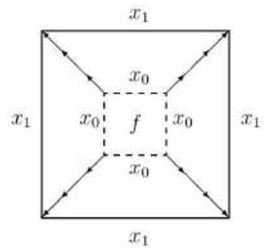
This exercise follows similar steps to Exercise 6.4. Prove the following:

is well-defined in homotopy.

is associative up to homotopy.

$\pi_2(X, x_0)$ has an identity. Prove that the constant map sending all of I^2 to x_0 is an identity for $*$ up to homotopy. Inverses exist. Let f be a map $I^2 \to X$, let $\bar{f}$ denote the map obtained by following f backwards in the first of the two times: $\bar{f}(t_1, t_2) = f(1 - t_1, t_2)$. Prove that $\bar{f}$ is the inverse of f up to homotopy.

If X is path-connected, it follows that $\pi_2(X, x_0)$ doesn't depend on the basepoint (up to isomorphism). Indeed, let f represent a homotopy class in $\pi_2(X, x_0)$. Given any other basepoint $x_1 \in X$, let γ be a path from x_0 to x_1. The following diagram represents a homotopy class in $\pi_2(X, x_1)$:

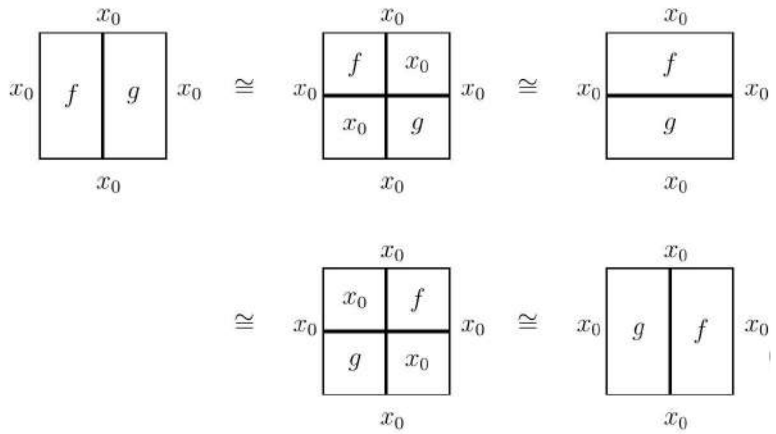


where the arrows represent the path γ going from x_0 to x_1. The next exercise asks you to show that this map is an isomorphism.

Exercise 6.9 ($\pi_2(X, x_0) \cong \pi_2(X, x_1)$)

Prove that the above map is an isomorphism $\pi_2(X, x_0) \cong \pi_2(X, x_1)$; and that the isomorphism only depends on the homotopy class of the path γ used to define it.

Hence for path-connected X it makes sense to talk about $\pi_2(X)$ without reference to the basepoint, provided that we are only interested in its isomorphism class. This group is a higher-dimensional analogue of the fundamental group $\pi_1(X)$. Unlike the fundamental group, $\pi_2(X)$ is always abelian. The
following sequence of homotopies proves this assertion:



This proof also shows that there is no ambiguity in the composition after all. Similarly one proves that also $\pi_k(X)$ for $k > 2$ are abelian.

Finally we discuss some examples of higher homotopy groups. Unlike the fundamental group, for which there are theorems (for instance, the Van Kampen theorem) allowing us to compute $\pi_1(X)$ starting from a decomposition of X into simpler spaces, the computation of the higher homotopy groups is a very difficult problem. For example, not all the homotopy groups of the 2-sphere S^2 are known! Nevertheless, here are some examples of higher homotopy groups:

$\pi_k\left(S^n\right) = 0$ for $k < n$, and $\pi_n\left(S^n\right) \cong \mathbb{Z}$. The first equality is the fact that one cannot k-lasso an n-orange, for $k < n$; whereas the last isomorphism is given by the degree of the map, as was the case for $n = 1$. $\pi_k\left(S^1\right) = 0$ for $k > 1$. This follows because S^1 can be covered by a contractible space: $S^1 \cong \mathbb{R}/\mathbb{Z}$. $\pi_k\left(T^n\right) = 0$ for $k > 1$, where T^n is an n-torus. This follows for the same reason as the above example: $T^n \cong \mathbb{R}^n/\Lambda$, where Λ is some lattice. $\pi_k\left(\Sigma_g\right) = 0$ for $k > 1$, where Σ_g is a compact Riemann surface of genus g. Again, the proof is as above, since Σ_g can be written as the quotient of the Poincaré upper half plane by a Fuchsian group: $\Sigma_g \cong H/\Gamma$. $\pi_2(G) = 0$, for any topological group G. This result of É. Cartan will play a very important role in the next section. The homotopy groups of the spaces determining a fibration $F \rightarrow E \rightarrow B$, where F is the typical fibre, E is the total space, and B is the base, are related by a useful gadget known as the exact homotopy sequence of the fibration:

$$\begin{aligned}
&\cdots \rightarrow \pi_n(B) \rightarrow \pi_{n-1}(F) \rightarrow \pi_{n-1}(E) \rightarrow \pi_{n-1}(B) \rightarrow \cdots \\
&\cdots \rightarrow \pi_2(E) \rightarrow \pi_2(B) \rightarrow \pi_1(F) \rightarrow \pi_1(E) \rightarrow \pi_1(B)
\end{aligned}$$

The "exactness" of this sequence simply means that every arrow is a group homomorphism such that its kernel (the normal subgroup sent to the identity) precisely agrees with the image of the preceding arrow. If the fibration is principal, so that F is a Lie group, then the sequence extends one more term to include a map $\pi_1(B) \rightarrow \pi_0(F)$. Where $\pi_0(F)$ is the set of connected components of the typical fibre. One can define $\pi_0(X)$ in this way for any space X, but for a general $X, \pi_0(X)$ is only a set. It is when $X = G$ is a group, that $\pi_0(G)$ also inherits a group operation. Indeed, $\pi_0(G) \cong G/G_0$, where G_0 is the connected component of the identity. It is in this case that it makes sense to speak of a group homomorphism $\pi_1(B) \rightarrow \pi_0(G)$.

A lot more could be said about higher homotopy groups, but this about covers all that we will need in the sequel.

12.1.2 6.1.2 Homotopy classification of finite-energy configurations

After this brief review of homotopy theory, we return to the problem at hand. Let us fix a basepoint $\vec{r}_0$ in the two-sphere at infinity; for example, we could choose the north pole. Let $\Phi\left(\vec{r}_0\right) = \phi_0 \in M_0$. The G-orbit of ϕ_0 will be the set $G \cdot \phi_0 = \{g \cdot \phi_0 \mid g \in G\} \subset M_0$. If we let $H = H_{\phi_0} \subset G$ denote the stability subgroup of $\phi_0 : H_{\phi_0} = \{h \in G \mid h \cdot \phi_0 = \phi_0\}$, then

$G \cdot \phi_0 \cong G/H$. The asymptotics of the Higgs field define a map from the two-sphere to G/H taking the basepoint to ϕ_0. In other words it defines an element in the second homotopy group $\pi_2(G/H, \phi_0)$.

Because G is connected, this class is gauge-invariant, as the next exercise asks you to show. This shows that the homotopy class is physical.

Exercise 6.10 (Gauge invariance of the homotopy class)

Prove that gauge related Higgs field configurations define asymptotics which are homotopic.

(Hint: Use the fact that G is connected to write down an explicit homotopy between the two configurations.)

As shown in the previous section, $\pi_2(G/H, \phi_0)$ is an abelian group, and because G/H is connected it does not depend (up to isomorphism) on ϕ_0. Because of this fact we will drop the reference to the basepoint when unnecessary. We will now prove that $\pi_2(G/H)$ is isomorphic to the subgroup of $\pi_1(H)$ given by those homotopy classes of loops in H which are null-homotopic in G. From what was said in the previous section, you will immediately recognise this statement as part of the exact homotopy sequence associated to the fibration $H \to G \to G/H$. But rather than appealing to such heavy machinery, we will prove most of this statement here using more pedestrian methods.

We first associate a loop in H with each Φ. Let $\Sigma^\pm$denote an open cover for the asymptotic two-sphere S^2_∞; more concretely, we take their union to be S^2_∞ and their intersection to be a small band around the equator. Since $\Sigma^\pm$are homeomorphic to disks (hence contractible), we can find local gauge transformations $g_\pm : \Sigma^\pm \to G$ such that

$$\Phi(x) = g_\pm(x) \cdot \phi_0 \quad \text{for } x \in \Sigma^\pm$$

Then on the intersection $\Sigma^+ \cap \Sigma^-$, we have

$$g_+(x) \cdot \phi_0 = g_-(x) \cdot \phi_0$$

whence $g_+(x)^{-1} g_-(x) \cdot \phi_0 = \phi_0$, whence $g_+(x)^{-1} g_-(x)$ defines an element of H. Restricting to the equator, we have a continuous map $x \mapsto h(x) = g_+(x)^{-1} g_-(x) \in H$; that is, a loop in H. Because $g_\pm(x)$ are defined only up to right multiplication by H, we can always arrange so that $h(x)$ is the identity for some x in the equator. This way h defines an element in $\pi_1(H, \mathbb{1})$. Furthermore, this loop is trivial in G. Indeed, since G is path-connected, there exist paths $t \mapsto g_\pm(t, x)$ from the identity to $g_\pm(x)$. Defining $h(t, x) = g_+(t, x)^{-1} g_-(t, x)$ provides a homotopy in G from the identity to the loop $h(x)$.

Alternatively we can understand this in a more explicit way. A map $\Phi : S^2_\infty \to G/H$ can be thought of as a loop of loops: In the above picture, there is a family parametrised by the interval $s \in [0, 1]$ of loops based at ϕ_0 and each loop in the family is in turn parametrised by the interval $t \in [0, 1]$, with the condition that the initial and final loops are trivial. Therefore we can redraw Figure 6.2 as a map from the square to G/H where the edges are mapped to ϕ_0:

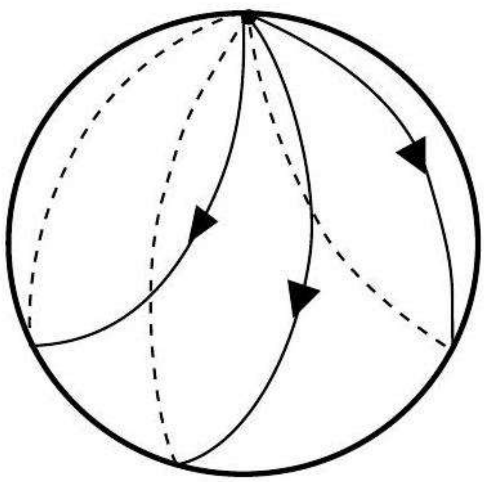
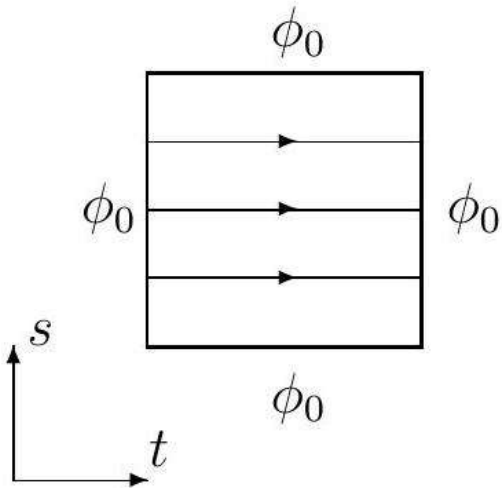


Figure 6.2: A map $S^2 \to G/H$ as a loop of loops.



where the three horizontal lines are precisely the three loops depicted above.

Consider now a fixed loop, that is, a fixed value of s. Because $D_i\Phi = 0$ on $S^2\infty$, we can solve for Φ.

Exercise 6.11 (Solving for Φ)

Prove that for a fixed $s, \Phi(s,t)$ given by

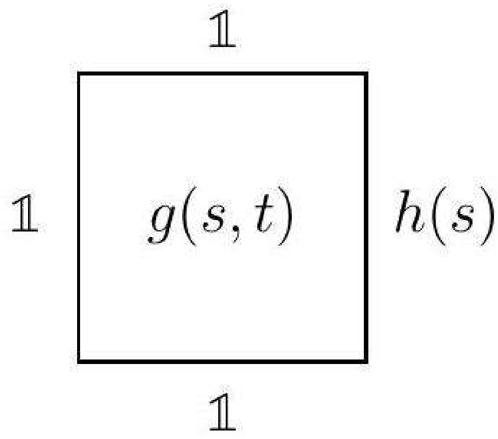
$$\Phi(s,t) = P\exp\left(e\int_0^t dt'\mathbf{W}_i\left(s,t'\right)\frac{\partial x^i}{\partial t'}\right)\cdot\phi_0 \tag{6.3}$$

is a solution of $D_i\Phi$ with the boundary conditions $\Phi(s,0) = \phi_0$. Here s,t are coordinates for S^2_∞ and $x(s,t)$ are the coordinates in $\mathbb{R}^3$. Similarly $\mathrm{W}_i(s,t)$ is short for $\mathrm{W}_i(x(s,t))$.

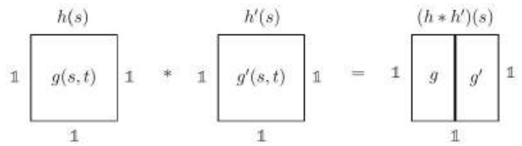
Let $g(s,t) \in G$ be the group element defined by $\Phi(s,t) = g(s,t)\cdot\phi_0$ in (6.3). From its definition it follows that $g(s,0) = 1$ and since $\frac{\partial x^i}{\partial t} = 0$ at $s = 0$ and $s = 1$, it follows that $g(0,t) = g(1,t) = 1$. How about $g(s,1)$? Because $\Phi(s,1) = \phi_0, g(s,1)\cdot\phi_0 = \phi_0$, whence $g(s,1) = h(s)$ is in H. Since $h(0) = h(1) = 1$, as s varies, $h(s)$ defines a loop in H:

$$h(s) \equiv P\exp\left(e\int_0^1 dt\mathbf{W}_i(s,t)\frac{\partial x^i}{\partial t}\right) \tag{6.4}$$

Moreover, this loop in H is trivial in G, the homotopy being given by $g(s,t)$ itself, as the following figure shows:



This approach has the added benefit that it is very easy to see that the map $\pi_2(G/H) \to \pi_1(H)$ sending the class of $\Phi(s,t)$ to the class of $h(s)$ is a group homomorphism. Indeed rotating the squares, we have the composition:

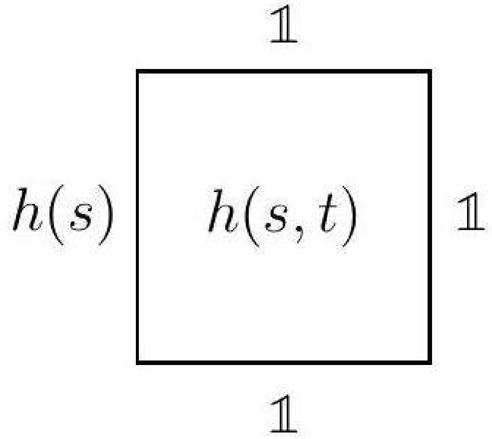


from where we see that if $\Phi \mapsto h$ and $\Phi' \mapsto h'$, then $\Phi * \Phi' \mapsto h * h'$. In summary we have a map $\pi_2(G/H) \to \pi_1(H)$, sending $\Phi(s,t)$ to $h(s)$, which is a homomorphism of groups. Notice that both groups are abelian by Exercise 6.7 and equation (6.2). Also, because any loop in H can be thought of as a loop in G, we have a natural map $\pi_1(H) \to \pi_1(G)$, which is also a homomorphism of abelian groups. We can thus compose these two maps, and what we have shown is that the composition:

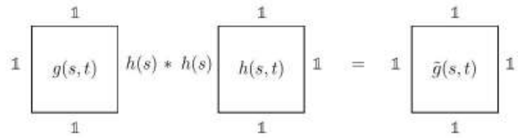
$$\pi_2(G/H) \to \pi_1(H) \to \pi_1(G) \quad \text{is zero.} \tag{6.5}$$

Conversely it follows readily from what we said above that any loop in H which is trivial in G of necessity comes from a map Φ in $\pi_2(G/H)$. Indeed, if $h(s)$ is a loop in H which is null-homotopic in G, let $g(s,t)$ denote the homotopy in G. We can then define $\Phi'(s,t) = g(s,t) \cdot \phi_0$. It is not hard to convince oneself that this $\Phi'(s,t)$ gives rise to a loop in H which is homotopic to the one from which we obtained it.

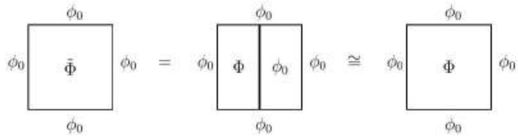
In other words, we have proven that the above sequence (6.5) is exact at $\pi_1(H)$, since the kernel of the arrow leaving $\pi_1(H)$ coincides with the image
of the arrow entering $\pi_1(H)$. We are still not done, though: for we still have to show that the map $\pi_2(G/H) \to \pi_1(H)$ is one-to-one; that is, that if the loop $h(s)$ in H defined by $\Phi(s,t)$ is null-homotopic in H, then the map $\Phi(s,t)$ was already null-homotopic in G/H. Indeed, suppose that there exists a homotopy in H interpolating between $h(s)$ and the constant loop based at the identity; that is, that there exists a map $h(s,t)$:



Then we can compose this with $g(s,t)$ as follows:



Now define $\tilde{\Phi}(s,t) = \tilde{g}(s,t) \cdot \phi_0$. Because $h(s,t) \in H$, this map is homotopic to $\Phi(s,t)$. In fact, acting on ϕ_0 with the above maps we find:



But now notice that $\tilde{g}(s,t)$ defines an element in $\pi_2(G)$. It is now that we must invoke the result of E. Cartan mentioned in the previous section, that
$\pi_2(G) = 0$]] This means that there exists a homotopy $H(s,t,u)$ interpolating continuously between $\tilde{g}(s,t)$ and $\mathbb{1}$. Acting on ϕ_0, we see that $H(s,t,u) \cdot \phi_0$ provides the desired homotopy between $\Phi(s,t)$ and the constant map ϕ_0.

In summary we have proven that

$$\pi_2(G/H) \cong \ker\left(\pi_1(H) \to \pi_1(G)\right) \tag{6.6}$$

or equivalently the exactness of the sequence:

$$0 \to \pi_2(G/H) \to \pi_1(H) \to \pi_1(G)$$

In particular, if G is simply-connected, as is often the case, then every loop in G is null-homotopic, and we find that $\pi_2(G/H) \cong \pi_1(H)$.

We can illustrate this theorem with a simple example. Suppose that $G = SU(2)$ and $H = U(1)$, then we have that $SU(2)/U(1) \simeq S^2$ and $U(1) \simeq S^1$, and indeed $\pi_2\left(S^2\right) \cong \pi_1\left(S^1\right) \cong \mathbb{Z}$. As an abelian group, $\mathbb{Z}$ is freely generated by 1, hence it suffices to determine where 1 gets sent to under the map. In the above case the map $\pi_2(SU(2)/U(1)) \to \pi_1(U(1))$ sends the generator to the generator since the two groups are isomorphic, SU(2) being simply-connected. On the other hand now consider $G = SO(3)$ and $H = SO(2)$. Again we have that $SO(3)/SO(2) \simeq S^2$ and $SO(2) \simeq S^1$, hence as abstract abelian groups $\pi_2(SO(3)/SO(2))$ and $\pi_1(SO(2))$ are both isomorphic to $\mathbb{Z}$, but the theorem says more. It says that the generator of $\pi_2(SO(3)/SO(2))$ cannot be sent to the generator of $\pi_1(SO(2))$, since the image of $\pi_2(SO(3)/SO(2))$ is not all of $\pi_1(SO(2))$ but only the kernel of the map $\pi_1(SO(2)) \cong \mathbb{Z} \to \mathbb{Z}_2 \cong \pi_1(SO(3))$. This map is simply reduction modulo 2 and its kernel consists of the even integers, that is, the subgroup generated by 2. Hence the generator of $\pi_2(SO(3)/SO(2))$ must get sent to twice the generator of $\pi_1(SO(2))$. This can also be understood more pictorially from the discussion surrounding Figure 6.1 and from Exercise 6.6, and it is a good exercise to do so.

Adding topological charges

We now briefly discuss to what extent we can add the topological charges of distant monopoles. It is physically intuitive, despite the fact that the equations describing monopoles are nonlinear, that one should be able to patch distant monopoles together to form a multi-monopole solution. It also seems physically intuitive that the charge of this monopole solution should be given purely in terms of the charges of the constituents and not depend on the details on how the solutions were patched together. We will now see, however that this is not quite right. We will see that when the unbroken gauge group H is disconnected there is an ambiguity in the addition of the monopole charges.

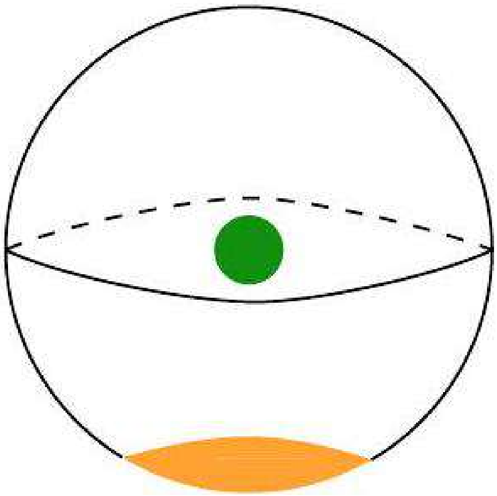


Figure 6.3: monopole.

Consider first a monopole configuration. For convenience we will draw monopole configurations in such a way that the asymptotic sphere at spatial infinity is brought forth to a finite distance from the origin. Physically, we are assuming that the fields reach their asymptotic values to a good approximation in a finite distance. More formally, we are since many physical quantities (e.g., the monopole charge) will turn out to be conformally invariant. The Higgs configuration is gauge related to one which is constant almost everywhere on the asymptotic sphere. This is the so-called unitary gauge. More precisely, the unitary gauge is one where the Higgs field is constant throughout the sphere. It follows that the unitary gauge is singular whenever the Higgs configuration has nontrivial topological charge, since we have seen that regular gauge transformations are homotopies. For our purposes it will be sufficient to consider gauges in which the Higgs is constant almost everywhere on the sphere. Such configurations are depicted in Figure 6.3, where the Higgs is constant everywhere but in the shaded region at the south pole.

Now suppose that we have two monopoles. They are assumed to be so separated that their asymptotic spheres do not intersect. In other words in the space on and outside their two asymptotic spheres, the fields have already attained their asymptotic values. It is as if the monopoles were non-intersecting bubbles in the Higgs vacuum.

*[1] There is to my knowledge no "simple" proof of this fact, but the interested reader is encouraged to go through the one in the book of Bröcker and tom Dieck BtD85.

We can make a 2-monopole solution by patching together two monopoles in the following way. Let us denote by Φ_1 and Φ_2 the Higgs fields for each of the monopoles. It is of course necessary that the image of the Higgs fields lie in the same G-orbit of the manifold of vacua. In a sense different G-orbits are like different superselection sectors. We start by gauge transforming the Higgs fields in such a way that they are equal to ϕ_0 almost everywhere on their asymptotic spheres. We can further orient the monopoles in such a way that the regions in which Φ_1 and Φ_2 are allowed to fluctuate do not intersect. We can do this independently for each monopole because we can always perform gauge transformations which are "compactly supported" in the sense that they are the identity far away from the centre of the monopole. After these gauge transformations, we have a configuration where on and outside the asymptotic spheres (except for the shaded regions in Figure 6.4) the Higgs field is constant and equal to ϕ_0. In particular we have continuity in the Higgs field along the dotted line in Figure 6.4.

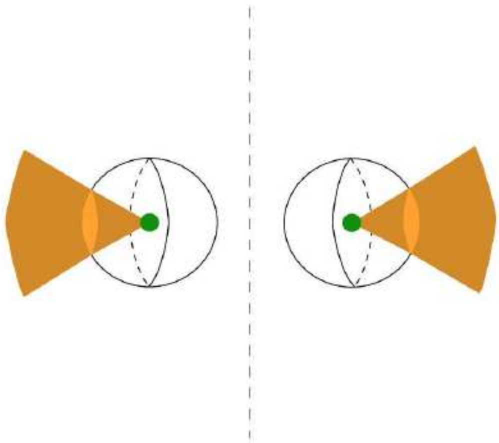


Figure 6.4: Patching two monopoles.

It is clear that the resulting field configuration is continuous, but is this procedure unambiguous? Suppose that we had used a different set of gauge transformations in order to make Φ_1 and Φ_2 equal almost everywhere asymptotically. Once the Higgs are set to ϕ_0 we are still allowed to make an gauge transformation in the stability subgroup H. Would the resulting two-monopole configuration be homotopic (i.e, gauge-equivalent) to the one resulting from our first attempt at patching? Clearly if H is connected, then we can always make a homotopy so that any discontinuity in the H-gauge transformation can be undone. But how about if H is disconnected? In this case there is a potential ambiguity in the patching prescription.

Another way to understand this is as follows. The question boils down to whether $\pi_2(G/H, \phi_1)$ and $\pi_2(G/H, \phi_2)$ can be composed meaningfully. Because G is path-connected, we know that G/H is path-connected, hence $\pi_2(G/H, \phi_1) \cong \pi_2(G/H, \phi_1)$. But remember that this isomorphism depends on the path used to connect ϕ_1 and ϕ_2. If all such paths were homotopicthat is, if G/H were simply-connected - then all such isomorphisms would be one and the same and we could unambiguously compose elements in $\pi_2(G/H, \phi_1)$ and $\pi_2(G/H, \phi_2)$. In other words, if G/H were simply-connected, then we could add without ambiguity the topological charges of each of the monopoles constituting a given two-monopole solution to derive its charge. It turns out, thanks to Theorem 16.11 in [Ste51], that it is enough to check that H be connected. If H is connected, the theorem states, that the isomorphism $\pi_2(G/H, \phi_1) \cong \pi_2(G/H, \phi_2)$ is independent of the path used to go between ϕ_1 and ϕ_2. If H is not connected, however, there is a potential ambiguity. We can patch separated monopoles together, but the topological charge of the resulting two-monopole configuration will not be given simply in terms of the topological charges of its constituents. We need more information: namely the details on how the solutions were put together.

12.2 6.2 The Dirac quantisation condition

In this section we start the analysis of the generalised Dirac quantisation condition obeyed by these monopole solutions. The results in this section are based on the seminal paper of Goddard, Nuyts and Olive GNO77.

We start by considering a monopole in the unitary gauge, where the Higgs field Φ is constant and equal to ϕ_0, say, almost everywhere on the asymptotic sphere. Looking back at Figure 6.3, we have $\Phi = \phi_0$ everywhere on the sphere but on the shaded region around the south pole. Because $D_i\Phi = 0$ everywhere on the sphere, on the part where Φ is constant, this condition becomes $\mathrm{W}_i \cdot \Phi = 0$, hence W_i takes values in the Lie algebra $\mathfrak{h}$ of the stability subgroup $H \subset G$, and so hence so does the path-ordered exponential in equation (6.3).

We will now assume that the field strength G_{ij} has the following asymptotic form:

$$\mathrm{G}_{ij} = \epsilon_{ijk}\frac{x^k}{|x|^3}\frac{\mathrm{Q}(x)}{4\pi} \tag{6.7.7}$$

where the magnetic charge $\mathrm{Q}(x)$ is Lie algebra valued, and hence takes values in $\mathfrak{h}$ almost everywhere on S^2_∞. It may seem surprising at first that the magnetic charge is not constant; but in the presence of a non-abelian gauge symmetry, constancy is not a gauge-invariant statement. The correct nonabelian generalisation of constancy is covariantly constant; and, as the next exercise asks you to show, this is indeed the case.

Exercise 6.12 (The magnetic charge is covariantly constant)

Prove that $D_i\mathrm{Q}(x) = 0$ on the sphere.

(Hint: Analyse the Bianchi identity and the equations of motion on the asymptotic sphere in the Ansatz (6.7) and show that

$$\begin{aligned} \text{Bianchi identity } &\Rightarrow x^k D_k \mathrm{Q}(x) = 0 \\ \text{equation of motion } &\Rightarrow \epsilon_{ijk} x^k D_j \mathrm{Q}(x) = 0 \end{aligned}$$

Deduce that these two equations together imply that $D_k\mathrm{Q} = 0$.)

The quantisation condition will come from demanding that the map $h(s)$ defined in (6.4) does indeed trace a loop in H, so that $h(1) = h(0)$. To make this condition into something amenable to computation we will derive another expression for $h(1)$ in the unitary gauge. This will bring to play a non-abelian version of Stoke's theorem.

Let us define the group element $g(s,t)$ by $\Phi(s,t) = g(s,t)\cdot\phi_0$ in (6.3). Because $D_i\Phi = 0$, it follows that $D_i g(s,t) = 0$, where the covariant derivative is now in the adjoint representation. Define the covariant derivative along
the curves of constant s, by $D_t = \frac{\partial x^i}{\partial t}D_i$, and the covariant derivative along the curves of constant t, by $D_s = \frac{\partial x^i}{\partial s}D_i$. The next exercise asks you to prove the non-abelian Stoke's theorem.

Exercise 6.13 (The non-abelian Stoke's theorem)

The point of this exercise is to prove the following formula.

$$h(s)^{-1}\frac{dh(s)}{ds} = -e\int_0^1 dt g(s,t)^{-1}\mathbf{G}_{ij}g(s,t)\frac{\partial x^i}{\partial t}\frac{\partial x^j}{\partial s} \tag{6.8}$$

We proceed in steps:

Use the fact that $D_t\Phi(s,t) = 0$ implies $D_t g(s,t) = 0$, provided the curve $s =$ constant lies in the region where $\Phi(s,t) = \phi_0$ is constant, and prove that this is equivalent to $D_t \circ g(s,t) = g(s,t)\circ\partial_t$ as operators.

Show that

$$\begin{aligned}\partial_t\left(g(s,t)^{-1}D_s g(s,t)\right) &= g(s,t)^{-1}\left[D_t,D_s\right]g(s,t)\\ &= -e\frac{\partial x^i}{\partial t}\frac{\partial x^j}{\partial s}g(s,t)^{-1}\mathrm{G}_{ij}g(s,t)\end{aligned}$$

Finally integrate the above expression over $t \in [0,1]$ and use the fact that $g(s,t)^{-1}D_s g(s,t)|_{t=0} = 0$ because $g(s,0) = 1$ and $x^i(s,0)$ is constant for all s, and that $g(s,t)^{-1}D_s g(s,t)|_{t=0} = h(s)^{-1}\frac{dh(s)}{ds}$.

With the Ansatz (6.7) for G_{ij} we now have that

$$g(s,t)^{-1}\mathrm{G}_{ij}g(s,t) = \frac{1}{4\pi}\epsilon_{ijk}\frac{x^k}{|x|^3}g(s,t)^{-1}\mathrm{Q}(x)g(s,t)$$

But now notice that because $\mathrm{Q}(x)$ is covariantly constant,

$$g(s,t)^{-1}\mathrm{Q}(x(s,t))g(s,t) = \mathrm{Q}(x(0,0)) \equiv \mathrm{Q} \in \mathfrak{h}$$

In other words,

$$h(s)^{-1}\frac{dh(s)}{ds} = -\frac{e}{4\pi}\mathbf{Q}\int_0^1 dt\epsilon_{ijk}\frac{x^k}{|x|^3}\frac{\partial x^i}{\partial t}\frac{\partial x^j}{\partial s}$$

which can be trivially solved for $h(s)$ to yield:

$$h(s) = \exp\left[-\frac{e}{4\pi}\mathbf{Q}\int_0^s ds'\int_0^1 dt\epsilon_{ijk}\frac{x^k}{|x|^3}\frac{\partial x^i}{\partial t}\frac{\partial x^j}{\partial s'}\right]\cdot h(0)$$

and in particular

$$h(1) = \exp\left[-\frac{e}{4\pi}\mathrm{Q}\int_0^1 ds\int_0^1 dt\epsilon_{ijk}\frac{x^k}{|x|^3}\frac{\partial x^i}{\partial t}\frac{\partial x^j}{\partial s}\right]\cdot h(0) \tag{6.9}$$

where $\exp : \mathfrak{h} \to H$ is the exponential map. Alternatively, if you are more familiar with matrix groups, you can embed H inside a matrix group (every Lie group has a faithful finite-dimensional matrix representation) and then the above differential equation for the matrix $h(s)$ is solved by the above expression, but where exp of a matrix is now defined by its power series.

Let us first compute the above integral. Notice that because the integrand is invariant under rescalings of x we can evaluate it on the unit sphere in $\mathbb{R}^3$; that is, we take $|x| = 1$. We then rewrite it in a more invariant looking form. To this end, it suffices to notice that the integrand is the pull back via the embedding $S^2 \to \mathbb{R}^3, (s,t) \mapsto x^i(s,t)$ of the form $\omega = \frac{1}{2}\epsilon_{ijk}x^k dx^i \wedge dx^j$, whose exterior derivative $d\omega = \frac{1}{2}\epsilon_{ijk}dx^i \wedge dx^j \wedge dx^k$ is precisely 3 times the volume form in $\mathbb{R}^3$ relative to the standard euclidean metric. Therefore using Stoke's theorem and understanding the unit sphere in $\mathbb{R}^3$ as the boundary of the unit ball $S^2 = \partial B^3$, we have that

$$\int_{\partial B^3}\omega = \int_{B^3} d\omega = 3\,\mathrm{vol}\left(B^3\right) = 4\pi$$

We can then rewrite equation (6.9) as follows

$$h(1) = \exp[-e\mathbf{Q}]\cdot h(0)$$

whence the the Dirac quantisation condition $h(1) = h(0)$, becomes

$$\exp e\mathrm{Q} = \mathbb{1} \in H \tag{6.10}$$

Before undertaking a general analysis of this equation, let us make sure that it reduces to the familiar condition (1.6) for $G = SO(3)$ and a nonzero Higgs field in the adjoint representation, as was the case in Chapter 1. The adjoint representation of $SO(3)$ is three-dimensional. Choose the Higgs field to point in the z-direction. The stability subgroup if the subgroup of rotations about the z-axis. It is the $SO(2)$ subgroup consisting of matrices of the form

$$\begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

where θ runs from 0 to 2π. The magnetic charge Q is given by:

$$\mathbf{Q} = g \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

where g is what appears in (1.5). With these definitions, we see that the Dirac quantisation condition (6.10) becomes

$$\exp eg \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos(eg) & \sin(eg) & 0 \\ -\sin(eg) & \cos(eg) & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbb{1}$$

whence $eg \in 2\pi\mathbb{Z}$ in complete agreement with (1.6). We are clearly on the right track. In order to analyse the Dirac quantisation condition properly we will need quite a bit of technology concerning compact Lie groups. This is the purpose of the following section. Readers who already know this material are encouraged to skim through the section for notation.

12.3 6.3 Some facts about compact Lie groups and Lie algebras

In this section we collect without proof those results from the theory of compact Lie groups that are relevant for the analysis of the Dirac quantisation condition. There are many fine books on the subject. A quick and efficient introduction to the main results can be found in the second chapter of Pressley and Segal's book on loop groups [PS86]. A fuller treatment of the parts we will need can be can be found in the book by Adams [Ada69] and also in the more comprehensive book by Bröcker and tom Dieck BtD85]. For the results on Lie algebras we have followed the book by Humphreys Hum72.

12.3.1 6.3.1 Compact Lie groups

Suppose that G is a compact connected Lie group. Any connected abelian subgroup is clearly a torus. Let T be a fixed maximal connected abelian subgroup of G; that is, a maximal torus. Maximal tori obviously exist because any one-parameter subgroup is a connected abelian subgroup. One of the key theorems in the structure of compact Lie groups is the fact that all maximal tori are conjugate in G. This implies, in particular, that the dimension of all maximal tori are the same: it is an invariant of G known as the rank of G. Another way to rephrase this theorem is that any element in G is conjugate
to an element in T, or simply that every group element in G lies in some maximal torus. Generic elements will lie in just one maximal torus: these are called regular elements, whereas there exist also singular elements which lie in more than one.

The prototypical compact connected Lie group is $U(n)$ and many of the results in the theory of compact Lie groups, when restricted to $U(n)$, reduce to well-known facts. For instance, a maximal torus in $U(n)$ can be taken to be the set of diagonal matrices; hence the rank of $U(n)$ is n, and the rank for the $SU(n)$ subgroup is $n-1$. The theorem about maximal tori being conjugate, is simply the fact that any set of commuting unitary matrices can be simultaneously diagonalised by a unitary transformation. The regular elements are those matrices which have distinct eigenvalues.

Let $\mathfrak{g}$ and $\mathfrak{t}$ denote the Lie algebras of G and T, respectively. $\mathfrak{t}$ is a maximal toral subalgebra. A lot can be learned about G by studying the action of T on $\mathfrak{g}$. Because T is abelian, any finite-dimensional complex representation is completely reducible into one-dimensional representations. But $\mathfrak{g}$ is a real representation, so we complexify it first: define $\mathfrak{g}_{\mathbb{C}} = \mathfrak{g} \otimes \mathbb{C}$, and extend the action of G (and hence the one of T) complex-linearly. We can now decompose $\mathfrak{g}_{\mathbb{C}}$ as representations of T as follows:

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{t}_{\mathbb{C}} \oplus \left(\bigoplus_{\alpha} \mathfrak{g}_{\alpha} \right)$$

where $\mathfrak{t}_{\mathbb{C}} = \mathfrak{t} \otimes \mathbb{C}$ is the subspace on which T acts trivially, and $\mathfrak{g}_{\alpha}$ is the subspace of $\mathfrak{g}_{\mathbb{C}}$ defined as follows:

$$v \in \mathfrak{g}_{\alpha} \Leftrightarrow \exp X \cdot v = e^{i\alpha(X)} v$$

where $X \in \mathfrak{t}$ and $\alpha : \mathfrak{t} \to \mathbb{R}$ is a real linear function. The α 's appearing in the above decomposition are known as the (infinitesimal) roots of G. Notice that if α is a root, so is $-\alpha$ since if $v \in \mathfrak{g}_{\alpha}$ its complex conjugate $\bar{v} \in \mathfrak{g}_{-\alpha}$.

The complexified Lie algebra of $U(n)$ is the Lie algebra of all $n \times n$ complex matrices. The roots are given by α_{ij} where $1 \leq i, j \leq n, i \neq j$, and the root subspace corresponding to α_{ij} is spanned by the matrices E_{ij} with a 1 in the (ij) entry and zeroes everywhere else. Acting on the diagonal matrix $X = \operatorname{diag}(x_1, x_2, \ldots, x_n) \in \mathfrak{t}, \alpha_{ij}(X) = x_i - x_j$.

Let us think of $U(1)$ as the group of complex numbers of unit norm. A homomorphism $\chi : T \to U(1)$ is called a character of T. Characters can be multiplied pointwise and indeed form a group called the character group of T. Characters are uniquely determined by their derivatives at the identity. In other words, if χ is a character and $\exp X$ belongs to T, then

$$\chi(\exp X) = e^{iw(X)} \tag{6.11}$$

where $w \in \mathfrak{t}^*$ is an infinitesimal character or a weight. The set of infinitesimal characters define a lattice in $\mathfrak{t}^*$ called the weight lattice of G and denoted $\Lambda_w(G)$. The roots are particular examples of weights, and taking integer linear combinations of the roots, we obtain a sublattice of the weight lattice known as the root lattice and denoted $\Lambda_r(G)$. The root lattice only depends on the Lie algebra, whence Lie groups sharing the same Lie algebra have the same root lattice. On the other hand, the weight lattice identifies the Lie group. If G is semisimple, then both the weight and root lattices span $\mathfrak{t}^*$. It means that the quotient $\Lambda_w(G)/\Lambda_r(G)$ is a finite abelian group. We will see later that it is the fundamental group of the dual group of G.

We can illustrate this with $SU(2)$ and $SO(3)$. The weights of $SU(2)$ form a one-dimensional integral lattice isomorphic to $\mathbb{Z}$, shown below. The weight $m \in \mathbb{Z}$ corresponds to twice the "magnetic quantum number," since for $SU(2)$ the magnetic quantum number, like the spin, can be half-integral. In the case of $SO(3)$ only integral spin representations can occur, hence its weight lattice (shown below with filled circles) corresponds to the sublattice consisting of even integers:

In the semisimple case, $\mathfrak{t}^*$ is called the root space of $\mathfrak{g}$. On the other hand, if G is not semisimple, the roots will only span a subspace of $\mathfrak{t}^*$ which is then the root space of its maximal semisimple subalgebra $[\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{g}$. This shows that the Lie algebra of a compact Lie group is reductive; that is, the direct product of a semisimple Lie algebra and an abelian algebra-namely, its centre.

From the definition of the roots of $U(n)$ above we see that they don't span $\mathfrak{t}^*$, since they annihilate the scalar matrices. This is to be expected since the scalar matrices are in the centre of the Lie algebra $u(n)$ of $U(n)$. The traceless matrices in $u(n)$ span the complement of the scalars matrices and generate the Lie algebra $su(n)$ of $SU(n)$, which is semisimple (in fact, simple). The space spanned by the roots is the root space of $su(n)$.

The root subspaces $\mathfrak{g}_\alpha$ are one-dimensional. Choose vectors $e_\alpha \in \mathfrak{g}_\alpha$ such that $e_{-\alpha} = \bar{e}_\alpha$. Then $e_\alpha, e_{-\alpha}$ and their bracket $h_\alpha = -i\left[e_\alpha, e_{-\alpha}\right] \in \mathfrak{t}$ define an embedding of $sl(2,\mathbb{C})$ in $\mathfrak{g}_{\mathbb{C}}$:

$$[h_\alpha, e_\alpha] = 2ie_\alpha \quad [h_\alpha, e_{-\alpha}] = -2ie_{-\alpha} \quad \text{and} \quad [e_\alpha, e_{-\alpha}] = ih_\alpha$$

Explicitly the embedding is given by $e \mapsto e_\alpha, f \mapsto e_{-\alpha}$ and $h \mapsto h_\alpha$, where

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad f = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad h = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

It therefore follows that $\exp(2\pi h_\alpha) = \mathbb{1}$. It also follows from the representation theory of $sl(2,\mathbb{C})$ that for any root $\beta \in \mathfrak{t}^*, \beta(h_\alpha) \in \mathbb{Z}$ and that, in particular, $\alpha(h_\alpha) = 2$. The h_α are known as coroots and their integer linear combinations span a lattice in $\mathfrak{t}$ called the coroot lattice and denoted $\Lambda_r^\vee(G)$. If Λ is a lattice, then the dual lattice is the set of linear functions $\Lambda \to \mathbb{Z}$ and is denoted Λ^*. This relation is reflexive because $\Lambda^{**} = \Lambda$. In this notation we now see that the coroot lattice is a sublattice of the dual root lattice: $\Lambda_r^\vee(G) \subseteq \Lambda_r(G)^*$. We will see later that the two lattices will agree when G is simply-connected.

Despite the fact that the coroot lattice lives naturally in $\mathfrak{t}$, one often sees in the literature where the coroot lattice is a lattice in $\mathfrak{t}^*$, just like the root and weight lattices. In my opinion this causes more confusion than it is worth, but for the sake of comparison let us see how this goes. In order to identify $\mathfrak{t}$ and $\mathfrak{t}^*$ we need a new piece of information: namely, a metric. We saw in Exercise 6.1 that the Lie algebra of every compact Lie group has an invariant metric, so we will fix one such G-invariant metric $(-,-)$ on $\mathfrak{g}$. Its restriction to T will also be denoted $(-,-)$. Because this metric is invariant and non-degenerate, we can use it to identify $\mathfrak{t}$ and $\mathfrak{t}^*$. In particular there is an element $\alpha^\vee \in \mathfrak{t}^*$ such that for all $X \in \mathfrak{t}, \alpha^\vee(X) = (h_\alpha, X)$. In terms of the root α, we have that $\alpha^\vee = 2\alpha/(\alpha,\alpha)$. We call $\alpha^\vee$ the inverse root corresponding to the root α. Taking integer linear combinations of the coroots, we span the inverse root lattice of G. The weight, root and inverse root lattices are all subsets of $\mathfrak{t}^*$, but notice that whereas the weight and root lattices are intrinsic, the inverse root lattice depends on the chosen metric. In particular the inverse root lattice cannot be meaningfully compared with either the root or weight lattices, since we can scale it at will by rescaling the metric. As we do not wish to advocate its use, we will not give it a symbol, but the reader should beware that sometimes the symbol we use for the coroot lattice is reserved for the inverse root lattice, relative to some "standard" metric.

Using the metric on $\mathfrak{t}^*$ one can measure lengths of roots, and it can be proven that if $\mathfrak{g}$ is simple, then there are at most two lengths of roots, called long and short roots. Simple Lie algebras for which all roots are the same length are called simply-laced. For these simple Lie algebras, we can choose the metric so that $(\alpha,\alpha) = 2$ for all roots. Under this metric, the roots and the inverse roots agree.

12.3.2 6.3.2 The Weyl group

Because the maximal torus T is abelian, conjugation by elements of T is trivial. Moreover generic elements of G will conjugate T to another maximal torus. However there are some elements of G which conjugate T back to T. The largest such subgroup of G is called the normaliser of T and is denoted $N(T)$; that is,

$$N(T)=\left\{h \in G \mid h T h^{-1}=T\right\}$$

It follows from this definition that $N(T)$ is indeed a subgroup of G and that T is contained in $N(T)$ as a normal subgroup. Because $T \subset N(T)$ is a normal subgroup, it follows that $N(T)/T$ is a group. This group is the Weyl group of G relative to the maximal torus T. It is the group of symmetries of the maximal torus. Although it is defined relative to T, the Weyl group $N\left(T^{\prime}\right)/T^{\prime}$ corresponding to any other maximal torus $T^{\prime}$ is conjugate (and hence isomorphic) to $N(T)/T$. Hence it makes sense to talk about the Weyl group W of G, up to isomorphism.

The Weyl group W is a finite group, generated by reflections corresponding to the roots. More precisely, if α is a root, then consider the group element

$$\exp \frac{\pi}{2}\left(e_{\alpha}+e_{-\alpha}\right) \in N(T)$$

The adjoint action of this group element on $\mathfrak{t}$ corresponds to a reflection ρ_{α} on the reflection hyperplane $H_{\alpha} \subset \mathfrak{t}$ defined by $H_{\alpha}=\{X \in \mathfrak{t} \mid \alpha(X)=0\}$. Indeed, one computes that for all $X \in \mathfrak{t}$,

$$\rho_{\alpha}(X)=X-\alpha(X) h_{\alpha}$$

It can be proven that the ρ_{α} generate W.

For example, the Weyl group of $U(n)$ is $\mathfrak{S}_{n}$, the symmetric group in n objects, and it acts by permuting the entries of the diagonal matrices in $\mathfrak{t}$. This is also the Weyl group of $SU(n)$ since the roots of $U(n)$ are the roots of $SU(n)$.

Elements of $\mathfrak{t}$ not belonging to any hyperplane H_{α} are called regular; whereas those who are not regular are called singular. Regular elements fall into connected components called Weyl chambers. The Weyl group permutes the Weyl chambers and no two elements in the same Weyl chamber are Weylrelated. Fix a Weyl chamber C and call it positive. The roots can then be split into two sets, positive and negative roots, according to whether they are positive or negative on C-they cannot be zero, because C does not intersect any hyperplane H_{α}. A positive root is called simple if H_{α} is a wall of C. If G is a simple group of rank ℓ, then there are ℓ simple roots. Every positive root is a linear combination of the simple roots with nonnegative integer coefficients, hence the simple roots generate the root lattice, and their associated reflections generate the Weyl group. The positive Weyl chamber is sometimes called the fundamental Weyl chamber. Its closure
(that is, including the walls) is a fundamental domain for the action of the Weyl group on $\mathfrak{t}$: every point in $\mathfrak{t}$ is Weyl-related to a unique point in the closure of the fundamental Weyl chamber.

In the case of $U(n)$, a choice of fundamental Weyl chamber consists in choosing diagonal matrices whose entries are ordered in a particular way. For instance, we can choose a descending order, in which case the positive roots of $U(n)$ are the α_{ij} with $i<j$. The simple roots are then clearly the $\alpha_{i,i+1}$.

Again using the metric on $\mathfrak{t}$ there is a dual picture of this construction in $\mathfrak{t}^{*}$, where the hyperplanes H_{α} now are defined as the hyperplanes perpendicular to the roots. This picture is independent of the metric since the notion of perpendicularity does not depend on the choice of G-invariant metric on $\mathfrak{t}$. The Weyl group acts on $\mathfrak{t}^{*}$ and it is once again generated by reflections associated to every root. If α and β are roots, we have

$$\rho_{\alpha}(\beta)=\beta-\left(\beta, \alpha^{\vee}\right) \alpha=\beta-(\beta, \alpha) \alpha^{\vee} \tag{6.12}$$

where $\alpha^\vee = 2\alpha/(\alpha, \alpha)$ is the inverse root. Once again the complement of the hyperplanes is divided into connected components called the dual Weyl chambers and one any one of them can be chosen to be the positive or fundamental dual Weyl chamber. The walls of the fundamental dual Weyl chamber are the hyperplanes perpendicular to the simple roots. Once again the closure of the fundamental dual Weyl chamber is a fundamental domain for the action of the Weyl group in t^*.

Those weights of G which lie in the closure of the fundamental dual Weyl chamber are called dominant. We write this set $\Lambda_w^+(G)$. It is a semigroup of $\Lambda_w(W)$; that is, if w_1 and w_2 are dominant, so is their sum $w_1 + w_2$, but there are no inverses. Every irreducible representation of G has a unique highest weight which is dominant. Therefore $\Lambda^+(W)$ is in one-to-one correspondence with the set of finite-dimensional irreducible representations of G.

12.3.3 6.3.3 Root systems and simple Lie algebras

We have seen that the Lie algebra of a compact Lie group is reductive. Because semisimple Lie algebras split in turn into their simple factors, we see that the Lie algebra of a compact Lie group is a direct sum of abelian and simple Lie algebras. This does not mean that any compact Lie group is the direct product of simple Lie groups and a torus, but it turns out that it is covered finitely by a compact Lie group of this type. Hence to a large extent it is enough to study simple Lie groups and abelian Lie groups separately and only at the end put the structures together. Let us therefore assume that G is a simple Lie group. It is a remarkable fact that compact simple Lie groups are essentially classified up to "finite ambiguity" by their root systems. We will begin to describe this process now.

First of all we need to axiomatise the notion of a root system. A subset Φ of a euclidean space E is a root system if the following conditions are obeyed:

Φ is finite, spans E and does not contain the origin;

If $\alpha \in \Phi$, then $-\alpha \in \Phi$ and no other multiples of α are in Φ;

The reflections ρ_α (see (6.12)) leave Φ invariant; and

For all $\alpha, \beta \in \Phi, (\alpha^\vee, \beta) \in \mathbb{Z}$, where $(-, -)$ is the metric in E.

This last condition is extremely restrictive. It essentially says that only very few angles can occur between roots. Indeed, notice that $(\alpha^\vee, \beta) = 2|\beta|/|\alpha| \cos\vartheta$, where ϑ is the angle between α and β. Now, $(\alpha^\vee, \beta)\,(\beta^\vee, \alpha) = 4\cos^2\vartheta$ is a non-negative integer. Taking into account that $(\alpha^\vee, \beta)$ and $(\beta^\vee, \alpha)$ have the same sign we are left with the possibilities listed in Table 6.1, where we have chosen $(\alpha, \alpha) < (\beta, \beta)$ for definiteness and have omitted the trivial case $\alpha = \pm\beta$.

$(\alpha, \beta^\vee)$	$(\alpha^\vee, \beta)$	ϑ	$(\beta, \beta)/(\alpha, \alpha)$
0	0	$\pi/2$	undetermined
1	1	$\pi/3$	1
-1	-1	$2\pi/3$	1
1	2	$\pi/4$	2
-1	-2	$3\pi/4$	2
1	3	$\pi/6$	3
-1	-3	$5\pi/6$	3

Table 6.1: Allowed angles between roots in a root system.

If a root system Φ admits a split $\Phi = \Phi_1 \cup \Phi_2$ into disjoint sets so that every element of Φ_1 is orthogonal to every element of Φ_2, we say that it is reducible; otherwise it is simple. The simple root systems have been classified. There are four infinite families: A_ℓ, for $\ell \geq 1, B_\ell$ and C_ℓ for $\ell \geq 2$, and D_ℓ for $\ell \geq 3$; and five exceptional root systems G_2, F_4, E_6, E_7 and E_8. There are two "accidental" isomorphisms in the above list: $B_2 = C_2$ and $A_3 = D_3$. In all cases, the subscript indicates the rank. The simply-laced root systems are those in the A, D and E series. They are listed in Table 6.2 in a graphical notation that will be explained shortly.

Notice that the above definition of a root system is symmetrical with respect with the interchange $\alpha \leftrightarrow \alpha^\vee$ of a root and the inverse root. In particular this shows that the set $\Phi^\vee \subset \mathrm{E}$ consisting of the dual roots $\alpha^\vee$ is again a root system, and that it is simple if Φ is. In this case $\Phi^\vee$ must be again one of the simple root systems listed above. From the definition of $\alpha^\vee$, it follows that for simply-laced root systems (where all roots have the same length) we can choose the metric so that $\alpha^\vee = \alpha$, hence simply laced root systems are self-dual. More generally, since $(\alpha^\vee, \alpha^\vee) = 4/(\alpha, \alpha)$, long and short roots are interchanged. A quick look at Table 6.2 reveals that G_2 and F_4 are also self-dual, whereas $B_\ell^\vee = C_\ell$, and viceversa, since $\Phi^{\vee\vee} = \Phi$.

Let $\alpha_1, \alpha_2, \ldots, \alpha_\ell$ be a set of simple roots. Let $\alpha_i^\vee = 2\alpha_i / (\alpha_i, \alpha_i)$. Then the inner product $a_{ij} = \left(\alpha_i, \alpha_j^\vee\right)$ is an integer. The set $\{a_{ij}\}$ of all such integers are called the Cartan integers and the matrix (a_{ij}) is known as the Cartan matrix. They are independent of the invariant metric chosen for $\mathfrak{t}$. If two complex simple Lie algebras have the same Cartan matrix, then they are isomorphic. The Cartan matrices of the simple Lie algebras are listed in Hum72, for example.

There is also a graphical notation for root systems. Let Φ be a root system of rank ℓ, and let $\alpha_1, \alpha_2, \ldots, \alpha_\ell$ be a set of simple roots. The Coxeter graph of Φ is the graph consisting of ℓ vertices and such that the i th vertex is joined to the j th vertex by $\left(\alpha_i, \alpha_j^\vee\right)\left(\alpha_i^\vee, \alpha_j\right)$ lines. From Table 6.1, we know that this number can be $0, 1, 2$ or 3. The Coxeter graph can be shown to determine the Weyl group, but does not determine the root system because when two vertices are connected by more than one line, it fails to tell us which of the two vertices corresponds to the shorter root. In other words, the Coxeter graph cannot tell between Φ and $\Phi^\vee$. In order to distinguish them it is necessary to decorate the diagram further: we colour those vertices which corresponds to the short roots, if any are present. The resulting diagram is called the Dynkin diagram. The Dynkin diagrams corresponding to the simple root systems are listed in Table 6.2, the vertex labelled i corresponds to α_i, and the filled vertices corresponds to the short roots.

Reconstructing the group

From the above discussion about compact Lie groups it follows that the root system associated to a compact simple Lie group is simple. Hence it has to be one of the roots systems listed above. This prompts the question of the reconstruction of the group from the root system. It turns out that this is possible up to a finite ambiguity. In a nutshell, given the root system of a compact group, one can obtain a finite covering group of the group in question. In this section we will consider only simple Lie groups.

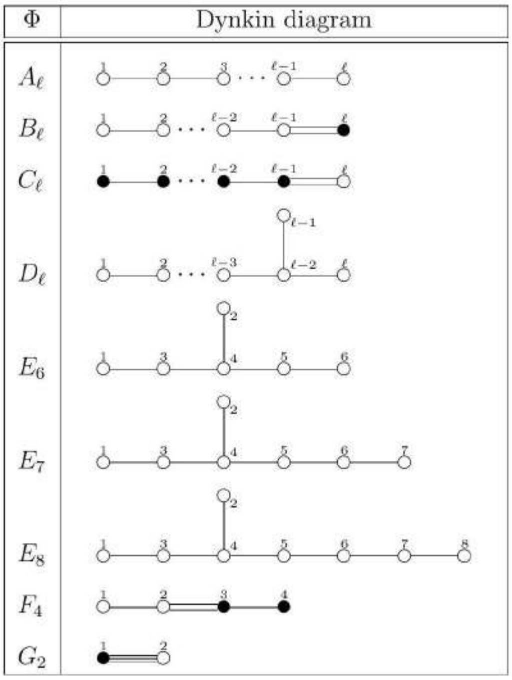


Table 6.2: Dynkin diagrams of the simple root systems.

Given an simple root system Φ of rank ℓ, one can construct a unique simple complex Lie algebra. Associated with each simple root α_i there exist three generators $e_i = e_{\alpha_i}$, $f_i = e_{-\alpha_i}$ and $h_i = h_{\alpha_i}$. These 3ℓ elements, subject to the so-called Serre relations (see Hum72, for example) generate a complex simple Lie algebra $\mathfrak{g}_{\mathbb{C}}$, where $\mathfrak{t}_{\mathbb{C}}$ is spanned by the h_i and whose root system relative to this maximal torus is Φ. As a linear space, the Lie algebra will be generated by ℓ elements h_i spanning the Cartan subalgebra and generators e_α for each root α, whose Lie brackets can be written down as follows:

$$[e_\alpha, e_\beta] = \begin{cases} n_{\alpha\beta} e_{\alpha+\beta} & \text{if } \alpha+\beta \text{ is a root} \\ h_\alpha & \text{if } \beta = -\alpha; \text{ and} \\ 0 & \text{otherwise} \end{cases}$$

where h_α belongs to the Cartan subalgebra. Furthermore, for every h in the Cartan subalgebra, $[h, e_\alpha] = \alpha(h) e_\alpha$. It is possible to choose a basis in which the $n_{\alpha\beta}$ are all nonzero integers. These integers obey $n_{\alpha\beta} = -n_{-\alpha,-\beta}$.

Every complex simple Lie algebra has in general several real forms; that is, real subalgebras. Among these real forms there is a unique compact real form; that is, one for which the Killing form is negative-definite. It is easy to write this form down explicitly. It is generated over the reals by the ih_α, and the combinations

$$\frac{i}{\sqrt{2}}(e_\alpha + e_{-\alpha}) \quad \text{and} \quad \frac{1}{\sqrt{2}}(e_\alpha - e_{-\alpha})$$

It is easy to check that the real linear combinations of these generators close under the Lie bracket. It is also easy to compute the Killing form and see that it is indeed negative-definite.

Next, each compact real form $\mathfrak{g}$ of a simple Lie algebra is the Lie algebra of a unique simply-connected compact simple Lie group $\tilde{G}$, whose maximal torus T is obtained by exponentiating the $\{ih_\alpha\}$, and whose root system relative to T agrees with the one of G. Therefore we have almost come full circle. I say almost, because we are left with a simply-connected compact simple Lie group, even though we started with a compact simple Lie group G which was not assumed to be simply-connected. Therefore we need more information. The information we need is of course the fundamental group $\pi_1(G)$ of G, which is a finite subgroup of the centre of $\tilde{G}$. Remarkably, the centre of $\tilde{G}$ can be read off simply from Lie algebraic data. We review this now.

The centre of $\tilde{G}$

Let $\{\alpha_i\}$ for $i = 1, \ldots, \ell$ denote the simple roots of G relative to a maximal torus T. We define ℓ fundamental weights $\{\lambda_i\}$ by the requirement: $\left(\lambda_i, \alpha_j^\vee\right) = \delta_{ij}$. Alternatively, $\lambda_i\left(h_{\alpha_j}\right) = \delta_{ij}$. In other words, the fundamental weights generate a lattice which is dual to the coroot lattice $\Lambda_r^\vee(G)$. The lattice generated by the fundamental weights is the weight lattice of the simply-connected Lie group $\tilde{G} : \Lambda_w(\tilde{G})$. This lattice contains the root lattice $\Lambda_r(\tilde{G})$ as a sublattice, and the quotient $\Lambda_w(\tilde{G})/\Lambda_r(\tilde{G})$ is a finite abelian group isomorphic to the centre $Z(\tilde{G})$ of $\tilde{G}$. It is sometimes called the fundamental group of the root system, since it is the fundamental group of the adjoint group, the Lie group whose weight lattice agrees with its root lattice.

Let us now explain why $\Lambda_w(\tilde{G})/\Lambda_r(\tilde{G})$ is isomorphic to the centre $Z(\tilde{G})$ of $\tilde{G}$. First of all, notice that $Z(\tilde{G})$ is contained in every maximal torus of $\tilde{G}$. In fact, $Z(\tilde{G})$ is the intersection of all the maximal tori of $\tilde{G}$.

Exercise

6.14 (The centre is the intersection of all the maximal tori)

Prove that the centre $Z(\tilde{G})$ is the intersection of all the maximal tori of $\tilde{G}$.

(Hint: Use that any element can be conjugated to any given maximal torus and the fact that an element of the centre is invariant under conjugation.)

Fix a maximal torus $\tilde{T}$ of $\tilde{G}$ and let $\exp : \mathfrak{t} \to \tilde{T}$ be the restriction of the exponential map to $\mathfrak{t}$. Because $\tilde{T}$ is abelian, the exponential map is a homomorphism of abelian groups. We find it convenient in what follows to include a factor of 2π in the exponential map. We will introduce then the reduced exponential map, denoted $\underline{\exp}$, and defined by $\underline{\exp} X = \exp 2\pi i X$. Clearly exp is also a group homomorphism and, in particular, its kernel is a lattice $\Lambda_I(\tilde{G})$, called the integer lattice of $\tilde{G}$. The reduced exponential map yields an isomorphism $\tilde{T} \cong \mathfrak{t}/\Lambda_I(\tilde{G})$, whence we see that the integer lattice is the lattice of periods of the maximal torus $\tilde{T}$. It follows from (6.11) that h belongs to the integer lattice if and only if for every weight w of $\tilde{G}, w(h) \in \mathbb{Z}$. In other words, the integer lattice and the weight lattice are dual:

$$\Lambda_I(\tilde{G}) = \Lambda_w(\tilde{G})^*$$

Let $\Lambda_Z(\tilde{G}) = \underline{\exp}^{-1} Z(\tilde{G})$ denote those elements of $\mathfrak{t}$ which the reduced exponential map sends to the centre of $\tilde{G}.\Lambda_Z(\tilde{G})$ too is a lattice called the central lattice of $\tilde{G}$, which by definition contains the integer lattice. Because $\underline{\exp}$ is a group homomorphism, we have that $Z(\tilde{G})$ is canonically isomorphic to $\Lambda_Z(\tilde{G})/\Lambda_I(\tilde{G})$. We now claim that the central lattice is the dual of the root lattice.

Exercise 6.15 (The central and root lattices are dual)

Prove that $X \in \mathfrak{t}$ belongs to the central lattice if and only if for every root $\alpha \in \mathfrak{t}^*$, $\alpha(X) \in \mathbb{Z}$.

(Hint: X belongs to the central lattice if and only $\exp 2\pi i X$ is central in $\tilde{G}$, which in turn is equivalent to the statement that for every root $\alpha, \exp(2\pi i X) \exp\left(t e_\alpha\right) = \exp\left(t e_\alpha\right) \exp(2\pi i X)$, for all t. Now use that $[X, e_\alpha] = \alpha(X) e_\alpha$.)

Therefore $Z(\tilde{G}) \cong \Lambda_Z(\tilde{G})/\Lambda_I(\tilde{G}) = \Lambda_r(\tilde{G})^*/\Lambda_w(\tilde{G})^*$, which as the next exercise asks you to show is isomorphic to $\Lambda_w(\mathcal{G})/\Lambda_r(\tilde{G})$ as we had claimed.

Exercise 6.16 (Some facts about lattices)

Let $\Lambda_1 \supseteq \Lambda_2$ be lattices. Prove the following:

Duality reverses inclusions: $\Lambda_1^* \subseteq \Lambda_2^*$; $\Lambda_1/\Lambda_2 \cong \Lambda_2^*/\Lambda_1^*$; and if $\Lambda_3 \subseteq \Lambda_2$ is a third lattice, then

$$\Lambda_1/\Lambda_2 \cong (\Lambda_1/\Lambda_3)\,/\,(\Lambda_2/\Lambda_3)$$

Since the root lattice is contained in the fundamental weight lattice, we can write $\alpha_i = \sum_j M_{ij}\lambda_j$, for some integers M_{ij}. Now, taking the inner product with $\alpha_j^\vee$ and using the definition of the fundamental weights, we find that $\left(\alpha_i, \alpha_j^v\right) = M_{ij}$. In other words, (M_{ij}) is the Cartan matrix. Hence in order to write the fundamental weights in terms of the roots, it is necessary to invert the Cartan matrix. If the Cartan matrix has unit determinant then its inverse has integer entries and the fundamental weights belong to the root lattice. In this case the root lattice and the fundamental weight lattice agree, and $\tilde{G}$ has no centre. In general, the order of the group $Z(\tilde{G})$ is given by the determinant of the Cartan matrix, since this is the only denominator in which we incur in the process of expressing the fundamental weights in terms of the roots. In many cases, the order of $Z(\tilde{G})$ is enough to determine the group uniquely. For example, the Cartan matrices of G_2, F_4 and E_8 have unit determinant, the ones of B_ℓ, C_ℓ and E_7 have determinant 2, and the one of E_6 has determinant 3. Hence the fundamental groups of the roots systems are respectively $1, \mathbb{Z}_2$ and $\mathbb{Z}_3$. In the other cases, the order does not generally determine the group, and one has to work a little harder: the Cartan matrix of A_ℓ has determinant $\ell+1$, from where it follows that if $\ell+1$ is prime, then the fundamental group of A_ℓ is $\mathbb{Z}_{\ell+1}$, since the only finite abelian group of prime order is the cyclic group. (Proof: Take any element not equal to the identity. It generates a cyclic subgroup whose order must divide the order of the group.) In fact, this persists for all ℓ, but this requires an explicit computation. Finally the Cartan matrix of D_ℓ has determinant 4, which again does not determine the fundamental group. It turns out that for ℓ even the fundamental group is $\mathbb{Z}_2 \times \mathbb{Z}_2$, whereas for ℓ odd it is $\mathbb{Z}_4$. A useful mnemonic in this case is to remember that $D_3 = A_3$.

An example: $A_3 = D_3$

Let us in fact work out this example to see how to go about these calculations. Let us consider the root system $A_3 = D_3$ whose simply-connected compact Lie group is $SU(4) \cong \mathrm{Spin}(6)$. We can read off the Cartan matrix from its Dynkin diagram listed in Table 6.2:

$$\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

which has indeed determinant 4. Inverting this matrix we can read off the expression for the fundamental weights in terms of the roots:

$$\begin{aligned} \lambda_1 &= \frac{3}{4}\alpha_1 + \frac{1}{2}\alpha_2 + \frac{1}{4}\alpha_3 \\ \lambda_2 &= \frac{1}{2}\alpha_1 + \alpha_2 + \frac{1}{2}\alpha_3 \\ \lambda_3 &= \frac{1}{4}\alpha_1 + \frac{1}{2}\alpha_2 + \frac{3}{4}\alpha_3 \end{aligned}$$

We can now compute the factor group Λ_w/Λ_r. Its elements are the cosets $0 + \Lambda_r, \lambda_1 + \Lambda_r, \lambda_2 + \Lambda_r$ and $\lambda_3 + \Lambda_r$, which possess the following multiplication table:

	0	λ_1	λ_2	λ_3
0	0	λ_1	λ_2	λ_3
λ_1	λ_1	λ_2	λ_3	0
λ_2	λ_2	λ_3	0	λ_1
λ_3	λ_3	0	λ_1	λ_2

where all entries are understood modulo Λ_r. It follows clearly that this is the cyclic group $\mathbb{Z}_4$.

Another example: D_4

Finally let us work a second example. We pick now one of our favourite root systems: D_4, whose simply-connected compact Lie group is Spin(8). From Table 6.2 we can read off the Cartan matrix:

$$\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & -1 \\ 0 & -1 & 2 & 0 \\ 0 & -1 & 0 & 2 \end{pmatrix}$$

and inverting it, we can read off the expression of the fundamental weights in terms of the roots:

$$\lambda_1 = \alpha_1 + \alpha_2 + \tfrac{1}{2}\alpha_3 + \tfrac{1}{2}\alpha_4 \quad \lambda_2 = \alpha_1 + 2\alpha_2 + \alpha_3 + \alpha_4$$
$$\lambda_3 = \tfrac{1}{2}\alpha_1 + \alpha_2 + \alpha_3 + \tfrac{1}{2}\alpha_4 \quad \lambda_4 = \tfrac{1}{2}\alpha_1 + \alpha_2 + \tfrac{1}{2}\alpha_3 + \alpha_4$$

We can now compute the factor group Λ_w/Λ_r. Because $\lambda_2 \in \Lambda_r$, it has as elements the cosets of $0, \lambda_1, \lambda_3$ and λ_4. The multiplication table for this group can be read off easily:

	0	λ_1	λ_3	λ_4
0	0	λ_1	λ_3	λ_4
λ_1	λ_1	0	λ_4	λ_3
λ_3	λ_3	λ_4	0	λ_1
λ_4	λ_4	λ_3	λ_1	0

where all entries are understood modulo Λ_r. It is clear that this group is $\mathbb{Z}_2 \times \mathbb{Z}_2$. It has three proper $\mathbb{Z}_2$ subgroups, each one generated by one of the cosets $\lambda_1 + \Lambda_r, \lambda_3 + \Lambda_r$ and $\lambda_4 + \Lambda_r$. The representations with highest weights λ_1, λ_3 and λ_4 are all eight-dimensional. They correspond to the vector and the two spinor representations $\Delta_\pm$of Spin(8). Alternatively, they correspond to the three inequivalent embeddings Spin(7) $\subset$ Spin(8). Given any one of these eight-dimensional representations there exists an Spin(7) subgroup of Spin(8) under which the representation remains irreducible and can be identified with the unique spinorial representation of Spin(7). The dihedral group $\mathcal{D}_3$ of automorphisms of the Dynkin diagram is the group of outer automorphisms of Spin(8). It is called the triality group in the physics literature, and it permutes the three inequivalent Spin(7) subgroups and thus the three eight-dimensional representations. In terms of the weights, it permutes λ_1, λ_3 and λ_4.

All the connected compact simple Lie groups

It is now time to summarise what we have learned so far in a table. Table 6.3 lists the simple root systems, their Weyl groups, the associated simple complex Lie algebras, their simply-connected simple compact Lie groups, and their centres. An eternal thorny issue about the notation in Table 6.3: the compact Lie group associated to the root system of type C_ℓ is called in the physics literature $USp(2\ell)$ and in the mathematics literature $Sp(\ell)$. From here until the end of this section, all Lie groups are connected, compact and simple unless otherwise explicitly stated.

We start by associating with every Lie group G a subgroup of the centre of its universal cover $\tilde{G}$ and viceversa. Representations of G are also representations of $\tilde{G}$, whence the weight lattice $\Lambda_w(G)$ is contained in the weight lattice $\Lambda_w(\tilde{G})$. We thus have the following inclusions of lattices in $\mathfrak{t}^*$:

$$\Lambda_r(\tilde{G}) \subseteq \Lambda_w(G) \subseteq \Lambda_w(\tilde{G})$$

Dualising and keeping in mind Exercise 6.16, we have in $\mathfrak{t}$ the following lattices:

$$\Lambda_Z(\tilde{G}) \supseteq \Lambda_w(G)^* \supseteq \Lambda_I(\tilde{G})$$

where we have used that $\Lambda_I(\tilde{G}) = \Lambda_w(\tilde{G})^*$, and that $\Lambda_Z(\tilde{G}) = \Lambda_r(\tilde{G})^*$. Applying the reduced exponential map $\exp : \mathfrak{t} \to \tilde{G}$ to these lattices and remembering that exp is a group homomorphism when restricted to the maximal torus, we find that $\exp \Lambda_w(G)^* = \Gamma_G \subseteq Z(\tilde{G})$ is a subgroup of the centre. The subgroup Γ_G is naturally isomorphic to $\Lambda_w(G)^*/\Lambda_I(\tilde{G})$, since the integral lattice is the kernel of the reduced exponential map. Using the fact that

Φ	W or $\lvert W\rvert$	$\mathfrak{g}_{\mathbb{C}}$	$\tilde{G}$	$Z(\tilde{G})$
A_ℓ	$\mathfrak{S}_{\ell+1}$	$sl(\ell+1,\mathbb{C})$	$SU(\ell+1)$	$\mathbb{Z}_{\ell+1}$
B_ℓ	$(\mathbb{Z}_2)^\ell \rtimes \mathfrak{S}_\ell$	$so(2\ell+1,\mathbb{C})$	$Spin(2\ell+1)$	$\mathbb{Z}_2$
C_ℓ	$(\mathbb{Z}_2)^\ell \rtimes \mathfrak{S}_\ell$	$sp(2\ell,\mathbb{C})$	$USp(2\ell)$	$\mathbb{Z}_2$
D_ℓ	$(\mathbb{Z}_2)^{\ell-1} \rtimes \mathfrak{S}_\ell$	$so(2\ell,\mathbb{C})$	$Spin(2\ell)$	$\{\mathbb{Z}_4$ is odd
$\mathbb{Z}_2 \times \mathbb{Z}_2 \ell$ is even				
G_2	$\mathcal{D}_6$			1
F_4	$2^7 3^2$	F_4	G_2	F_4
E_6	$2^7 3^4 5$	E_6	E_6	1
E_7	$2^{10} 3^4 5 7$	E_7	E_7	$\mathbb{Z}_3$
E_8	$2^{10} 3^5 5^2 7$	E_8	E_8	$\mathbb{Z}_2$

Table 6.3: Simple root systems, their Weyl groups, their complex Lie algebras, compact Lie groups and their centres.

$\Lambda_I(\tilde{G}) = \Lambda_w(\tilde{G})^*$, we have

$$\Gamma_G = \Lambda_w(G)^*/\Lambda_w(\tilde{G})^* \cong \Lambda_w(\tilde{G})/\Lambda_w(G)$$

where we have again used Exercise 6.16. Since G is determined by its weight lattice, this actually tells us that $G \cong \tilde{G}/\Gamma_G$. Since $\tilde{G}$ is simply-connected, this implies that $\pi_1(G) \cong \Gamma_G$.

Conversely, if $\Gamma \subseteq Z(\tilde{G})$ is a subgroup of the centre of $\tilde{G}$. The preimage of Γ via the reduced exponential map $\exp : \mathfrak{t} \to \tilde{G}$, is a sublattice of the central lattice and contains the integer lattice:

$$\Lambda_I(\tilde{G}) \subseteq \Lambda_\Gamma \subseteq \Lambda_Z(\tilde{G}) \tag{6.13}$$

which upon dualising gives in $\mathfrak{t}^*$ the following series of lattices:

$$\Lambda_w(\tilde{G}) \supseteq \Lambda_\Gamma^* \supseteq \Lambda_r(\tilde{G}) \tag{6.14}$$

It is not hard to see that Λ_Γ^* is the weight lattice of the group G defined by $\tilde{G}/\Gamma$.

In summary, there is a one-to-one correspondence between Lie groups with the same universal covering group $\tilde{G}$ and subgroups of the centre $Z(\tilde{G})$; or, equivalently, between Lie groups with the same Lie algebra $\mathfrak{g}$ and lattices Λ containing the root lattice and contained in the fundamental weight lattice. Since the centre $Z(\tilde{G})$ is finite, it has a finite number of subgroups, and hence there are only a finite number of Lie groups covered by the same simplyconnected Lie group. This is what we meant earlier by "finite ambiguity."

Minding Table 6.3, we can now list all the connected compact simple Lie groups.

For the root systems E_8, F_4 and G_2, the centre is trivial, so they are the only groups with that root system. Similarly, the centres of E_6, E_7, B_ℓ and C_ℓ are not trivial but have no proper non-trivial subgroups, hence there are only two groups associated with each of those root systems: the simply-connected group and the adjoint group: E_6 and $E_6/\mathbb{Z}_3, E_7$ and $E_7/\mathbb{Z}_2, \mathrm{Spin}(2\ell+1)$ and $SO(2\ell+1) = \mathrm{Spin}(2\ell+1)/\mathbb{Z}_2$, and $USp(2\ell)$ and $USp(2\ell)/\mathbb{Z}_2$. A similar story holds for A_ℓ with $\ell+1$ prime: there are only two groups with that root system, $SU(\ell+1)$ and $SU(\ell+1)/\mathbb{Z}_{\ell+1}$. For general ℓ, however, the centre of $SU(\ell+1)$ has subgroups corresponding to the divisors of $\ell+1 : \mathbb{Z}_m \subset \mathbb{Z}_{\ell+1}$ if and only if m divides $\ell+1$. So we have a whole hierarchy of groups $SU(\ell+1)/\mathbb{Z}_m$ where m runs over the divisors or $\ell+1$, interpolating between the simply-connected $SU(\ell+1)$ and the adjoint group $SU(\ell+1)/\mathbb{Z}_{\ell+1}$. For the root system $D_{2\ell+1}$, the centre is $\mathbb{Z}_4$ which has a single nontrivial proper subgroup isomorphic to $\mathbb{Z}_2$. Hence there are three groups: $\mathrm{Spin}(4\ell+2), SO(4\ell+2) = \mathrm{Spin}(4\ell+2)/\mathbb{Z}_2$ and $\mathrm{Spin}(4\ell+2)/\mathbb{Z}_4$. Finally, the root systems $D_{2\ell}$ has centre $\mathbb{Z}_2 \times \mathbb{Z}_2$ which has three proper subgroups isomorphic to $\mathbb{Z}_2$. Hence there are five groups in this family: $\mathrm{Spin}(4\ell), SO(4\ell) = \mathrm{Spin}(4\ell)/\mathbb{Z}_2, \mathrm{Spin}(4\ell)/\mathbb{Z}_2', \mathrm{Spin}(4\ell)/Z_2''$, and the adjoint group $\mathrm{Spin}(4\ell)/\left(\mathbb{Z}_2 \times \mathbb{Z}_2\right)$. In the next section we will see that many of these groups are mapped to each other by a duality transformation.

12.3.4 6.3.4 Some simple examples

We illustrate some of the results above with some simple examples: the simple root systems of rank 2: $A_2, B_2 = C_2$ and G_2.

The simple root system A_2

The root system A_2 is defined by the Cartan matrix

$$(A_{ij}) = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

Therefore the simple roots are given in terms of the fundamental weights as follows: $\alpha_1 = 2\lambda_1 - \lambda_2$ and $\alpha_2 = -\lambda_1 + 2\lambda_2$. Inverting these relations we see that $\lambda_1 = \frac{2}{3}\alpha_1 + \frac{1}{3}\alpha_2$ and $\lambda_2 = \frac{1}{3}\alpha_1 + \frac{2}{3}\alpha_2$. This clearly shows that the order of the fundamental group of A_2 is 3, and hence that $\Lambda_w/\Lambda_r \cong \mathbb{Z}_3$. Indeed, notice that this group has as elements the cosets $0+\Lambda_r$ and $\lambda_1 + \Lambda_r$ and

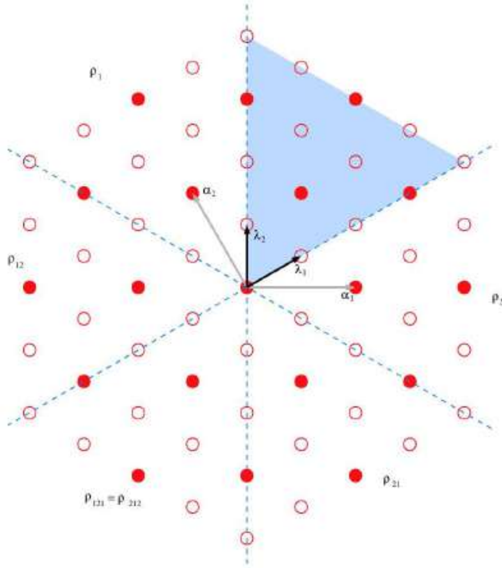


Figure 6.5: The root system A_2.
$\lambda_2 + \Lambda_r$, with multiplication table:

	0	λ_1	λ_2
0	0	λ_1	λ_2
λ_1	λ_1	λ_2	0
λ_2	λ_2	0	λ_1

where all entries are to be understood modulo Λ_r. We can choose a euclidean metric on $\mathbb{R}^2$ and represent these lattices pictorially. This is done in Figure 6.5. which also shows the hyperplanes perpendicular to the roots as dashed lines, and the positive dual Weyl chamber as shaded. The Weyl group is the dihedral group $\mathcal{D}_3 \cong \mathfrak{S}_3$, the symmetries of an equilateral triangle, and it clearly permutes the dual Weyl chambers. Indeed, in Figure 6.5 all chambers but the fundamental are labelled with the element of the Weyl with which it is associated. Since the Weyl group is generated by reflections on the hyperplanes perpendicular to the simple roots, I have chosen to write the Weyl group elements in this way: the notation ρ_i means the reflection ρ_{α_i} and $\rho_{ij\cdots k} = \rho_i\rho_j\cdots\rho_k$. The filled circles defines the root lattice Λ_r and these together with the open circle define the weight lattice Λ_w. The fundamental weights and simple roots are also shown. Λ_w is the weight lattice of the group $SU(3)$, whereas Λ_r is the weight lattice of the adjoint group $SU(3)/\mathbb{Z}_3$.

The simple root systems $B_2 = C_2$

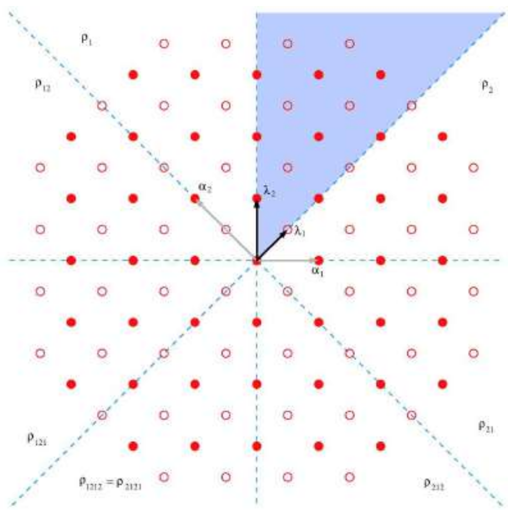


Figure 6.6: The root system B_2.
The root system B_2 is defined by the Cartan matrix

$$(A_{ij}) = \begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix}$$

Therefore the simple roots are given in terms of the fundamental weights as follows: $\alpha_1 = 2\lambda_1 - \lambda_2$ and $\alpha_2 = -2\lambda_1 + 2\lambda_2$. Inverting these relations we see that $\lambda_1 = \alpha_1 + \frac{1}{2}\alpha_2$ and $\lambda_2 = \alpha_1 + \alpha_2$. This clearly shows that the order of the fundamental group of B_2 is 2, and hence that $\Lambda_w/\Lambda_r \cong \mathbb{Z}_2$. Again, one can see this directly: the cosets $0 + \Lambda_r$ and $\lambda_1 + \Lambda_r$ are the elements of the fundamental group with multiplication table

	0	λ_1
0	0	λ_1
λ_1	λ_1	0

where all entries are again to be understood modulo Λ_r. We can choose a euclidean metric on $\mathbb{R}^2$ and represent these lattices pictorially. This is done in Figure 6.6, which also shows the hyperplanes perpendicular to the roots as dashed lines, and the positive dual Weyl chamber as shaded. The Weyl group is isomorphic to the dihedral group $\mathcal{D}_4$ of symmetries of the square, and again the Weyl chambers have been decorated with the corresponding element of the Weyl group. Once again the filled circles define the root lattice Λ_r and these together with the open circles define the weight lattice Λ_w. The fundamental weights and simple roots are also shown. Λ_w is the weight lattice of the group $\mathrm{Spin}(5) \cong USp(4)$, whereas Λ_r is the weight lattice of the group $SO(5)$. The weight λ_1 is the highest weight of the irreducible spinorial representation Δ of $\mathrm{Spin}(5)$ obtained as the unique irreducible representation of the Clifford algebra in five-dimensional euclidean space.

The dual root system C_2 has as Cartan matrix the transpose of the Cartan matrix of B_2. They are of course isomorphic root systems, but the isomorphism interchanges long and short roots: $\alpha_1 \leftrightarrow \alpha_2$. This essentially rotates the root diagram by $\pi/4$, and chooses a different fundamental dual Weyl chamber.

The simple root system G_2

The root system G_2 is defined by the Cartan matrix

$$(A_{ij}) = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}$$

Therefore the simple roots are given in terms of the fundamental weights as follows: $\alpha_1 = 2\lambda_1 - \lambda_2$ and $\alpha_2 = -3\lambda_1 + 2\lambda_2$. Inverting these relations we see that $\lambda_1 = 2\alpha_1 + \alpha_2$ and $\lambda_2 = 3\alpha_1 + 2\alpha_2$. Hence the root and weight lattices agree. We can choose a euclidean metric on $\mathbb{R}^2$ and represent this lattice pictorially. This is done in Figure 6.7, which also shows the hyperplanes perpendicular to the roots as dashed lines, and the positive dual Weyl chamber as shaded. The Weyl group is now $\mathcal{D}_6$, the symmetries of the regular hexagon, and it permutes the Weyl chambers as shown in the figure. Now the open circles define the root/weight lattice Λ. The fundamental weights and simple roots are also shown. Notice that the long roots form a root system of type A_2, indicative of the fact that $SU(3)$ is a maximal subgroup of G_2.

12.3.5 6.4 The magnetic dual of a compact Lie group

We now start to analyse the Dirac quantisation condition (6.10) in more detail. The punch-line is that the Dirac quantisation condition says that the

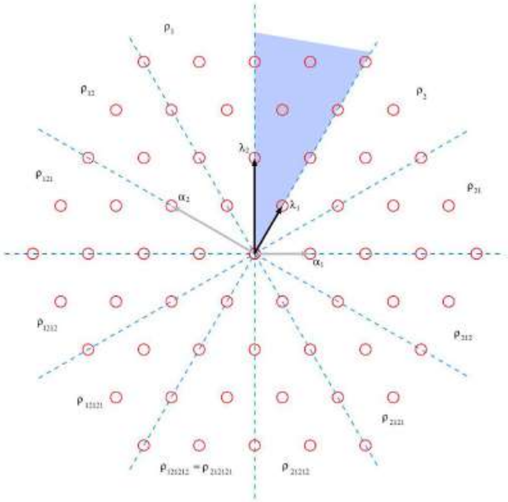


Figure 6.7: The root system G_2.

magnetic charge (suitably normalised) is a dominant weight of a connected compact Lie group $H^\vee$, the (magnetic) dual group of H.

First of all notice that it is irrelevant for these purposes that H be connected, since the image of the exponential map lies in the connected component of the identity. (Proof: If $g = \exp X$, then $g(t) = \exp(1-t)X$ is a path to the identity.) Therefore we will assume from now on that H is connected. It is also compact since it is a closed subgroup of a compact Lie group. So we are in the situation that we have just discussed. Because physics is gauge invariant, we have to identify different charges Q which are gauge related via the unbroken gauge group H. Q belongs to the Lie algebra $\mathfrak{h}$ of H and H acts on its Lie algebra via conjugation. One way to fix this gauge invariance is to to choose a fixed maximal torus T in H, with Lie algebra $\mathfrak{t}$ and use our gauge freedom to conjugate Q to lie in $\mathfrak{t}$. As discussed above, this does not fix the gauge completely, because there will be elements of H which stabilise T; in other words, we have to still take into account the action of the Weyl group. The action of the Weyl group is fixed by choosing Q in the closure of the fundamental Weyl chamber C, since this is a fundamental domain for the action of the Weyl group.

Therefore in the Dirac quantisation condition (6.10), we can take $e\mathbf{Q}$ to lie in the closure $\bar{C}$ of the fundamental Weyl chamber in $\mathfrak{t}$. The exponential map in (6.10) is then the exponential map $\mathfrak{t} \to T$, and (6.10) says that $e\mathrm{Q}/2\pi$ belongs to the integer lattice $\Lambda_I(H)$ of H. We saw above that $\Lambda_I(H) = \Lambda_w(H)^*$, whence the Dirac quantisation condition becomes:

$$e\mathrm{Q}/2\pi \in \Lambda_w(H)^*/W \cong \Lambda_w(H)^* \cap \bar{C}$$

where W is the Weyl group. On the other hand, the integer lattice can be thought of as the weight lattice of a connected compact Lie group $H^\vee$ known as the (magnetic) dual group of H. As we will see below, this group is a quotient of the simply-connected compact simple Lie group whose root system is dual to the root system of H. Now, dual root systems share the same Weyl group. This follows from the fact that the Weyl group is generated by those reflections ρ_α in (6.12) corresponding to simple roots. But from (6.12) it follows that $\rho_\alpha = \rho_\alpha v$. Therefore we can fix the Weyl symmetry by going to the fundamental dual Weyl chamber of $H^\vee$. In other words, the Dirac quantisation condition can be rewritten as

$$e\mathrm{Q}/2\pi \in \Lambda_w\left(H^\vee\right)/W \cong \Lambda_w^+\left(H^\vee\right)$$

where $\Lambda_w^+\left(H^\vee\right)$ are the dominant weights of $H^\vee$, which are in one-to-one correspondence with the finite-dimensional irreducible representations of $H^\vee$. We now turn to a more detailed description of the dual group.

6.4.1 Some lattices and dual groups

In our flash review of compact Lie groups, we have already encountered several lattices. We will now review their interrelations and in particular how they can be used to describe the dual of a connected compact Lie group. For the purposes of studying the Dirac quantisation condition, we will take the unbroken gauge group H to be a compact connected Lie group. Such a group is covered finitely by a compact group $\tilde{H} = \tilde{K} \times S$, where $\tilde{K}$ is a simply-connected compact Lie group (hence semisimple, and in turn the product of simple factors $\tilde{K}_1 \times \cdots \times \tilde{K}_p$) and S is a torus. S is the connected component of the identity of the centre of $\tilde{H}$. We define the dual group $H^\vee$ of H to be the compact connected Lie group whose weight lattice is dual to the weight lattice of H. From this it follows that just like H is a finite quotient of $\tilde{H}$, so will $H^\vee$ be a finite quotient of $\tilde{H}^\vee = \tilde{K}_1^\vee \times \cdots \times \tilde{K}_p^\vee \times S^\vee$. It is impractical to treat the general case, so we will discuss separately the cases of H abelian and H simple. From these ingredients it should be possible to treat the case of general H should the urge arise. All Lie groups in this section are compact and connected unless stated otherwise.

H abelian

If H is abelian, then it is a torus. Let $\mathfrak{h}$ be its Lie algebra. The (reduced) exponential map is surjective and defines a diffeomorphism $H \cong \mathfrak{h}/\Lambda$, where $\Lambda \subset \mathfrak{h}$ is the lattice of periods of H. As we reviewed above, this lattice is dual to the weight lattice $\Lambda_w(H) \subset \mathfrak{h}^*$. By definition this is the weight lattice of the dual group $H^\vee$. Hence we have a diffeomorphism $H^\vee \cong \mathfrak{h}^*/\Lambda_w(H)$. Notice that $\mathfrak{h}^\vee$ is identified with $\mathfrak{h}^*$.

H simple

Let H be a simple Lie group, $\tilde{H}$ its universal covering group, and $\mathfrak{h}$ its Lie algebra. Let T be a fixed maximal torus and $\mathfrak{t} \subset \mathfrak{h}$ its Lie algebra. We let $\mathfrak{t}^*$ be the space of linear forms $\mathfrak{t} \to \mathbb{R}$. The root lattices $\Lambda_r(H)$ and $\Lambda_r(\tilde{H})$ in $\mathfrak{t}^*$ agree, since as explained above they only depend on the Lie algebra. We will therefore write it as $\Lambda_r(\mathfrak{h})$. The weight lattices $\Lambda_w(H)$ and $\Lambda_w(\tilde{H})$ are different, with $\Lambda_w(\tilde{H})$ depending only on the Lie algebra again, since it is the lattice of fundamental weights. We will then often write it as $\Lambda_w(\mathfrak{h})$. We have the following inclusions:

$$\Lambda_r(\mathfrak{h}) \subseteq \Lambda_w(H) \subseteq \Lambda_w(\mathfrak{h})$$

where the first inclusion is an equality when H is the adjoint group, and the last inclusion is an equality when $H = \tilde{H}$. From Table 6.3 we see that for E_8, F_4 and H_2, the adjoint group is simply-connected, so that in these cases, and in these cases only, are both inclusions equalities.

The dual of these lattices give rise to lattices in $\mathfrak{t}$. Dualising the lattices reverses the inclusions in (6.15), so we have

$$\Lambda_r(\mathfrak{h})^* \supseteq \Lambda_w(H)^* \supseteq \Lambda_w(\mathfrak{h})^* \tag{6.16}$$

We have met some of these lattices before. $\Lambda_w(\mathfrak{h})^* = \Lambda_I(\tilde{H}) = \Lambda_r^\vee(\mathfrak{h})$, which again only depends on the Lie algebra. Similarly, $\Lambda_w(H)^* = \Lambda_I(H)$ is the integer lattice of H: those elements $h \in T$ such that $2\pi h$ lies in the kernel of the exponential map $\exp : \mathfrak{t} \to T$. It clearly depends on H, as it will be different for H and for $\tilde{H}$.

Now with the same notation as above, let us consider the inverse root system $\Phi^\vee$. As mentioned above, it is a simple root system. Therefore by the construction outlined above, there will be a complex simple Lie algebra $\mathfrak{h}_{\mathbb{C}}^\vee$ associated to $\Phi^\vee$, which has a unique compact real form $\mathfrak{h}^\vee$, which can be integrated to a unique connected compact simply-connected Lie group $\tilde{H}^\vee$. We will now exhibit the dual group $H^\vee$ of H as a quotient of $\tilde{H}^\vee$ by a finite subgroup of its centre.

The inverse root system requires for its very definition the existence of the metric: $\alpha^\vee = 2\alpha/(\alpha,\alpha)$. We can undo this dependence by using the metric to map each $\alpha^\vee \in \mathfrak{t}^*$ to a unique $\alpha^* \in \mathfrak{t}$ such that if $\beta \in \mathfrak{t}^*, \beta(\alpha^*) = (\beta, \alpha^\vee)$. We have met these α^* before: they are nothing but the coroots h_α. The coroots generate a root system in $\mathfrak{t}$ whose root lattice is the coroot lattice $\Lambda^\vee(\mathfrak{h})$ of $\mathfrak{h}$ and whose fundamental weight lattice is the dual lattice to the root lattice of $\mathfrak{h}$.

Exercise 6.17 (The dual fundamental weights)

Let $\Lambda_w(\mathfrak{h}^\vee) \subset \mathfrak{t}$ denote the lattice of fundamental weights of the dual root system. Prove that

$$\Lambda_w(\mathfrak{h}^\vee) = \Lambda_r(\mathfrak{h})^*$$

In other words, if we let $\mathfrak{t}^\vee = \mathfrak{t}^*$, then on $\mathfrak{t}^{\vee *} = \mathfrak{t}$, we have a root lattice $\Lambda_r(\mathfrak{h}^\vee) = \Lambda_r^\vee(\mathfrak{h}) = \Lambda_w(\mathfrak{h})^*$ and a fundamental weight lattice $\Lambda_w(\mathfrak{h}^\vee) = \Lambda_r(\mathfrak{h})^*$. On $\mathfrak{t}^\vee = \mathfrak{t}^*$ there is also a notion of reduced exponential map exp: $\mathfrak{t}^\vee \to T^\vee$ which is given by the canonical projection $\mathfrak{t}^* \to \mathfrak{t}^*/\Lambda_r(\mathfrak{h})$. The centre of $\tilde{H}^\vee$ is given by

$$Z\left(\tilde{H}^\vee\right) \cong \Lambda_w(\mathfrak{h}^\vee)/\Lambda_r(\mathfrak{h}^\vee) \cong \Lambda_r(\mathfrak{h})^*/\Lambda_w(\mathfrak{h})^* \cong \Lambda_w(\mathfrak{h})/\Lambda_r(\mathfrak{h}) \cong Z(\tilde{H})$$

where we have used Exercise 6.16.

Now, by definition, the weight lattice $\Lambda_w(H^\vee)$ of $H^\vee$ is dual of the weight lattice $\Lambda_w(H)$ of H:

$$\Lambda_w(H^\vee) \equiv \Lambda_w(H)^*$$

which sits between the above two lattices in $\mathfrak{t}$:

$$\Lambda_r(\mathfrak{h}^\vee) \subseteq \Lambda_w(H^\vee) \subseteq \Lambda_w(\mathfrak{h}^\vee)$$

From the above discussion surrounding equations (6.13) and (6.14), we know that $H^\vee$ is given by $\tilde{H}^\vee/\Gamma^\vee$ where $\Gamma^\vee \subseteq Z\left(\tilde{H}^\vee\right)$ is the subgroup of the centre of $\tilde{H}^\vee$ defined by $\Lambda_w(H^\vee)^*/\Lambda_w(\mathfrak{h}^\vee)^* = \Lambda_w(H)/\Lambda_r(\mathfrak{h})$. But consider now the subgroup $\Gamma \equiv \Gamma_H \subseteq Z(\tilde{H})$ which defines $H = \tilde{H}/\Gamma$. Taking into account Exercise 6.16, we find that it is given by

$$\Gamma \cong \Lambda_w(\mathfrak{h})/\Lambda_w(H) \cong (\Lambda_w(\mathfrak{h})/\Lambda_r(\mathfrak{h}))/(\Lambda_w(H)/\Lambda_r(\mathfrak{h})) \cong Z(\tilde{H})/\Gamma^\vee$$

whence

$$|\Gamma|\,|\Gamma^\vee| = |Z(\tilde{H})| \tag{6.17}$$

Let us now look at examples of dual groups. Above we listed the connected compact simple Lie groups. We now do the same for their duals. This has been done in GNO77. We list the results in Table 6.4. Most cases

H	$H^\vee$
$SU(pq)/\mathbb{Z}_p$	$SU(pq)/\mathbb{Z}_q$
$\mathrm{Spin}(2\ell+1)$	$USp(2\ell)/\mathbb{Z}_2$
$SO(2\ell+1)$	$USp(2\ell)$
$SO(2\ell)$	$SO(2\ell)$
$\mathrm{Spin}(4\ell+2)$	$\mathrm{Spin}(4\ell+2)/\mathbb{Z}_4$
$\mathrm{Spin}(4\ell)$	$\mathrm{Spin}(4\ell)/\left(\mathbb{Z}_2\times\mathbb{Z}_2\right)$
$\mathrm{Spin}(8\ell)/\mathbb{Z}_2'$	$\mathrm{Spin}(8\ell)/\mathbb{Z}_2''$
$\mathrm{Spin}(8\ell)/\mathbb{Z}_2''$	$\mathrm{Spin}(8\ell)/\mathbb{Z}_2''$
$\mathrm{Spin}(8\ell+4)/\mathbb{Z}_2'$	$\mathrm{Spin}(8\ell+4)/\mathbb{Z}_2''$
G_2	G_2
F_4	F_4
E_6	$E_6/\mathbb{Z}_3$
E_7	$E_7/\mathbb{Z}_2$
E_8	E_8

Table 6.4: The connected compact simple Lie groups and their duals.

can be determined without any computation, but some of the D_ℓ series turn out to be subtle, and require an explicit description of the weight and root lattices. They are listed, for example, in Hum72.

Equation (6.17) tells us that the orders of Γ and $\Gamma^\vee$ are complementary in $|Z(\tilde{H})| = \left|Z\left(\tilde{H}^\vee\right)\right|$. This means that the dual of the simply-connected group $\tilde{H}$ is the adjoint group $\tilde{H}^\vee/Z\left(\tilde{H}^\vee\right)$. This already tells us the last five entries of Table 6.4 as well as the second, third, fifth and sixth entries, and the special case $p=1$, in the first entry. But, in fact, the rest of the first entry also requires no further calculation. Since any subgroup of a cyclic group is cyclic and is moreover unique, the dual of $SU(pq)/\mathbb{Z}_p$ has to be $SU(pq)/\mathbb{Z}_q$, since given $\mathbb{Z}_p\subset\mathbb{Z}_{pq}$ there is a unique subgroup of $\mathbb{Z}_{pq}$ of order q, and it is $\mathbb{Z}_q$. The same argument also applies to $D_{2\ell+1}$, since the centre is cyclic in this case: whence $SO(4\ell+2)$ is self-dual. For the groups with root system $D_{2\ell}$ one has to work harder.

An example: Spin(8) and its quotients

As an example we will work out the example of Spin(8) and its factor groups. The root system of Spin(8) is D_4 and we worked out the Cartan matrix, the centre and the fundamental weights above. The lattice of fundamental weights $\Lambda_w = \mathbb{Z}\langle\lambda_i\rangle$ is the integer span of the fundamental weights λ_i. The
root lattice is the sublattice of the fundamental weights generated by the combinations:

$$\begin{aligned} &\alpha_1 = 2\lambda_1-\lambda_2 \quad \alpha_2 = -\lambda_1+2\lambda_2-\lambda_3-\lambda_4 \\ &\alpha_3 = -\lambda_2+2\lambda_3 \ \alpha_4 = -\lambda_2+2\lambda_4. \end{aligned}$$

Equivalently it is the lattice consisting of elements $\sum_{i=1}^4 n_i\lambda_i$ where $n_i\in\mathbb{Z}$ such that n_1, n_3 and n_4 are either all even or all odd. There are three intermediate lattices corresponding to the weight lattices of the three subgroups $SO(8), \mathrm{Spin}(8)/\mathbb{Z}_2'$ and $\mathrm{Spin}(8)/\mathbb{Z}_2''$: $\Lambda_1 = \Lambda_r\cup\left(\Lambda_r+\lambda_1\right), \Lambda_3 = \Lambda_r\cup(\Lambda_r+\lambda_3)$, and $\Lambda_4=\Lambda_r\cup(\Lambda_r+\lambda_4)$. Equivalently,

$$\Lambda_1 = \left\{ \sum_{i=1}^4 n_i\lambda_i \mid n_3 \equiv n_4(2) \right\}$$
$$\Lambda_3 = \left\{ \sum_{i=1}^4 n_i\lambda_i \mid n_1 \equiv n_4(2) \right\}$$
$$\Lambda_4 = \left\{ \sum_{i=1}^4 n_i\lambda_i \mid n_1 \equiv n_3(2) \right\}$$

all other integers n_i unconstrained. We can easily find a $\mathbb{Z}$-basis for these lattices as follows:

$$\begin{aligned} \Lambda_1 &= \mathbb{Z}\langle \lambda_1, \lambda_2, \lambda_3 \pm \lambda_4 \rangle \\ \Lambda_3 &= \mathbb{Z}\langle \lambda_2, \lambda_3, \lambda_1 \pm \lambda_4 \rangle \\ \Lambda_4 &= \mathbb{Z}\langle \lambda_2, \lambda_4, \lambda_1 \pm \lambda_3 \rangle \end{aligned}$$

The dual picture is as follows. Take as a basis the canonical dual basis $\{\alpha^i\}$ to the roots: $\alpha^i(\alpha_j) = \delta^i{}_j$. Their $\mathbb{Z}$-span is the lattice Λ_r^* and all lattices of interest are contained in it, so their elements will be integer linear combinations of the α^i. Given a sublattice $\Lambda \subseteq \Lambda_w$ described as the $\mathbb{Z}$-span of some vectors v_i in the weight lattice Λ_w, the dual lattice will be the sublattice $\Lambda^* \subseteq \Lambda_r^*$ given by the $\mathbb{Z}$-span of the canonical dual basis v^i to the v_i. Let $v_i = \sum_j M_i{}^j\lambda_j$, where $M_i{}^j \in \mathbb{Z}$ since Λ is a sublattice of Λ_w. Similarly $v^i = \sum_j N^i{}_j\alpha^j$, where $N^i{}_j \in \mathbb{Z}$. We can solve for N in terms of M and the Cartan matrix C as follows. By definition, $v^i(v_j) = \delta^i{}_j$, whence

$$\begin{aligned} \delta^i{}_j &= v_j\left(v^i\right) \\ &= N^i{}_k M_j^\ell \lambda_\ell\left(\alpha^k\right) \\ &= N^i{}_k M_j^\ell \left(C^{-1}\right)_\ell{}^k \end{aligned}$$

where we have used that $\lambda_\ell = \left(C^{-1}\right)_\ell{}^k\alpha_k$. In other words, $N = \left(CM^{-1}\right)^t$. Computing this for each of the lattices above, we find:

$$\begin{aligned} \Lambda_1^* &= \mathbb{Z}\left\langle 2\alpha^1 - \alpha^2, \alpha^1 - \alpha^2 + \alpha^3 + \alpha^4, \alpha^2 - \alpha^3 - \alpha^4, \alpha^3 - \alpha^4 \right\rangle \\ \Lambda_3^* &= \mathbb{Z}\left\langle \alpha^1 - \alpha^2 + \alpha^4, \alpha^1 - \alpha^2 + \alpha^3 + \alpha^4, \alpha^2 - 2\alpha^3, \alpha^1 - \alpha^4 \right\rangle \\ \Lambda_4^* &= \mathbb{Z}\left\langle \alpha^1 - \alpha^2 + \alpha^3, \alpha^1 - \alpha^2 + \alpha^3 + \alpha^4, \alpha^1 - \alpha^3, \alpha^2 - 2\alpha^4 \right\rangle \end{aligned}$$

We can understand these lattices as sublattices of Λ_r^* by changing basis to the α^i and constraining the coefficients. We find

$$\Lambda_1^* = \left\{ \sum_{i=1}^4 n_i\alpha^i \mid n_3 \equiv n_4(2) \right\}$$
$$\Lambda_3^* = \left\{ \sum_{i=1}^4 n_i\alpha^i \mid n_1 \equiv n_4(2) \right\}$$
$$\Lambda_4^* = \left\{ \sum_{i=1}^4 n_i\alpha^i \mid n_1 \equiv n_3(2) \right\}$$

whence we conclude that all three lattices are self-dual, in agreement with Table 6.4.

Another example: Spin(12) **and its quotients**

As a final example, and to illustrate the other behaviour of the $D_{2\ell}$ series, we will work out the example of Spin(12) and its factor groups. The root system of Spin(12) is D_6, whose Cartan matrix follows from Table 6.2:

$$\begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & -1 \\ 0 & 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & -1 & 0 & 2 \end{pmatrix}$$

The fundamental weights are given by

$$\begin{aligned}
\lambda_1 &= \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \frac{1}{2}\alpha_5 + \frac{1}{2}\alpha_6 \\
\lambda_2 &= \alpha_1 + 2\alpha_2 + 2\alpha_3 + 2\alpha_4 + \alpha_5 + \alpha_6 \\
\lambda_3 &= \alpha_1 + 2\alpha_2 + 3\alpha_3 + 3\alpha_4 + \frac{3}{2}\alpha_5 + \frac{3}{2}\alpha_6 \\
\lambda_4 &= \alpha_1 + 2\alpha_2 + 3\alpha_3 + 4\alpha_4 + 2\alpha_5 + 2\alpha_6 \\
\lambda_5 &= \frac{1}{2}\alpha_1 + \alpha_2 + \frac{3}{2}\alpha_3 + 2\alpha_4 + \frac{3}{2}\alpha_5 + \alpha_6 \\
\lambda_6 &= \frac{1}{2}\alpha_1 + \alpha_2 + \frac{3}{2}\alpha_3 + 2\alpha_4 + \alpha_5 + \frac{3}{2}\alpha_6
\end{aligned}$$

It follows that the centre $\Lambda_w/\Lambda_r \cong \mathbb{Z}_2 \times \mathbb{Z}_2$ consists of the following Λ_r-cosets: $0, \lambda_1, \lambda_5$ and λ_6, with multiplication table:

	0	λ_1	λ_5	λ_6
0	0	λ_1	λ_5	λ_6
λ_1	λ_1	0	λ_6	λ_5
λ_5	λ_5	λ_6	0	λ_1
λ_6	λ_6	λ_5	λ_1	0

where as usual all entries are modulo Λ_r.

Letting $\Lambda_r = \mathbb{Z}\langle\lambda_i\rangle$, the root lattice is the sublattice $\Lambda_r = \mathbb{Z}\langle\alpha_i\rangle$ spanned by the following combinations:

$$\begin{aligned}
\alpha_1 &= 2\lambda_1 - \lambda_2 & \alpha_2 &= -\lambda_1 + 2\lambda_2 - \lambda_3 \\
\alpha_3 &= -\lambda_2 + 2\lambda_3 - \lambda_4 & \alpha_4 &= -\lambda_3 + 2\lambda_4 - \lambda_5 - \lambda_6 \\
\alpha_5 &= -\lambda_4 + 2\lambda_5 & \alpha_6 &= -\lambda_4 + 2\lambda_6
\end{aligned}$$

Equivalently it is the lattice consisting of elements $\sum_{i=1}^{6} n_i\lambda_i$ where $n_i \in \mathbb{Z}$ such that $n_1 + n_3, n_5$ and n_6 are either all even or all odd. There are three intermediate lattices corresponding to the weight lattices of the three subgroups $SO(12), \mathrm{Spin}(12)/\mathbb{Z}_2'$ and $\mathrm{Spin}(12)/\mathbb{Z}_2''$: $\Lambda_1 = \Lambda_r \cup (\Lambda_r + \lambda_1)$, $\Lambda_5 = \Lambda_r \cup (\Lambda_r + \lambda_5)$, and $\Lambda_6 = \Lambda_r \cup (\Lambda_r + \lambda_6)$. Equivalently,

$$\Lambda_1 = \left\{ \sum_{i=1}^{6} n_i \lambda_i \mid n_5 \equiv n_6(2) \right\}$$
$$\Lambda_5 = \left\{ \sum_{i=1}^{6} n_i \lambda_i \mid n_1 + n_3 \equiv n_6(2) \right\}$$
$$\Lambda_6 = \left\{ \sum_{i=1}^{6} n_i \lambda_i \mid n_1 + n_3 \equiv n_5(2) \right\}$$

all other integers n_i unconstrained. We can easily find a $\mathbb{Z}$-basis for these lattices as follows:

$$\Lambda_1 = \mathbb{Z}\left\langle \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5 \pm \lambda_6 \right\rangle$$
$$\Lambda_5 = \mathbb{Z}\left\langle \lambda_2, \lambda_4, \lambda_5, \lambda_1 + \lambda_3, \lambda_1 + \lambda_6, \lambda_3 + \lambda_6 \right\rangle$$
$$\Lambda_6 = \mathbb{Z}\left\langle \lambda_2, \lambda_4, \lambda_6, \lambda_1 + \lambda_3, \lambda_1 + \lambda_5, \lambda_3 + \lambda_5 \right\rangle$$

Following the discussion given in the previous example, the dual lattices are given by

$$\Lambda_1^* = \mathbb{Z}\big\langle 2\alpha^1 - \alpha^2, \alpha^1 - 2\alpha^2 + \alpha^3, \alpha^2 - 2\alpha^3 + \alpha^4, \alpha^3 - 2\alpha^4 + \alpha^5 + \alpha^6$$
$$\alpha^5 - \alpha^6, \alpha^4 - \alpha^5 - \alpha^6 \big\rangle$$
$$\Lambda_5^* = \mathbb{Z}\big\langle \alpha^1 - \alpha^2 + \alpha^3 - \alpha^6, \alpha^1 - \alpha^3 + \alpha^4 - \alpha^6, \alpha^3 - 2\alpha^4 + \alpha^5 + \alpha^6$$
$$\alpha^1 - 2\alpha^2 + \alpha^3, \alpha^5 - 2\alpha^6, \alpha^1 - \alpha^3 + \alpha^6, \big\rangle$$
$$\Lambda_6^* = \mathbb{Z}\big\langle \alpha^1 - \alpha^2 + \alpha^3 - \alpha^5, \alpha^1 - \alpha^3 + \alpha^4 - \alpha^5, \alpha^3 - 2\alpha^4 + \alpha^5 + \alpha^6$$
$$\alpha^1 - 2\alpha^2 + \alpha^3, \alpha^1 - \alpha^3 + \alpha^5, \alpha^4 - 2\alpha^6 \big\rangle$$

We can understand these lattices as sublattices of Λ_r^* by changing basis to the α^i and constraining the coefficients. We find

$$\Lambda_1^* = \left\{ \sum_{i=1}^{6} n_i \alpha^i \mid n_5 \equiv n_6(2) \right\}$$
$$\Lambda_5^* = \left\{ \sum_{i=1}^{6} n_i \alpha^i \mid n_1 + n_3 \equiv n_5(2) \right\}$$
$$\Lambda_6^* = \left\{ \sum_{i=1}^{6} n_i \alpha^i \mid n_1 + n_3 \equiv n_6(2) \right\}$$

whence we conclude that Λ^1 is self-dual, whereas duality interchanges the groups whose weight lattices are Λ_5 and Λ_6. It can be shown that the group whose weight lattice is Λ_1 is $SO(12)$. Again this is in agreement with Table 6.4.

13 Duality Transformations in Supersymmetric Yang-Mills Theories coupled to Supergravity by Ceresole, Ferrara, Proeyen

Abstract

We consider duality transformations in $N = 2, d = 4$ Yang-Mills theory coupled to $N = 2$ supergravity. A symplectic and coordinate covariant framework is established, which allows one to discuss stringy 'classical and quantum duality symmetries' (monodromies), incorporating T and S dualities. In particular, we shall be able to study theories (like

$N = 2$ heterotic strings) which are formulated in symplectic basis where a 'holomorphic prepotential' F does not exist, and yet give general expressions for all relevant physical quantities. Duality transformations and symmetries for the $N = 1$ matter coupled Yang-Mills supergravity system are also exhibited. The implications of duality symmetry on all $N > 2$ extended supergravities are briefly mentioned. We finally give the general form of the central charge and the $N = 2$ semiclassical spectrum of the dyonic BPS saturated states (as it comes by truncation of the $N = 4$ spectrum).

13.1 1 Introduction

Recently, proposals for the quantum moduli space of $N = 2$ rigid Yang-Mills theories [1] have been given in terms of particular classes of genus r Riemann surfaces parametrized by r complex moduli[2], r being the rank for the gauge group G broken to $U(1)^r$ for generic values of the moduli. The effective action for such theories, with terms up to two derivatives, is described by $N = 2$ supersymmetric lagrangians of r abelian massless vector multiplets[3], whose dynamics is encoded in a holomorphic prepotential $F\left(X^A\right)$, function of the moduli coordinates $X^A (A = 1, \ldots, r)$. According to Seiberg and Witten [1] this effective theory has classical, perturbative and non perturbative duality symmetries which reflect on monodromy properties of certain holomorphic symplectic vectors $\left(X^A, F_A(X)\right)$, eventually related to periods of holomorphic one-forms [1]

$$\omega = X^A \alpha_A + F_A \beta^A \tag{1.1}$$

where α_A, β^A is a basis for the $2r$ homology cycles of a genus r Riemann surface. The Picard-Fuchs equations satisfied by the holomorphic vector one-form $U_i = \left(\partial_i X^A, \partial_i F_A\right) (i = 1, \ldots, r)$ can be regarded as differential identities for "rigid special geometry" [7]. To attach a particular algebraic curve to "rigid special geometry" is therefore equivalent to exactly compute the holomorphic data U_i, and thus to exactly reconstruct the effective action for the self interaction of the r massless gauge multiplets once the massive states, both perturbative and non perturbative, have been integrated out. Indeed it is a virtue of $N = 2$ supersymmetry that all the couplings in the effective Lagrangian, including 4 -fermion terms, can be computed purely in terms of the holomorphic data. Quite remarkably the quantum monodromies dictate the monopole and dyon spectrum of the effective theory [17.27] which turns out to be "dual" to non-perturbative instanton effects [5] in the original G-invariant microscopic theory [6]:7].

This paper considers several issues in order to extend the approach pursued in the rigid case to the more challenging case of coupling an $N = 2$ Yang-Mills theory to gravity. In particular we shall include in the $N = 2$ supergravity theory a dilaton-axion vector multiplet which is an essential ingredient to describe effective $N = 2$ theories which come from the low energy limit of $N = 2$ heterotic string theories in four dimensions [8]. Another ingredient is the extension of the "classical monodromies" to $N = 2$ local supersymmetry. For rigid theories the classical metric is essentially the Cartan matrix of the group G and the classical monodromies are related to the Weyl group of the Cartan subalgebra of G [2]. For $N = 2$ supergravity theories coming from $N = 2$ heterotic strings, the classical metric of the moduli space of the pure gauge sector is based on the homogeneous space $O(2, r)/O(2) \times O(r)$ [3. [8 [10] and the classical monodromies are related to the T-duality group $O(2, r; \mathbb{Z})$ which in particular is an invariance of the massive charged states 11]. This state of affair is quite analogous to the analysis performed by Sen and Schwarz 12] for the $N = 4$ heterotic string compactifications, in which case an exact quantum duality symmetry $SL(2, \mathbb{Z}) \times O(6, r; \mathbb{Z})$ was conjectured [12 [16] and a resulting spectrum for BPS states with both electric and magnetic states was proposed. In the $N = 4$ theory the $SL(2, \mathbb{Z}) \times O(6, r; \mathbb{Z})$ symmetry, using general arguments [17,18, has a natural embedding in $Sp(2(6+r); \mathbb{Z})$, acting on the $6+r$ vector self-dual field strengths $\mathcal{F}^{+A}_{\mu\nu}$ and their "dual" defined through $G_{+A}^{\mu\nu} \equiv -i\frac{\delta\mathcal{L}}{\delta\mathcal{F}_{\mu\nu}^{+A}}$. In generic $N=2$ theories, because of quantum corrections [19, 20], we do not expect such factorized $S-T$ duality to occur anymore [4]. Indeed this can be argued with a pure supersymmetry argument, related to the fact that once the classical moduli space $O(2,r)/O(2)\times O(r)$ is deformed by quantum corrections, then the factorized structure with the dilaton degrees of freedom is lost and a non trivial moduli space, mixing the S and T degrees of freedom should emerge. This result is in fact a consequence of a theorem on "special geometry" [21,22] which asserts that the only factorized special manifolds are the $\frac{SU(1,1)}{U(1)}\times\frac{O(2,r)}{O(2)\times O(r)}$ series, which precisely describe the "classical moduli space" of $S-T$ moduli. Because of the coupling to gravity, the symplectic structure and identification of periods, coming from special geometry, is also remarkably different from rigid special geometry. Indeed the interpretation of $\left(X^\Lambda, F_\Lambda\right), \Lambda=0,1,\ldots,r+1$ as periods of algebraic curves is no longer appropriate to genus r Riemann surfaces, as it can be seen from the Picard-Fuchs equations [23.24] and from the form of the metric $g_{i\bar{\jmath}}=-\partial_i\partial_{\bar{\jmath}}\log i\left(\bar{F}_A X^A-\bar{X}^A F_A\right)$ of the moduli space 23 29. In fact special geometry is known to be appropriate to a particular class of complex manifolds (Calabi-Yau manifolds or their mirrors) and to describe the deformations of the complex structure [23]. It is therefore tempting to argue that the quantum moduli space including $S-T$ duality and its monodromies is related to 3 -manifolds (or their mirrors) with $h_{(2,1)}=r+1$.

The paper is organized as follows: In chapter 2 we give a résumé of rigid theories, also discussing duality for the fermionic sector and the physical significance of monodromies and geometrical data, such as the holomorphic tensor C_{ijk}, related to the gaugino anomalous magnetic moment. In chapter 3 we describe in detail the coupling to gravity, the extension of duality to the fermionic sector and the existence of
symplectic bases which do not admit a prepotential function F, as it occurs in certain formulations of $N=2$ supergravities coming from $N=2$ heterotic strings. The general form of duality transformations and symmetries as they occur in $N=1$ locally supersymmetric Yang-Mills theories coupled to matter is also described. In chapter 4 we use such a formulation where all the perturbative duality symmetries become invariances of the action. Then, we discuss the implementation of duality symmetries in $N>2$ extended supergravities for the spectrum of dyonic states. In chapter 5 we analyze classical and quantum duality symmetries and give generic formulae for the spectrum of the BPS states and the "semiclassical formulae" when the non perturbative spectrum is computed in terms of the "classical periods". The explicit expression for the $r=2$ case is given as an example, and the special occurrence of enhanced symmetry points is described. The paper ends with some concluding remarks.

13.2 2 Résumé of rigid special geometry

13.2.1 2.1 Basics

$N=2$ supersymmetric gauge theory on a group G broken to $U(1)^r$, with $r=\operatorname{rank} G$, corresponds to a particular case of the most general $N=1$ coupling of r chiral multiplets $\left(X^A, \chi^A\right)$ to $rN=1$ abelian vector multiplets $\left(\mathcal{A}_\mu^A, \lambda^A\right)$ in which the Kähler potential K and the holomorphic kinetic term function $f_{AB}\left(X^A\right)$ are given by

$$\begin{aligned} K &= i\left(\bar{F}_A X^A-F_A\bar{X}^A\right), \quad \left(F_A=\partial_A F\right) \\ f_{AB} &= \partial_A\partial_B F \equiv F_{AB} \end{aligned} \tag{2.1}$$

in terms of the single prepotential $F(X)$ [3]. One can show that the Kähler geometry is constrained because the Riemann tensor satisfies the identity [26,4]

$$R_{A\bar{B}C\bar{D}} = -\partial_A \partial_C \partial_P F \partial_{\bar{B}} \partial_{\bar{D}} \partial_{\bar{Q}} \bar{F} g^{P\bar{Q}} \tag{2.2}$$

with

$$g_{P\bar{Q}} = \partial_P \partial_{\bar{Q}} K = 2 \,\mathrm{Im}\, \partial_P \partial_Q F \tag{2.3}$$

The lagrangian has the form

$$\begin{aligned} \mathcal{L} &= g_{A\bar{B}} \partial_\mu X^A \partial_\mu \bar{X}^B + \left(g_{A\bar{B}} \lambda^{IA} \sigma^\mu \mathcal{D}_\mu \bar{\lambda}_I^{\bar{B}} + \text{ h.c. } \right) \\ &+ \mathrm{Im} \left(F_{AB} \mathcal{F}_{\mu\nu}^{-A} \mathcal{F}_{\mu\nu}^{-B} \right) + \mathcal{L}_{\text{Pauli}} + \mathcal{L}_{4-\text{ fermi}} \end{aligned} \tag{2.4}$$

where $A, B, \ldots$ run on the adjoint representation of the gauge group G, $I = 1, 2$ and $\mathcal{F}_{\mu\nu}^{+A} = \mathcal{F}_{\mu\nu}^{A} - \frac{i}{2} \epsilon_{\mu\nu\rho\sigma} \mathcal{F}^{A\rho\sigma}$ (and $\mathcal{F}_{\mu\nu}^{-A} = \overline{\mathcal{F}}_{\mu\nu}^{+A}$). As we shall see, also $\mathcal{L}_{\text{Pauli}}$ and $\mathcal{L}_{4-\text{ Fermi}}$ contain the function F and its derivatives up to the fourth.

The previous formulation, derived from tensor calculus, is incomplete because it is not coordinate covariant. It is written in a particular coordinate system ("special coordinates") which is not uniquely selected. In fact, eq.(2.1) is left invariant under particular coordinate changes of the $X^A \to \widetilde{X}^A$ with some new function $\widetilde{F}(\widetilde{X})$ described by

$$\begin{aligned} \widetilde{X}^A(X) &= A_B^A X^B + B^{AB} F_B(X) + P^A \\ \widetilde{F}_A \left(\widetilde{X}^A(X) \right) &= C_{AB} X^B + D_A^B F_B(X) + Q_A \end{aligned} \tag{2.5}$$

where $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is an $Sp(2r, \mathbf{R})$ matrix

$$A^T C - C^T A = 0 \quad , \quad B^T D - D^T B = 0 \quad , \quad A^T D - C^T B = \mathbb{1} \tag{2.6}$$

and P^A, Q_A can be complex constants which from now on will be set to zero.

It can be shown that a function $\widetilde{F}$ exists such that [3]

$$\widetilde{F}_A = \frac{\partial \widetilde{F}}{\partial \widetilde{X}^A} \tag{2.7}$$

provided the mapping $X^A \to \widetilde{X}^A$ is invertible.

It is well known that the equations of motion and the Bianchi identities [3] [17] [18]

$$\begin{aligned} \partial^\mu \,\mathrm{Im}\, \mathcal{F}_{\mu\nu}^{-A} &= 0 \quad \text{Bianchi identities} \\ \partial_\mu \,\mathrm{Im}\, G_{-A}^{\mu\nu} &= 0 \quad \text{Equations of motion} \end{aligned} \tag{2.8}$$

transform covariantly under (2.5) (with $P^A = Q_A = 0$), so that $\left(\mathcal{F}_{\mu\nu}^{-A}, G_{-A}^{\mu\nu} \right)$ is a symplectic vector. Here, $G_{-A}^{\mu\nu} \equiv i \frac{\delta \mathcal{L}}{\delta \mathcal{F}_{\mu\nu}^{-A}} = \overline{\mathcal{N}}_{AB} \mathcal{F}_{\mu\nu}^{-B} +$ fermionic terms, where we have set $F_{AB} = \overline{\mathcal{N}}_{AB}$ in order to unify the notations to the gravitational case [3]. The transformations (2.5) leave invariant the whole lagrangian but the vector kinetic term. Indeed, neglecting for the moment fermion terms (see section 2.2) and setting
for simplicity $\mathcal{F}_{\mu\nu}^{-A} = \mathcal{F}^A$ and $G_{-A}^{\mu\nu} = G_A$ the vector kinetic lagrangian transforms as follows

$$\begin{aligned} \mathrm{Im}\, \mathcal{F}^A \overline{\mathcal{N}}_{AB} \mathcal{F}^B &\to \mathrm{Im}\, \widetilde{\mathcal{F}}^A \widetilde{G}_A = \\ &= \mathrm{Im} \left(\mathcal{F}^A G_A + 2 \mathcal{F}^A \left(C^T B \right)_A^B G_B + \right. \\ &\left. + \mathcal{F}^A \left(C^T A \right)_{AB} \mathcal{F}^B + G_A \left(D^T B \right)^{AB} G_B \right) \end{aligned} \tag{2.9}$$

If $C = B = 0$ the lagrangian is invariant. If $C \neq 0, B = 0$ it is invariant up to a four-divergence. In presence of a topologically non-trivial $\mathcal{F}^{-A}_{\mu\nu}$ background, $\left(C^T A\right)_{AB} \int \operatorname{Im} \mathcal{F}^{-A}_{\mu\nu} \mathcal{F}^{-B}_{\mu\nu} \neq 0$, one sees that in the quantum theory duality transformations must be integral valued in $Sp(2r, \mathbb{Z})$ [1] and transformations with $B = 0$ will be called perturbative duality transformations.

If $B \neq 0$ the lagrangian is not invariant. As it is well known, then the duality transformation is only a symmetry of the equations of motion and not of the lagrangian.

Since $\widetilde{G}^{\mu\nu}_{-A} = \widetilde{\mathcal{N}}_{AB} \widetilde{\mathcal{F}}^{-B}_{\mu\nu}$ one also has

$$\widetilde{\mathcal{N}} = (C + D\mathcal{N})(A + B\mathcal{N})^{-1} \tag{2.10}$$

A duality transformation will be a symmetry of the theory if $\widetilde{\mathcal{N}}(\widetilde{X}) = \mathcal{N}(\widetilde{X})$, which implies $\widetilde{F}(\widetilde{X}) = F(\widetilde{X})$.

Note that $B \neq 0$ means that the coupling constant $\widetilde{\mathcal{N}}$ is inverted and symmetry transformations with $B \neq 0$ will be called quantum non perturbative duality symmetries.

The perturbative duality rotations are of the form

$$\begin{pmatrix} A & 0 \\ C & \left(A^T\right)^{-1} \end{pmatrix}, A \subset GL(r), A^T C \text{ symmetric} \tag{2.11}$$

In rigid supersymmetry the tree level symmetries are of the form $\begin{pmatrix} A & 0 \\ 0 & \left(A^T\right)^{-1} \end{pmatrix}$ while the quantum perturbative monodromy introduces a $C \neq 0$.

The general form of the central charge for BPS states in a generic $N = 2$ rigid theory is given by [1]

$$|Z| = M = \left| n^A_{(m)} F_A - n^{(e)}_A X^A \right| \tag{2.12}$$

where $n^A_{(m)}, n^{(e)}_A$ denote the values of magnetic and electric charges of the state of mass M. The above expression is manifestly symplectic covariant provided the vector $\left(n^A_{(m)}, n^{(e)}_A\right)$ is also transformed under $Sp(2r; \mathbb{Z})$. This equation shows again that a duality symmetry can only be a (perturbative) symmetry if $B = 0$, otherwise the vector subspace with $n^A_{(m)} = 0$ cannot be left invariant.

If the original unbroken gauge group is $G = SU(r+1)$, then $A \in$ Weyl group and $A^T C$ is the Cartan matrix $\langle \alpha_i \mid \alpha_j \rangle$ of $SU(r+1)$ [2].

Eq. (2.10) shows that $A + B\mathcal{N}$ has to be invertible in order that the new tensor $\widetilde{\mathcal{N}}$ exists. This is insured by the positive definiteness of $\operatorname{Im} \mathcal{N}$, which is the kinetic matrix. Here $A + B\mathcal{N} = \partial\widetilde{X}/\partial X$, so this implies the invertibility of the mapping $X \to \widetilde{X}$. As explained in (2.7), this then also implies the existence of $\widetilde{F}$. We will see that in local supersymmetry $\mathcal{N}_{AB} \neq \bar{F}_{AB}$, so that the existence of $\widetilde{F}$ is not equivalent to the invertibility of $\operatorname{Im} \mathcal{N}$, and $\widetilde{F}$ not always exists.

Special coordinates do not give a coordinate independent description of the effective action. A coordinate independent description is obtained by introducing a holomorphic symplectic bundle $V = \left(X^A(z), F_A(z)\right)$ and holomorphic $(1,0)$ forms on the Kähler manifold 4 , 1

$$U_i \equiv \partial_i V = \left(\partial_i X^A, \partial_i F_A\right) \quad \text{with } i = 1, \ldots, r \tag{2.13}$$

In rigid special geometry the U_i satisfy the constraints [4]

$$\begin{aligned} \mathcal{D}_i U_j &= i C_{ijk} g^{k\bar{l}} \bar{U}_{\bar{l}} \\ \partial_i \bar{U}_{\bar{j}} &= 0 \end{aligned} \tag{2.14}$$

Taking then the metric

$$g_{i\bar{\jmath}} = \partial_i \partial_{\bar{\jmath}} K = i \left(\partial_{\bar{\jmath}} \bar{F}_A \partial_i X^A - \partial_{\bar{\jmath}} \bar{X}^A \partial_i F_A \right) \\ = i \partial_i X^A \partial_{\bar{\jmath}} \bar{X}^B \left(\mathcal{N}_{AB} - \overline{\mathcal{N}}_{AB} \right) \tag{2.15}$$

where we used

$$\partial_{\bar{\imath}} \bar{F}_A = \mathcal{N}_{AB} \partial_{\bar{\imath}} \bar{X}^B \tag{2.16}$$

one may derive the tensor C_{ijk}

$$C_{ikp} = \partial_i X^A \mathcal{D}_k \partial_p F_A - \partial_i F_A \mathcal{D}_k \partial_p X^A \\ = \partial_i X^B \left(\partial_k \partial_p F_B - \partial_k \partial_p X^A \overline{\mathcal{N}}_{AB} \right) \tag{2.17}$$

The integrability conditions on (2.14) yields

$$R_{i\bar{\jmath}k\bar{l}} = -C_{ikp} \bar{C}_{\bar{\jmath}\bar{l}\bar{p}} g^{p\bar{p}} \tag{2.18}$$

The Bianchi identities of (2.18) also imply that C_{ijk} is a holomorphic completely symmetric tensor obeying $\mathcal{D}_{[i} C_{j]kl} = 0$.

Note that from (2.17) it also follows

$$C_{ijk} = \partial_i X^A \partial_j X^B \partial_k X^C \partial_A \partial_B \partial_C F \tag{2.19}$$

which in special coordinates reduces to

$$C_{ABC} = \partial_A \partial_B \partial_C F \tag{2.20}$$

13.2.2 2.2 Symplectic transformations in the fermionic sector

In the total supersymmetric action, the vectors also couple to fermions by terms linear in the field strength. We will first give the general features of the formulation of symplectic transformations in the presence of a fermionic sector, which could even be non-supersymmetric. Afterwards, we will specify the formulae for generic fermionic terms which we encounter in $N = 2$ lagrangians.

The general form of the Lagrangian, deleting terms which are by themselves symplectic invariant, is

$$\mathcal{L} = -\frac{i}{2} \overline{\mathcal{N}}_{AB} \mathcal{F}^{-A\mu\nu} \mathcal{F}^{-B}_{\mu\nu} - i \mathcal{F}^{-A\mu\nu} H^{-}_{A\mu\nu} + \text{c.c.} \ + \mathcal{L}_{4f} \tag{2.21}$$

where $H^{-}_{A\mu\nu}$are quadratic in the fermions, and $\mathcal{L}_{4f}$ are the quartic terms in fermions. Then

$$G^{-}_{A\mu\nu} \equiv i \frac{\delta \mathcal{L}}{\delta \mathcal{F}^{-A\mu\nu}} = \overline{\mathcal{N}}_{AB} \mathcal{F}^{-B}_{\mu\nu} + H^{-}_{A\mu\nu} = G^{-}_{bA\mu\nu} + H^{-}_{A\mu\nu} \tag{2.22}$$

As argued in ref [17], the point where the equations of motions (2.8) are satisfied is an invariant point. Thus, the first term of the action is (omitting the obvious A indices)

$$
\begin{aligned}
\mathcal{L}_V &\equiv -\frac{i}{2}\overline{\mathcal{N}}\mathcal{F}^-_{\mu\nu}\mathcal{F}^{-\mu\nu} + \text{ c.c.} \\
&= -\frac{i}{2}G^-_{b\mu\nu}\mathcal{F}^{\mu\nu} + \text{ c.c.} \\
&= i\partial^\mu G^-_{b\mu\nu}A^\nu + \text{ c.c.} \\
&= -i\partial^\mu H^-_{\mu\nu}A^\nu + \text{ c.c. } - 2\partial^\mu \operatorname{Im} G^-_{\mu\nu}A^\nu \\
&= \frac{i}{2}H^-_{\mu\nu}\mathcal{F}^{-\mu\nu} + \text{ c.c. } - 2\partial^\mu \operatorname{Im} G^-_{\mu\nu}A^\nu
\end{aligned}
\tag{2.23}
$$

Therefore

$$
\mathcal{L}|_{\frac{\delta\mathcal{L}}{\delta\mathcal{A}}=0} = -\frac{i}{2}H^-_{\mu\nu}\mathcal{F}^{-\mu\nu} + \text{ c.c. } + \mathcal{L}_{4f} \equiv \mathcal{L}_{inv} \tag{2.24}
$$

which should thus be invariant. The Lagrangian (2.21) is then

$$
\mathcal{L} = -\frac{i}{2}\mathcal{F}^{-A\mu\nu}G^-_{A\mu\nu} + \text{c.c. } + \mathcal{L}_{inv} \tag{2.25}
$$

Now we suppose $H^-_{A\mu\nu}$to be of the form

$$
H^-_{A\mu\nu} = \left(P_{Aa} - \overline{\mathcal{N}}_{AB}Q^B_a\right)\mathcal{T}^{-a}_{\mu\nu} \tag{2.26}
$$

where a denotes a new index, whose meaning depends on the model. $\mathcal{T}^{-a}_{\mu\nu}$ is a tensor not transforming under the symplectic group. Then

$$
\begin{aligned}
\mathcal{L}_{inv} &= -\frac{i}{2}\mathcal{F}^{-A\mu\nu}\left(P_{Aa} - \overline{\mathcal{N}}_{AB}Q^B_a\right)\mathcal{T}^{-a}_{\mu\nu} + \text{ c.c. } + \mathcal{L}_{4f} \\
&= -\frac{i}{2}\left(\mathcal{F}^{-A\mu\nu}P_{Aa} - G^{-\mu\nu}_{bA}Q^A_a\right)\mathcal{T}^{-a}_{\mu\nu} + \text{ c.c. } + \mathcal{L}_{4f}
\end{aligned}
\tag{2.27}
$$

Invariance of $\mathcal{L}_{\text{inv}}$ is then guaranteed if $\left(Q^A, P_A\right)$ is a symplectic vector, and $\mathcal{L}_{4f}$ is constructed as the completion of G_b to G in the above formula (plus possible completely invariant terms). These completions are thus

$$
\mathcal{L}_{4f} = \frac{i}{2}H^{-\mu\nu}_A Q^A_a \mathcal{T}^{-a}_{\mu\nu} + \text{ c.c. } + \text{ invariant terms} \tag{2.28}
$$

13.2.3 2.3 Fermions in $N=2$ rigid Yang-Mills theory

The coordinate independent description of fermions is given by $SU(2)$ doublets $\left(\lambda^{iI}, \lambda^{\bar{\imath}}_I\right)$ where upper and lower $SU(2)$ indices I mean positive and negative chiralities respectively [3] 26]. As such the spinors are symplectic invariant and contravariant world vector fields. The antiselfdual field strength $\mathcal{F}^{-A}_{\alpha\beta}$ and positive chiralities spinors are in the same $N=2$ multiplet, which is, in two component spinor notation, *

$$
\left(X^A, \partial_i X^A \lambda^{iI}_\alpha, \mathcal{F}^{-A}_{\alpha\beta}\right) \tag{2.29}
$$

with $\alpha, \beta \in SL(2,\mathbb{C})$.

$\star\mathcal{F}^{-A}_{\alpha\beta}$ is $\sigma^{\mu\nu}_{\alpha\beta}\mathcal{F}^{-A}_{\mu\nu}$.

In our application of (2.26) only $\mathcal{T}$ is dependent on the fermions λ^{iI}, while P and Q depend on the scalars X^A. The index a is now replaced by $\bar{\imath}$, and we have

$$Q_{\bar{\imath}}^{A}=\partial_{\bar{\imath}}\bar{X}^{A};\quad P_{A\bar{\imath}}=\partial_{\bar{\imath}}\bar{F}_{A}$$
$$\mathcal{T}_{\alpha\beta}^{\bar{\imath}}=kg^{\bar{\imath}j}C_{jkp}\lambda_{\alpha}^{kI}\lambda_{\beta}^{pJ}\epsilon_{IJ} \tag{2.30}$$

where k is a constant to be determined by supersymmetry. Then

$$H_{-A}^{\alpha\beta}=k\partial_{\bar{\imath}}\bar{X}^{B}\left(\mathcal{N}_{BA}-\overline{\mathcal{N}}_{BA}\right)g^{\bar{\imath}j}C_{jkp}\lambda^{\alpha kI}\lambda^{\beta pJ}\epsilon_{IJ} \tag{2.31}$$

This yields

$$\mathcal{L}_{\text{Pauli}}=-i(\mathcal{N}-\overline{\mathcal{N}})_{AB}\partial_{\bar{\imath}}\bar{X}^{A}\mathcal{T}_{\alpha\beta}^{\bar{\imath}}\mathcal{F}^{B\alpha\beta}+\text{ c.c.}$$
$$\mathcal{L}_{4\text{f}}=\frac{i}{2}\partial_{\bar{\imath}}\bar{X}^{A}\partial_{\bar{\jmath}}\bar{X}^{B}\left(\overline{\mathcal{N}}_{AB}-\mathcal{N}_{AB}\right)\mathcal{T}_{\alpha\beta}^{\bar{\imath}}\mathcal{T}^{\bar{\jmath}\alpha\beta}+c.c.+\text{ invariant terms} \tag{2.32}$$

in agreement with Cremmer et al. 30].
In special coordinates, setting $\lambda_{\alpha}^{i1}=\chi_{\alpha}^{i},\lambda_{\alpha}^{i2}=\lambda_{\alpha}^{i}$, the Pauli term reduces to

$$\mathcal{L}_{\text{Pauli}}=-k\partial_{A}\partial_{B}\partial_{C}F\left(\chi_{\alpha}^{A}\lambda_{\beta}^{B}-\lambda_{\alpha}^{A}\chi_{\beta}^{B}\right)\mathcal{F}^{-C\alpha\beta}+\text{ c.c.} \tag{2.33}$$

in agreement with the standard $N=1$ supersymmetric action with $f_{AB}=F_{AB}$ [30]. We see from (2.32) that in rigid supersymmetry the physical meaning of C_{ijk} is that of an anomalous magnetic moment. Note that C_{ijk} vanishes at tree-level and it is $\sim\frac{1}{\langle X\rangle}$ at one loop-level as it must be [19] 20] [1]. It is obviously singular at $\langle X\rangle=0$. In the $SU(2)$ quantum theory [1], the $SU(2)$ symmetry is not restored at $X=0$, and then one rather expects such terms to behave as $\frac{c_0}{\Lambda}$ where c_0 is a dimensionless number. The vanishing at tree-level of both Pauli terms and the corresponding four fermions terms is consistent with renormalizability arguments.

The other fermionic terms which are already duality invariant read

$$\lambda_{\alpha}^{iI}\lambda_{\beta}^{kJ}\epsilon^{\alpha\beta}\bar{\lambda}_{\dot{\alpha}I}^{\bar{\jmath}}\bar{\lambda}_{\dot{\beta}J}^{\bar{l}}\epsilon^{\dot{\alpha}\dot{\beta}}R_{i\bar{\jmath}k\bar{l}} \tag{2.34}$$

and

$$\mathcal{D}_{i}C_{jlm}\lambda_{\alpha}^{iI}\lambda_{\beta}^{jK}\epsilon^{\alpha\beta}\lambda_{\gamma}^{lJ}\lambda_{\delta}^{mL}\epsilon^{\gamma\delta}\epsilon_{IJ}\epsilon_{KL} \tag{2.35}$$

Note that, because of eq. (2.18), all couplings in the lagrangian are expressed through the tensors C_{ijk}.

From a tensor calculus point of view, all quartic terms but the last come from the equations of motion of the Y_{IJ}^{i} auxiliary field triplet [3].

13.2.4 2.4 Positivity and monodromies

Let us consider a submanifold $\mathcal{M}_{r}$ of the moduli space of a Riemann surface of genus r such that its tangent space is isomorphic to the Hodge bundle. In particular the dimension of $\mathcal{M}_{r}$ is equal to the genus r of the Riemann surface $\mathcal{C}_{r}{}^{\dagger}$. In this case, decomposing an abelian differential in terms of the $2r$ harmonic forms dual to the canonical basis of cycles, we have

$$\omega=X^{A}\left(z^{i}\right)\alpha_{A}+F_{A}\left(z^{i}\right)\beta^{A}\quad A,i=1,\ldots,r$$
$$\int\alpha_{A}\wedge\beta^{B}=\delta_{A}^{B}\quad,\int\alpha_{A}\wedge\alpha_{B}=\int\beta_{A}\wedge\beta_{B}=0 \tag{2.36}$$

where z^{i} are coordinates on the moduli space submanifold, and

$$\partial_i \omega = \partial_i X^A \alpha_A + \partial_i F_A \beta^A \tag{2.37}$$

Then the metric, given by the norm

$$g_{\bar{i}j} = i \int \partial_i \omega \wedge \partial_{\bar{j}} \bar{\omega} = i \partial_i \partial_{\bar{j}} \int \omega \wedge \bar{\omega} \tag{2.38}$$

is manifestly positive. Using eqs. (2.36), (2.37) we find

$$g_{i\bar{j}} = i \partial_i \partial_{\bar{j}} \left(\bar{F}_A X^A - \bar{X}^A F_A \right)$$

which coincides with the metric of $N = 2$ rigid special geometry (2.15) [1.[4].

Formula (2.37) implies by supersymmetry a similar expansion for the full multiplet (2.29). For the upper component $\mathcal{F}_{\mu\nu}^{-A}$ we get a self dual three form

$$w = \mathcal{F}^A \alpha_A + G_A \beta^A \tag{2.39}$$

on $\mathrm{R}_4 \times \mathcal{C}_r$ when (2.8) hold. We observe that an $N = 2, 4D$ abelian vector multiplet can be obtained from dimensional reduction from six dimensions either of a vector multiplet or of a tensor multiplet containing a self-dual field strength. This remarkable coincidence actually suggests a physical picture for the characterization of this subclass $\mathcal{C}_r$ of Riemann surfaces. Namely, they should appear in the compactification on $\mathbf{R}_4 \times \mathcal{C}_r$ of $N = 1$ six-dimensional theory of a self interacting tensor multiplet.

As shown in ref. [4], the Picard-Fuchs equations for $\mathcal{C}_r$ have a general form dictated by the differential constraints of rigid special geometry. A general proposal for $\mathcal{C}_r$ has been given in [2] and can be used to write down the Picard-Fuchs equations for the periods and to determine their monodromies. Such proposal can be checked by comparing the explicit form of the Picard-Fuchs equations with their general form given by rigid special geometry.

In the one parameter case ($G = SU(2)$), where $\mathcal{C}_1$ is given by the elliptic curve of ref. [1], the special geometry equations reduce to one ordinary second order equation

$$\left(\frac{d}{dz} + \widehat{\Gamma} \right) C^{-1} \left(\frac{d}{dz} - \widehat{\Gamma} \right) U = 0 \tag{2.40}$$

where $\widehat{\Gamma} = \frac{d}{dz} \log e, e = \frac{dX}{dz}$ and C is the 3 -tensor appearing in (2.14). This agrees with the Picard-Fuchs equations derived from $\mathcal{C}_1$. The general solution of this equation is [4]

$$U = \left(e, e \frac{d^2 F}{dX^2} \right) \tag{2.41}$$

with $\tau = \frac{d^2 F}{dX^2}$ being the uniformizing variable for which the differential equation reduces to $\frac{d^2}{d\tau^2}() = 0$.

13.3 3 Coupling to gravity

13.3.1 3.1 Special geometry and symplectic transformations

The coupling to gravity modifies the constraints of rigid special geometry because of the introduction of a $U(1)$ connection due to the $U(1)$ Kähler -Hodge structure of moduli space. For n vector multiplets one introduces $2(n+1)$ covariantly holomorphic sections [26, 23, 27, 29]

$$V = \left(L^\Lambda, M_\Lambda \right) \quad (\Lambda = 0, \ldots, n) \tag{3.1}$$

where 0 is the graviphoton index.

The new differential constraints of special geometry are

$$\begin{aligned} U_i &\equiv \left(\mathcal{D}_i L^\Lambda, \mathcal{D}_i M_\Lambda\right) = \left(f_i^\Lambda, h_{i\Lambda}\right) \\ \mathcal{D}_i U_j &= i C_{ijk} g^k \bar{U}_{\bar{l}} \\ \mathcal{D}_i \bar{U}_{\bar{j}} &= g_i \bar{V} \\ \mathcal{D}_i \bar{V} &= 0 \end{aligned} \tag{3.2}$$

where now $\mathcal{D}_i$ is the covariant derivative with respect to the usual Levi-Civita connection and the Kähler connection $\partial_i K$. That is, under $K \to K + f + \bar{f}$ a generic field ψ^i which under $U(1)$ transforms as $\psi^i \to e^{-\left(\frac{p}{2}f + \frac{\bar{p}}{2}\bar{f}\right)}\psi^i$ has the following covariant derivative

$$\mathcal{D}_i \psi^j = \partial_i \psi^j + \Gamma^j_{ik}\psi^k + \frac{p}{2}\partial_i K \psi^j \tag{3.3}$$

and analogously for $\mathcal{D}_{\bar{\imath}}$ with $p \to \bar{p}$. This $U(1)$ is related to the $U(1)$ in the $N = 2$ superconformal group, and the weights for all the fields were determined in 331 ($\bar{p} = c$). In our notations, $\left(L^\Lambda, M_\Lambda\right)$ have been given conventionally weights $p = -\bar{p} = 1$.

Since L^Λ, M_Λ are covariantly holomorphic, it is convenient to introduce holomorphic sections $X^\Lambda = e^{-K/2}L^\Lambda, F_\Lambda = e^{-K/2}M_\Lambda$.

The Kähler potential is fixed by the condition [3] [26]

$$i\left(\bar{L}^\Lambda M_\Lambda - L^\Lambda \bar{M}_\Lambda\right) = 1 \tag{3.4}$$

to be

$$K = -\log i\left(\bar{X}^\Lambda F_\Lambda - X^\Lambda \bar{F}_\Lambda\right) \tag{3.5}$$

As it is well known [3] [32], the differential constraints (3.2) can in general be solved in terms of a holomorphic function homogeneous of degree two $F(X)$. However, as we will see in the sequel, there exist particular symplectic sections for which such prepotential F does not exist. In particular this is the case appearing in the effective theory of the $N = 2$ heterotic string. For this reason it is convenient to have the fundamental formulas of special geometry written in a way independent of the existence of F.

First of all we note that quite generally we may write

$$M_\Lambda = \mathcal{N}_{\Lambda\Sigma}L^\Sigma; \quad h_{\Lambda i} = \overline{\mathcal{N}}_{\Lambda\Sigma}f_i^\Sigma \tag{3.6}$$

From (3.6) we can define the two $(n+1)\times(n+1)$ matrices

$$h_{\Lambda \bar{I}} = \left(h_{\Lambda\bar{0}} \equiv M_\Lambda, h_{\Lambda\bar{\imath}}\right) \quad , \quad f_{\bar{I}}^\Lambda = \left(f_{\bar{0}}^\Lambda \equiv L^\Lambda, f_{\bar{\imath}}^\Lambda\right) \tag{3.7}$$

to obtain an explicit expression for $\mathcal{N}_{\Lambda\Sigma}$ in terms of $\left(L^\Lambda, M_\Lambda\right)$ as

$$\mathcal{N}_{\Lambda\Sigma} = h_{\Lambda\bar{I}}\left(f^{-1}\right)^{\bar{I}}_{\Sigma} \tag{3.8}$$

Note that $h_{\Lambda\bar{I}}, f_{\bar{I}}^\Sigma$ are invertible matrices and the above expression implies the transformation law (2.10).

When F exists, $\mathcal{N}_{\Lambda\Sigma}$ has the form [3] 27]

$$\mathcal{N}_{\Lambda\Sigma} = \bar{F}_{\Lambda\Sigma} + 2i\frac{\left(\operatorname{Im}F_{\Lambda\Gamma}\right)\left(\operatorname{Im}F_{\Sigma\Pi}\right)L^\Gamma L^\Pi}{\left(\operatorname{Im}F_{\Xi\Omega}\right)L^\Xi L^\Omega} \tag{3.9}$$

which turns out to be the coupling matrix appearing in the kinetic term of the vector fields. However, as we show below, (3.6) are symplectic covariant and therefore they always hold even in some specific coordinate system in which F does not exist.

In the same way as in the rigid case, from eqs. (3.2) and (3.4) we find

$$g_{i\bar{\jmath}} = i\left(f_i^{\Lambda}\bar{h}_{\bar{\jmath}\Lambda} - h_{i\Lambda}\bar{f}_{\bar{\jmath}}^{\Lambda}\right) = i\left(\mathcal{N}_{\Lambda\Sigma} - \overline{\mathcal{N}}_{\Lambda\Sigma}\right) f_i^{\Lambda} f_{\bar{\jmath}}^{\Sigma} \tag{3.10}$$

$$C_{ijk} = f_i^{\Lambda}\mathcal{D}_j h_{k\Lambda} - h_{i\Lambda}\mathcal{D}_j f_k^{\Lambda} = f_i^{\Lambda}\partial_j\overline{\mathcal{N}}_{\Lambda\Sigma} f_k^{\Sigma} \tag{3.11}$$

which are symplectic invariant. (Note that $\mathcal{N}_{\Lambda\Sigma}$ has zero Kähler weight).
Furthermore, the integrability conditions (3.2) give [26 [25 230 29 270

$$R_{i\bar{\jmath}l\bar{k}} = g_{i\bar{\jmath}}g_{l\bar{k}} + g_{i\bar{k}}g_{l\bar{\jmath}} - C_{ilp}C_{\bar{\jmath}\bar{k}\bar{p}}g^{p\bar{p}} \tag{3.12}$$

replacing eq. (2.6).
Here C_{ilp} is a covariantly holomorphic tensor of weight $p = -\bar{p} = 2$,

$$\mathcal{D}_{\bar{l}}C_{ijk} = \partial_{\bar{l}}C_{ijk} - \partial_{\bar{l}}KC_{ijk} = 0 \tag{3.13}$$

which implies $\partial_{\bar{l}}W_{ijk} = 0$ with $C_{ijk} = e^K W_{ijk}$.
Some additional consequences of the previous formulae are the following: from $\mathcal{D}_i F_\Lambda = \overline{\mathcal{N}}_{\Lambda\Sigma}\mathcal{D}_i X^{\Sigma}$, applying $\mathcal{D}_{\bar{\jmath}}$ to both sides we also find

$$\mathcal{D}_{\bar{\jmath}}\mathcal{D}_i F_\Lambda = \partial_{\bar{\jmath}}\overline{\mathcal{N}}_{\Lambda\Sigma}\mathcal{D}_i X^{\Sigma} + \overline{\mathcal{N}}_{\Lambda\Sigma}\mathcal{D}_{\bar{\jmath}}\mathcal{D}_i X^{\Sigma} \tag{3.14}$$

which implies, using the third line of (3.2),

$$\left(F_\Lambda - \overline{\mathcal{N}}_{\Lambda\Sigma}X^{\Sigma}\right) g_{i\bar{\jmath}} = \partial_{\bar{\jmath}}\overline{\mathcal{N}}_{\Lambda\Sigma}\mathcal{D}_i X^{\Sigma} \tag{3.15}$$

Note that the left-hand side of (3.15) defines the graviphoton projector

$$T_\Lambda = M_\Lambda - \overline{\mathcal{N}}_{\Lambda\Sigma}L^{\Sigma} \tag{3.16}$$

From the first of equations (3.6) it also follows that

$$\partial_{\bar{\imath}}\mathcal{N}_{\Lambda\Sigma}L^{\Sigma} = 0 \quad , \quad h_{i\Lambda} = \mathcal{N}_{\Lambda\Sigma}f_i^{\Sigma} + \partial_i\mathcal{N}_{\Lambda\Sigma}L^{\Sigma} \tag{3.17}$$

and therefore

$$\partial_i\mathcal{N}_{\Lambda\Sigma}L^{\Sigma} = \left(\overline{\mathcal{N}}_{\Lambda\Sigma} - \mathcal{N}_{\Lambda\Sigma}\right) f_i^{\Sigma} \tag{3.18}$$

by contraction with $f_{\bar{\jmath}}^{\Lambda}$ we get

$$f_{\bar{\jmath}}^{\Lambda}\partial_i\mathcal{N}_{\Lambda\Sigma}L^{\Sigma} = ig_{i\bar{\jmath}} \tag{3.19}$$

Taking the complex conjugate of (3.19) and using (3.15) it follows that

$$T_\Lambda\bar{L}^{\Lambda} = -i \tag{3.20}$$

which is nothing but (3.4). An alternative form for the Kähler potential is

$$K = -\log i\left(\mathcal{N}_{\Lambda\Sigma} - \overline{\mathcal{N}}_{\Lambda\Sigma}\right) X^{\Lambda}\bar{X}^{\Sigma} \tag{3.21}$$

Duality transformations are now in $Sp(2n+2,\mathbb{Z})$ and act on X^Λ, F_Λ as in the rigid case. The symplectic action on $\left(L^\Lambda, M_\Lambda\right)$ (or $\left(X^\Lambda, F_\Lambda\right)$) is

$$\begin{pmatrix} L \\ M \end{pmatrix}' = \begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} L \\ M \end{pmatrix} = \mathcal{S}\begin{pmatrix} L \\ M \end{pmatrix} \quad \mathcal{S} \in Sp(2n+2,\mathbb{Z}) \tag{3.22}$$

Then it follows, because of eq. (3.2) and (3.6),

$$\begin{pmatrix} f_i^\Lambda \\ h_{i\Lambda} \end{pmatrix}' = \begin{pmatrix} A & B\overline{\mathcal{N}} \\ C & D\overline{\mathcal{N}} \end{pmatrix} \begin{pmatrix} f_i^\Lambda \\ f_i^\Lambda \end{pmatrix} \tag{3.23}$$

which implies again (2.10). These two transformations laws imply the covariance of (3.6).

The symplectic action on $\mathcal{F}_{\mu\nu}^{+\Lambda}, G_{+\Lambda}^{\mu\nu}$ is the same as on $\left(L^\Lambda, M_\Lambda\right)$, so eq. (2.8) is unchanged. Therefore the discussion of the previous section on perturbative and non perturbative duality transformations in the rigid case remains unchanged when gravity is turned on.

When the sections $\left(X^\Lambda, F_\Lambda\right)$ are chosen in such a way that a function F exists *, from (3.4) and the degree two homogeneity of F it follows that [26] 27]

$$\operatorname{Im} F_{\Lambda\Sigma} L^\Lambda \overline{f}_{\bar{i}}^\Sigma = 0 \tag{3.24}$$

so that the second of eq. (3.6) becomes $h_{i\Lambda} = F_{\Lambda\Sigma} f^{\Sigma i}$. Furthermore from (3.11) and (3.24) it also follows

$$e^{K/2} C_{ijk} = f_i^\Lambda f_j^\Gamma f_k^\Sigma F_{\Lambda\Gamma\Sigma} \tag{3.25}$$

By the same token, we have

$$\begin{pmatrix} f_i^\Lambda \\ h_{i\Lambda} \end{pmatrix}' = \begin{pmatrix} A & B\mathcal{F} \\ C & D\mathcal{F} \end{pmatrix} \begin{pmatrix} f_i^\Lambda \\ f_i^\Lambda \end{pmatrix} \tag{3.26}$$

where $\mathcal{F} = F_{\Lambda\Sigma}$. Note that in these cases

$$\begin{aligned} 2\widetilde{F}(\widetilde{X}) = \widetilde{F}_\Lambda \widetilde{X}^\Lambda = \\ 2F + 2X^\Lambda \left(C^T B\right)_\Lambda^{\ \Sigma} F_\Sigma + X^\Lambda \left(C^T A\right)_{\Lambda\Sigma} X^\Sigma + F_\Lambda \left(D^T B\right)^{\Lambda\Sigma} F_\Sigma \end{aligned} \tag{3.27}$$

Note also that the homogeneity of F implies

$$\widetilde{X} = (A + B\mathcal{F})X \tag{3.28}$$

where $\mathcal{F} = F_{\Lambda\Sigma}$ and

$$\widetilde{F} = (C + D\mathcal{F})X \tag{3.29}$$

Special coordinates in supergravity are defined by $t^\Lambda = X^\Lambda / X^0$ since we now have a set of $n+1$ homogeneous coordinates. If we assume that $\mathcal{D}_i \left(\frac{X^\Lambda}{X^0}\right)$ is an invertible matrix, then we may choose a frame for which $\partial_i \left(\frac{X^\Lambda}{X^0}\right) = \delta_i^\Lambda$. This is possible only if X^Λ are unconstrained variables and so $F_\Lambda = F_\Lambda(X)$, which implies $F_\Lambda = \partial_\Lambda F(X)$ with F homogeneous of degree 2.

A résumé of the duality transformations for this case, including the supergravity corrections has been given in appendix C of 32.

We now discuss the possible non-existence of $F(X)$. If we start with some special coordinates $X^\Lambda, F_\Lambda(X)$, it is possible that in the new basis the $\widetilde{X}^\Lambda$ are not good special coordinates in the sense that the mapping $X \to \widetilde{X}$ is not invertible. This happens whenever the $(n+1) \times (n+1)$ matrix $A + B\mathcal{F}$ is not invertible (its determinant vanishes). This does not mean that $\widetilde{X}, \widetilde{F}$ are not good symplectic sections since the symplectic matrix $\mathcal{S} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is always invertible. It simply means that $\widetilde{F}_\Lambda \neq \widetilde{F}_\Lambda(\widetilde{X})$ and therefore a prepotential $\widetilde{F}(\widetilde{X})$ does not exist. However our formulation of special geometry never explicitly used the fact that F_Λ be a functional of the X 's and indeed the quantities $\left(X^\Lambda, F_\Lambda\right), \left(f_i^\Lambda, h_{i\Lambda}\right), \mathcal{N}_{\Lambda\Sigma}$ and $C_{ijk}, g_{i\bar{j}}$ are well defined for any choice of the symplectic sections $\left(X^\Lambda, F_\Lambda\right)$ since they are symplectic invariant or covariant. For example, to compute the "gauge coupling" $\widetilde{\mathcal{N}}$ in such a basis $(\widetilde{X}^\Lambda, \widetilde{F}_\Lambda)$ one uses the formula

$$\widetilde{\mathcal{N}}(\widetilde{X}, \widetilde{F}) = (C + D\mathcal{N}(X))(A + B\mathcal{N}(X))^{-1} \tag{3.30}$$

and expresses the $X = X(\tilde{X}, \widetilde{F})$ by using the fact that the symplectic mapping can be inverted. All other quantities can be computed in this way.

We will see the relevance of this observation in the sequel, while discussing low energy effective action of $N = 2$ heterotic string. A simple example is the following. Consider $F = iX^0X^1$, leading to

$$\mathcal{N} = \begin{pmatrix} i\frac{X^1}{X^0} & 0 \\ 0 & i\frac{X^0}{X^1} \end{pmatrix} \tag{3.31}$$

This appears in the $N = 2$ reduction of pure $N = 4$ supergravity in the so-called $SO(4)$ formulation [33]. Consider now the symplectic mapping defined by

$$A = D = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}; \quad C = -B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.32}$$

Then the transformation is

$$\begin{aligned} \widetilde{X}^0 &= X^0 \widetilde{X}^1 = -F_1 \\ \widetilde{F}_0 &= F_0 \widetilde{F}_1 = X^1 \end{aligned} \tag{3.33}$$

Using in the first line $F_1 = iX^0$ would lead to a non-invertible mapping $X \to \widetilde{X}$, and using (3.27) would lead to $\widetilde{F} = 0$. One observes also that $A + B\mathcal{F}$ is non-invertible. However, $A + B\mathcal{N}$ is invertible, and one obtains $\widetilde{\mathcal{N}} = iX^1 \left(X^0\right)^{-1} \mathbb{1} = i\widetilde{F}_1 \left(\widetilde{X}^0\right)^{-1} \mathbb{1}$. This form appears in the $N = 2$ reduction of the $SU(4)$ formulation of pure $N = 4$ supergravity 34]. These two forms of the $N = 2$ reduced action and the duality transformation have been studied in (35] to relate electric and magnetic charges of black holes.

13.3.2 3.2 The fermionic sector

As far as the fermions are concerned, the vector $N = 2$ multiplet is now

$$\left(L^\Lambda, f_i^\Lambda \lambda^{iI}_\alpha, \mathcal{F}^{-\Lambda}_{\alpha\beta}\right) \tag{3.34}$$

The tensor $\mathcal{T}^{\bar\imath}_{\alpha\beta}$ is still the same as in (2.30), and

$$Q^\Lambda_{\bar\imath} = \mathcal{D}_{\bar\imath} \bar{L}^\Lambda; \quad P_{\Lambda\bar\imath} = \mathcal{D}_{\bar\imath} \bar{M}_\Lambda \tag{3.35}$$

Correspondingly, the gaugino Pauli terms have the form

$$i\left(\mathcal{D}_{\bar\imath} \bar{L}^\Lambda G^{\alpha\beta}_{b-\Lambda} - \mathcal{D}_{\bar\imath} \bar{M}_\Lambda \mathcal{F}^{-\Lambda\alpha\beta}\right) \mathcal{T}^{\bar\imath}_{\alpha\beta} \tag{3.36}$$

quite analogous to eq. (2.32).

Gravitino Pauli and quartic terms [3] 30 27] are defined by the formulas (2.21) and (2.28) with *

$$\begin{aligned} Q^\Lambda &= L^\Lambda; \quad P_\Lambda = M_\Lambda \\ \mathcal{T}^{\mu\nu} &= k_1 \bar\psi^I_\rho \psi^J_\sigma \epsilon_{IJ} \epsilon^{\mu\nu\rho\sigma} \end{aligned} \tag{3.37}$$

for the purely gravitino terms, in which case the index a of the general treatment is obsolete. For the mixed gaugino-gravitino Pauli terms we use

$$Q_{\bar{\imath}}^{\Lambda} = \mathcal{D}_{\bar{\imath}} \bar{L}^{\Lambda}; \quad P_{\Lambda\bar{\imath}} = \mathcal{D}_{\bar{\imath}} \bar{M}_{\Lambda}$$
$$\mathcal{T}_{\alpha\beta}^{\bar{\imath}} = k_2 \bar{\lambda}_I \gamma_\rho \psi_{\sigma J} \epsilon^{IJ} \epsilon^{\mu\nu\rho\sigma} \tag{3.38}$$

and the index $\bar{\imath}$ plays again the role of a. The constants k, k_1 and k_2 should also be fixed by supersymmetry. So, as before, the unique quartic terms are generated by requiring duality invariance of the action. Of course many of these terms are absent in $N = 1$ [30] theories because of the absence of the second gravitino. This is one of the differences between rigid supersymmetry and local supersymmetry. What happens is that in $N = 2$ supergravity, one introduces an extra $\left(\frac{3}{2}, 1\right)$ multiplet, with respect to the $N = 1$ case. This has the effect of having extra auxiliary fields in the supergravity multiplet 36

$$\mathcal{V}_{J\mu}^{I}, A_\mu, T_{\mu\nu}^{-}, D \tag{3.39}$$

⋆ The Kähler weights of the fermions are $p = -\bar{p} = \frac{1}{2}$ for $\psi_{\mu I}$, and $p = -\bar{p} = -\frac{1}{2}$ for λ^{iI}. The scalars and the fermions of the hypermultiplets, not discussed here, have respectively Kähler weights $p = \bar{p} = 0$ and $p = -\bar{p} = -\frac{1}{2}$.
other than the matter auxiliary field of the vector multiplet Y^{iIJ} (traceless, real, symmetric in IJ), $i, j = 1, 2$, i.e. a real $SU(2)$ triplet. The meaning of the auxiliary fields is straightforward. The Y' 's correspond to the three auxiliary fields of a $N = 1$ vector multiplet and a chiral multiplet. The D auxiliary field gives the equation (3.4) (i.e (3.5)), $T_{\mu\nu}^{-}$is the graviphoton (symplectic invariant) combination of the gauge fields $T_{\mu\nu}^{-} = T_\Lambda \mathcal{F}_{\mu\nu}^{-\Lambda}$, and $\mathcal{V}_{J\mu}^{I}, A_\mu$ are the composite $SU(2)$ and $U(1)$ connections of the quaternionic manifold and Kähler-Hodge manifold respectively. Note that comparison between $N = 1$ and $N = 2$ theories shows that the spinors χ^i of the scalar multiplet and λ^Σ of the vector multiplet of the $N = 1$ theory are related to the doublet λ^{iI} of the $N = 2$ theory by

$$\chi^i = \lambda^{i1} \quad , \lambda^\Sigma = f_i^\Sigma \lambda^{i2} \tag{3.40}$$

13.3.3 3.3 The three-form cohomology

We recall that special geometry in $N = 2$ supergravity, unlike rigid special geometry, is suitable for three-form cohomology for Calabi-Yau manifolds. Let's define a holomorphic three-form 25,23

$$\Omega = X^\Lambda \alpha_\Lambda + F_\Lambda \beta^\Lambda \tag{3.41}$$

where $\alpha_\Lambda, \beta^\Lambda$ is a $2n + 2$ dimensional cohomology basis dual to the $2n + 2$ homology cycles $(n = h_{21})$.Ω is a holomorphic section of a line bundle. Then it follows that if one defines

$$e^{-K} = i \int \Omega \wedge \bar{\Omega} > 0 \tag{3.42}$$

then

$$g_{i\bar{\jmath}} = \frac{-i \int \mathcal{D}_i \Omega \wedge \mathcal{D}_{\bar{\jmath}} \bar{\Omega}}{i \int \Omega \wedge \bar{\Omega}} = -\partial_i \partial_{\bar{\jmath}} \log i \int \Omega \wedge \bar{\Omega} > 0 \tag{3.43}$$

The $(2n+2)$ three-forms $\mathcal{D}_i\Omega, \mathcal{D}_i\bar{\Omega}, \Omega, \bar{\Omega}$ with the cohomology basis $\left(\alpha_\Lambda, \beta^\Lambda\right)$ correspond to the decomposition

$$H^3(\mathbf{R}) = H^{(2,1)}(\mathbb{C}) + H^{(1,2)}(\mathbb{C}) + H^{(3,0)}(\mathbb{C}) + H^{(0,3)}(\mathbb{C}) \tag{3.44}$$

Note that since $\Omega = \left(X^\Lambda, F_\Lambda\right)$, then $\mathcal{D}_i\Omega = \left(\mathcal{D}_i X^\Lambda, \mathcal{D}_i F_\Lambda\right)$, with $f_i^\Lambda = e^{\frac{K}{2}}\mathcal{D}_i X^\Lambda, h_{i\Lambda} = e^{\frac{K}{2}}\mathcal{D}_i F_\Lambda$. The relations

$$\int \Omega \wedge \Omega = \int \Omega \wedge \mathcal{D}_i\bar{\Omega} = 0 \tag{3.45}$$

are obvious since $\mathcal{D}_i\Omega = \partial_i\Omega - \frac{1}{(\Omega,\bar{\Omega})}\left(\partial_i\Omega, \bar{\Omega}\right)\Omega$. However the relation

$$\int \Omega \wedge \mathcal{D}_i\Omega = 0 \tag{3.46}$$

which is suitable for three-form cohomology, implies

$$\int \Omega \wedge \partial_i\Omega = 0 \tag{3.47}$$

i.e.

$$\partial_i X^\Lambda F_\Lambda - \partial_i F_\Lambda X^\Lambda = 0 \tag{3.48}$$

for any choice of the symplectic section. Eq. (3.48) is equivalent to

$$X^\Lambda \mathcal{D}_i F_\Lambda - \mathcal{D}_i X^\Lambda F_\Lambda = 0 \tag{3.49}$$

13.3.4 3.4 Duality transformations in $N = 1$ locally supersymmetric Yang-Mills theories

In $N = 1$ super Yang-Mills theories coupled to supergravity [30], duality transformations are implemented as follows. Define the symplectic $Sp(2r)$ vectors

$$\begin{aligned} \mathcal{V} &= \left(\mathcal{F}^{-A}_{\mu\nu}, G^{-\mu\nu}_A = i\frac{\partial\mathcal{L}}{\partial\mathcal{F}^{-A}_{\mu\nu}}\right) \\ \mathcal{U}_\alpha &= \left(\lambda^A_\alpha, f_{AB}(z)\lambda^B_\alpha\right) \end{aligned} \tag{3.50}$$

where $\left(\lambda^A, \mathcal{F}^{-A}_{\mu\nu}\right)$ is the vector field strength multiplet and $f_{AB}(z)$ is the holomorphic coupling introduced in [30] *, which depends on the scalars of chiral multiplets, and which plays here the role of $\overline{\mathcal{N}}_{AB}$ in the general treatment of sections 2.1 and 2.2. Then the $N = 1$ supergravity lagrangian is invariant under the symplectic transformations

$$\mathcal{V} \to \mathcal{S}\mathcal{V}, \quad \mathcal{U} \to \mathcal{S}\mathcal{U} \quad , \quad f \to (C + Df)(A + Bf)^{-1} \quad , \quad \mathcal{S} \in Sp(2, r; \mathbf{R}) \tag{3.51}$$

This is best seen using the $N = 1$ tensor calculus (or superfield) notation of ref. [30]. The part of the action which contains the field strength chiral multiplet

$$W^A_\alpha = T\left(\mathcal{D}_\alpha V^A\right) \tag{3.52}$$

We replaced the f in 30] by $2if$.

where T is the generalisation of to local supersymmetry of the chiral projection $\overline{DD}$ (similar to the operation obtaining kinetic multiplets introduced in 37]), can be written in first order form by introducing an unconstrained chiral multiplet W^A_α and a (vector) real lagrangian multiplier U_A (f_{AB} is a chiral superfield)

$$4 \operatorname{Im} W_\alpha^A \mathcal{D}_\beta U_A \epsilon^{\alpha\beta}\Big|_D + i f_{AB}(z) W_\alpha^A W_\beta^B \epsilon^{\alpha\beta}\Big|_F \tag{3.53}$$

Variation with respect to U_A yields the Bianchi identity

$$\mathcal{D}^\alpha W_\alpha^A = \overline{\mathcal{D}}_{\dot\alpha} \bar W^{\dot\alpha A} \tag{3.54}$$

which is solved by

$$W_\alpha^A = T\left(\mathcal{D}_\alpha V^A\right) \tag{3.55}$$

which leads to the original form of the action. The dual form of the theory is obtained, in a manner analogous to the rigid case [1], by varying the same lagrangian with respect to W_α^A. Defining $W_{\alpha A}^{(D)} \equiv T\left(\mathcal{D}_\alpha U_A\right)$, and using the fact that the first term in (3.53) can also be written as $-\,2i W_\alpha^A W_{\beta B}^{(D)} \epsilon^{\alpha\beta}\Big|_F$, yields

$$W_\alpha^A = \left(f^{-1}\right)^{AB} W_{\alpha B}^{(D)} \tag{3.56}$$

which implies the Bianchi identity also for $W^{(D)}$. The dual lagrangian is

$$\mathcal{L}^D = -\,i\left(f^{-1}\right)^{AB} W_{\alpha A}^{(D)} W_{\beta B}^{(D)} \epsilon^{\alpha\beta}\Big|_F \tag{3.57}$$

This realises the symplectic transformation of (3.51) with $B = -C = \mathbb{1}$ and $A = D = 0$.

A duality rotation is a symmetry if for some coordinate changes $z \to \widetilde{z}$ (z is the first component of a chiral multiplet)

$$\widetilde{f}_{AB}(\widetilde{z}) = f_{AB}(\widetilde{z}) \tag{3.58}$$

and the superpotential W is a symplectic invariant section of a Hodge bundle, i.e.

$$\|\widetilde{W}(\widetilde{z})\|^2 = \|W(\widetilde{z})\|^2 \tag{3.59}$$

where $\|W(z)\|^2 = |W(z)|^2 e^K \equiv e^G$. In component form, we can exhibit the symplectic invariance of the gaugino kinetic term and the Pauli terms by noticing that they can be written as

$$\begin{aligned}
e^{-1}\mathcal{L}_{\text{kin}}\left(\lambda, \bar\lambda\right) &= i\boldsymbol{\mathcal{U}}_\alpha \Omega \left(\sigma^\mu\right)^{\alpha\dot\alpha} \mathcal{D}_\mu \overline{\boldsymbol{\mathcal{U}}}_{\dot\alpha} \\
e^{-1}\mathcal{L}_{\text{Pauli}}\left(\psi, \lambda\right) &= \operatorname{Im}\left(\overline{\boldsymbol{\mathcal{U}}}_{\dot\alpha} \Omega \left(\sigma^\mu\right)^{\dot\alpha\beta} \mathcal{V}_{b\beta\gamma} \psi_\mu^\gamma\right) \\
e^{-1}\mathcal{L}_{\text{Pauli}}\left(\chi, \lambda\right) &= \operatorname{Im}\left(\partial_i f_{AB} \lambda_\alpha^A \bar\chi_\beta^i \mathcal{F}^{-B\alpha\beta}\right)
\end{aligned} \tag{3.60}$$

where Ω is the symplectic metric $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ (such that $\mathcal{S}^T \Omega \mathcal{S} = \Omega$) and $\mathcal{V}_b$ is the bare $\mathcal{V}$ (only bosonic part).

The (ψ, λ) Pauli term can be written in the form as in (2.21) and we identify in (2.26) the symplectic vector (Q, P) with $\overline{\boldsymbol{\mathcal{U}}}_{\dot\alpha}$, and

$$\mathcal{T}_{\beta\gamma}^{\dot\alpha} = -\frac{1}{2}\left(\sigma^\mu\right)^{\dot\alpha}{}_\beta \psi_{\mu\gamma}.$$

The last Pauli term, $e^{-1}\mathcal{L}_{\text{Pauli}}\left(\chi, \lambda\right)$, has the form (2.21), with

$$H_{A\alpha\beta} = \frac{1}{2}\partial_i f_{AB} \lambda_\alpha^B \bar\chi_\beta^i$$

This we rewrite in the form (2.26)using the following identifications (note that $(\mathrm{Im}\, f)_{AB}$ is the matrix of the kinetic terms of the vectors, and is thus invertible)

$$\begin{aligned} Q^A_{i\alpha} &\equiv (\mathrm{Im}\, f)^{-1AB}\partial_i f_{BC}\lambda^C_\alpha; \quad P_{Ai\alpha} \equiv \bar{f}_{AB}Q^B_{i\alpha} \\ \mathcal{T}^{i\alpha}_{\beta\gamma} &= \tfrac{i}{4}\delta^\alpha_{(\beta}\chi^i_{\gamma)} \end{aligned}$$

To prove that these (Q,P) form a symplectic vector, one uses the following relations (which are in general true for f_{AB} replaced by $\overline{\mathcal{N}}_{AB}$):

$$\begin{aligned} \tilde{f} &= (C+Df)(A+Bf)^{-1} = \left(A^T + fB^T\right)^{-1}\left(C^T + fD^T\right) \\ \partial_i\tilde{f} &= D\partial_i f(A+Bf)^{-1} - (C+Df)(A+Bf)^{-1}B\partial_i f(A+Bf)^{-1} \\ &= \left(A^T + fB^T\right)^{-1}\partial_i f(A+Bf)^{-1} \\ \mathrm{Im}\,\tilde{f} &= \left(A^T + fB^T\right)^{-1}(\mathrm{Im}\, f)(A+B\bar{f})^{-1} \\ \widetilde{\lambda} &= (A+Bf)\lambda \end{aligned} \tag{3.61}$$

These formulas then give automatically quartic fermionic terms as discussed in section 2.

We observe that the requirements for having symplectic transformations, (3.58) and (3.59), are in principle weaker than what is necessary to have an $N=2$ theory.

13.4 4 Duality symmetries

13.4.1 4.1 The facts

Duality transformations in generic $N=2$ supergravity theories are a different choice of the symplectic representative $\left(X^\Lambda, F_\Lambda\right)$ of the underlying special geometry. If the fields $\mathcal{F}^{+\Lambda}_{\mu\nu}, G^+_{\Lambda\mu\nu}$have no electric or magnetic sources these dualities are simply a different equivalent choice of sections $\left(X^\Lambda, F_\Lambda\right)$ since they are defined up to a symplectic transformation [3]. However if the gauge fields are coupled to (abelian) sources then duality transformations map theories into different theories with a duality transformed source. Since the matrix $\mathcal{N}_{\Lambda\Sigma}$ plays the role of a coupling constant it is clear that in perturbation theory the only possible duality transformations are those with $B=0$ and have a lower triangular block form

$$\mathcal{S} = \begin{pmatrix} A & 0 \\ C & A^{T-1} \end{pmatrix} \tag{4.1}$$

Under such change, the action changes in a total derivative which, up to fermion terms, is

$$\mathcal{L}'(A,C) = \mathcal{L} + \mathrm{Im}\,\mathcal{F}^{-\Lambda}\left(C^T A\right)_{\Lambda\Sigma}\mathcal{F}^{-\Sigma} \tag{4.2}$$

So the lagrangian is invariant up to a surface term. A duality transformation is a symmetry if

$$\tilde{\mathcal{N}}(\widetilde{X}, \widetilde{F}) = \mathcal{N}(\widetilde{X}, \widetilde{F}) \tag{4.3}$$

If $F_\Lambda = F_\Lambda(X)$ this implies

$$\widetilde{F}(\widetilde{X}) = F(\widetilde{X}) \tag{4.4}$$

Then using (3.27) we should have [3] 38]

$$\begin{aligned} 2F[(A+B\mathcal{F})X] = {} & 2F + 2X^\Lambda\left(C^T B\right)_\Lambda{}^\Sigma F_\Sigma \\ & + X^\Lambda\left(C^T A\right)_{\Lambda\Sigma}X^\Sigma + F_\Lambda\left(D^T B\right)^{\Lambda\Sigma}F_\Sigma \end{aligned} \tag{4.5}$$

which is a functional relation for F given A, B, C, D. Note that because of (3.27) it may happen that $\widetilde{F}(\widetilde{X})=0$. This is so when $\frac{\partial \widetilde{X}^{\Lambda}}{\partial X^{\Sigma}}$ is not an invertible matrix.

13.4.2 4.2 Heterotic $N=2$ superstring theories

In $N=2$ heterotic string theories, as the one obtained by the fermionic construction or by compactification on $T_2 \times K_3$, one often encounters classical moduli spaces which are locally of the form [39] 40 (19] (41 42

$$\frac{O\left(2, n_v\right)}{O(2) \times O\left(n_v\right)} \times \frac{O\left(4, n_h\right)}{O(4) \times O\left(n_h\right)} \tag{4.6}$$

where n_v and n_h are respectively the number of the moduli in vector and hypermultiplets. If there are no charged massless hypermultiplets with respect to the gauge group $U(1)^r$, with $r=n_v$, we may avoid holomorphic anomalies 43 46] and the situation for this theory may be similar to the rigid Yang-Mills theory coupled to supergravity with an additional dilaton axion multiplet. According to the previous discussion, all perturbative duality symmetries are those for which the previous formula holds for a subgroup of lower triangular matrices

$$\begin{pmatrix} A & 0 \\ C & A^{T-1} \end{pmatrix} \tag{4.7}$$

with $A^T C$ symmetric.

The $(r+2) \times (r+2)$ block A contains the target space T duality and C contains the Peccei-Quinn axion symmetry [12] (for the definition of S in the $N=2$ context, see below)

$$S \rightarrow S+1 \tag{4.8}$$

These are the tree level stringy symmetries of the massive states with $M=|Z|$ where Z is the central charge of the $N=2$ supersymmetry algebra. If the number of T-moduli is r then the duality symmetries are in $Sp(2r+4;\mathbb{Z})$.

An important point is that we would like to make the tree level (string) symmetry manifest. This means that the gauge fields

$$\mathcal{A}_\mu^{\Lambda}=\left(G_\mu, B_\mu, \mathcal{A}_\mu^A\right) \quad A=2, \ldots, r+1 \tag{4.9}$$

(G_μ is the graviphoton and the B_μ is the vector of the dilaton-axion multiplet) should transform in the $2+r$ dimensional (vector) representation of the target space duality symmetry

$$\mathcal{A}^{\prime}=A \mathcal{A} ; \quad A^T \eta A=\eta ; \quad \eta_{\Lambda \Sigma}=\operatorname{Diag}(1,1,-1,-1, \ldots) \tag{4.10}$$

with $A \in O(2, r ; \mathbb{Z})$. Under the axion Peccei-Quinn symmetry $S \rightarrow S+1$

$$\mathcal{A}^{\Lambda^{\prime}}=\mathcal{A}^{\Lambda} \quad, \quad G_{\Lambda \mu \nu} \rightarrow G_{\Lambda \mu \nu}+\eta_{\Lambda \Sigma} \mathcal{F}_{\mu \nu}^{\Sigma} \tag{4.11}$$

where

$$\mathcal{N}_{\Lambda \Sigma}(S+1)=\mathcal{N}_{\Lambda \Sigma}(S)+\eta_{\Lambda \Sigma} \tag{4.12}$$

This formulation is directly obtained by $N=2$ reduction of the standard form of the $N=4$ supergravity action 12 with a moduli space of the type $O(6, r)/O(6) \times O(r)/\Gamma$ and duality group $\Gamma=O(6, r ; \mathbb{Z})$. However to get this in a standard $N=2$ supergravity form, one must introduce $2+r$ symplectic sections $\left(X^{\Lambda}, F_{\Lambda}\right)$ $(\Lambda=0,1, \ldots, r+1)$ for which $O(2, r)$ is block diagonal and the $S \rightarrow S+1$ shift is lower triangular. This formulation can be obtained by making a symplectic rotation, with $\mathcal{S}$ given by

$$\mathcal{S} = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1} & -\mathbb{1} \\ \mathbb{1} & \mathbb{1} \end{pmatrix} \tag{4.13}$$

from a representation in which only $O(2) \times O(r)$ is block diagonal [47, namely

$$\begin{aligned} O(2,r): \quad & \begin{pmatrix} A & 0 \\ 0 & \eta A \eta \end{pmatrix} = \mathcal{S} A_1 \mathcal{S}^{-1} \\ S \to S+1: \quad & \begin{pmatrix} \mathbb{1} & 0 \\ \eta & \mathbb{1} \end{pmatrix} = \mathcal{S} A_2 \mathcal{S}^{-1} \end{aligned} \tag{4.14}$$

where A_1, A_2 are the matrices given in ref. [47]. The new sections are given explicitly by eqs. (3.28), (3.29),

$$\begin{aligned} \widehat{X}^\Lambda &= \frac{1}{\sqrt{2}} \left(\delta_{\Lambda\Sigma} - F_{\Lambda\Sigma}\right) X^\Sigma \\ \widehat{F}_\Lambda &= \frac{1}{\sqrt{2}} \left(\delta_{\Lambda\Sigma} + F_{\Lambda\Sigma}\right) X^\Sigma \end{aligned} \tag{4.15}$$

where the function

$$F = -\sqrt{X_i^2}\sqrt{X_\alpha^2} \quad i = 0,1; \quad \alpha = 2, \dots, r+1 \tag{4.16}$$

was obtained in ref. 47]. From (4.15),(4.16) one can verify that the $\widehat{X}^\Lambda, \widehat{F}_\Lambda$ satisfy the constraints $\widehat{X}^\Lambda \eta_{\Lambda\Sigma} \widehat{X}^\Sigma = \widehat{F}_\Lambda \eta^{\Lambda\Sigma} \widehat{F}_\Sigma = \widehat{X}^\Lambda \widehat{F}_\Lambda = 0$. In particular, the new variables $\widehat{X}^\Lambda$ are not independent. The previous constraints imply that we may set

$$\widehat{F}_\Lambda = S \eta_{\Lambda\Sigma} \widehat{X}^\Sigma \tag{4.17}$$

and from eq. (3.27) we find $\widehat{F}(\widehat{X}) = 0$. Note that this is precisely the case for which $\widehat{F}_\Lambda = \widehat{F}_\Lambda \left(\widehat{X}^\Lambda\right)$ does not hold.

Since $O(2,r)$ is block diagonal, the new sections $\left(\widehat{X}^\Lambda, \widehat{F}_\Lambda\right)$ are $O(2,r)$ vectors. Recalling that the manifold $\frac{O(2,r)}{O(2)\times O(r)}$ can be described by the following equations

$$\begin{aligned} \eta_{\Lambda\Sigma} \Phi^\Lambda \Phi^\Sigma &= 0 \\ \eta_{\Lambda\Sigma} \Phi^\Lambda \bar{\Phi}^\Sigma &= 1 \end{aligned} \tag{4.18}$$

where Φ^Λ are coordinates in $CP(1,r)$, we may actually set

$$\Phi^\Lambda = \frac{\widehat{X}^\Lambda}{\sqrt{\widehat{X}^\Sigma \eta_{\Sigma\Pi} \bar{\widehat{X}}^\Pi}} \tag{4.19}$$

The Kähler potential is

$$K = -\log i \left(\widehat{X}^\Lambda \bar{\widehat{F}}_\Lambda - \bar{\widehat{X}}^\Lambda \widehat{F}_\Lambda \right) = -\log i(\bar{S} - S) - \log \widehat{X}^\Lambda \eta_{\Lambda\Sigma} \bar{\widehat{X}}^\Sigma \tag{4.20}$$

Under $S \to S+1$

$$\begin{aligned} \widehat{X}^\Lambda &\rightarrow \widehat{X}^\Lambda \\ \widehat{F}_\Lambda &\rightarrow \widehat{F}_\Lambda + \eta_{\Lambda\Sigma}\widehat{X}^\Sigma. \end{aligned} \tag{4.21}$$

In the same basis the (non-perturbative) inversion $S \rightarrow -\frac{1}{S}$ is given by the symplectic matrix $\begin{pmatrix} 0 & \eta \\ -\eta & 0 \end{pmatrix}$. This element, together with the one corresponding to $S \rightarrow S + 1$ generates an $Sl(2,\mathbb{Z})$ commuting with the $O(2r,\mathbb{Z})$ in $Sp(2r+4,\mathbb{Z})$. The inversion is actually the only symmetry generator with $B \neq 0$. It leaves invariant (4.20) up to a Kähler transformation and it will be a symmetry of the classical spectrum (as it comes by truncation of the $N = 4$ spectrum [12]) of electrically and magnetically charged states discussed in chapter 5.

The holomorphic sections $\widehat{X}^\Lambda$ can be written as [8]

$$\widehat{X}^\Lambda = \left(\frac{1}{2}\left(1 + y_\alpha^2\right), \frac{i}{2}\left(1 - y_\alpha^2\right), y^\alpha\right) \tag{4.22}$$

where the y^α are coordinates of the $O(2,r)/O(2)\times O(r)$ manifold. In terms of the Φ variables the kinetic matrix $\widehat{\mathcal{N}}_{\Lambda\Sigma}$ turns out to be [80 [10] [12]

$$\widehat{\mathcal{N}}_{\Lambda\Sigma}(\widehat{X}) = (S - \bar{S})\left(\Phi_\Lambda\bar{\Phi}_\Sigma + \bar{\Phi}_\Lambda\Phi_\Sigma\right) + \bar{S}\eta_{\Lambda\Sigma} \tag{4.23}$$

where $\Phi_\Lambda = \eta_{\Lambda\Sigma}\Phi^\Sigma$, and we will also further raise or lower indices with η.

Notice that (4.23) cannot be computed directly from (3.9) since in the new basis the denominator identically vanishes. On the other hand, one can use the formula (2.10), which in our case becomes

$$\widehat{\mathcal{N}}(\widehat{X},\widehat{F}) = (\mathbb{1} + \mathcal{N}(X))(\mathbb{1} - \mathcal{N}(X))^{-1} \tag{4.24}$$

and substitute for X^Λ the right hand side of the inverse transformations of (4.15)

$$\begin{aligned} X^\Lambda &= \frac{1}{\sqrt{2}}\left(\delta_{\Lambda\Sigma} + S\eta_{\Lambda\Sigma}\right)\widehat{X}^\Sigma \\ F_\Lambda &= \frac{1}{\sqrt{2}}\left(-\delta_{\Lambda\Sigma} + S\eta_{\Lambda\Sigma}\right)\widehat{X}^\Sigma \end{aligned} \tag{4.25}$$

Formula (4.23) is precisely what is obtained from $N = 4$ supergravity. Because of target space duality we expect that also the $\widehat{X}^\Lambda, \widehat{F}_\Lambda$ become, because of one loop corrections, a lower triangular representation of $Sp(2r+4,\mathbb{Z})$

$$\begin{pmatrix} \widehat{X}^\Lambda \\ \widehat{F}_\Lambda \end{pmatrix} \rightarrow \begin{pmatrix} A & 0 \\ A^{T-1}C & A^{T-1} \end{pmatrix} \begin{pmatrix} X^\Lambda \\ F_\Lambda \end{pmatrix} \tag{4.26}$$

where the matrix C comes from the monodromy of the one-loop term [1,2].

It is interesting to compute explicitly the coupling of the dilaton to the vector fields. The vector kinetic term is

$$\mathrm{Im}\,\overline{\mathcal{N}}_{\Lambda\Sigma}\mathcal{F}^{-\Lambda}_{\mu\nu}\mathcal{F}^{-\Sigma\mu\nu} = -2\,\mathrm{Im}\,\mathcal{N}_{\Lambda\Sigma}\mathcal{F}^{\Lambda}_{\mu\nu}\mathcal{F}^{\Sigma\mu\nu} + \mathrm{Re}\,\mathcal{N}_{\Lambda\Sigma}\mathcal{F}^{\Lambda\mu\nu}\widetilde{\mathcal{F}}^{\Sigma}_{\mu\nu} \tag{4.27}$$

and, in particular, setting in (4.22)$y^\alpha = 0$, it becomes

$$-2\,\mathrm{Im}\,S\left(\mathcal{F}^0\mathcal{F}^0 + \mathcal{F}^1\mathcal{F}^1 + \mathcal{F}^\alpha\mathcal{F}^\alpha\right) + \mathrm{Re}\,S\left(\mathcal{F}^0\widetilde{\mathcal{F}}^0 + \mathcal{F}^1\widetilde{\mathcal{F}}^1 - \mathcal{F}^\alpha\widetilde{\mathcal{F}}^\alpha\right) \tag{4.28}$$

We see that the dilaton couples in a universal way to the vectors while in the topological term we have a coupling with lorentzian signature.

13.4.3 4.3 Duality symmetries in $N > 2$ supergravities

The general considerations of section 2 about duality symmetries will apply to any higher $N > 2$ extended supergravity theory. Therefore, it is worth to briefly mention the implications of duality symmetries for some non-perturbative properties that these theories may exhibit. The important fact about $N > 2$ theories is that the scalar field space is (at least locally) a homogeneous symmetric space G/H, where G is some non compact subgroup of $Sp(2n)$ (n is the total number of vector fields existing in the theory). H is its maximal compact subgroup, as it must be for the kinetic matrix of the scalar field space to be positive definite.

On general grounds, we also know that the fields $\left(\mathcal{F}^{-A}, G_A^-\right)$must belong to a linear representation of G which is given by the decomposition of the ($2n$-dimensional) vector representation of $Sp(2n)$ under G. Thus, it is obvious that if this representation remains irreducible in G, the duality symmetry will necessarily mix electrically and magnetically charged states, since the $Sp(2n)$ vector $\left(n^A_{(m)} = 0, n^{(e)}_A\right)$ cannot be an invariant vector of G.

It is now a fact of life that the full duality (continuous) symmetry G of any $N > 2$ theory has a $2n$ dimensional representation which remains irreducible under $Sp(2n)$ (see table below [48]). This immediately implies that, if we assume, as conjectured in ref. [49], that the full $G(\mathbb{Z})$ is a symmetry of the dyonic states, then $G(\mathbb{Z})$ must be non-perturbative since the matrix B (see eq. (2.5)) in $G(\mathbb{Z})$ will not be vanishing. $N = 3, 5, 6$ supergravities can be obtained as low energy limits of $d = 4$ string models 50.

Another implication of this conjecture, for the case of $N = 4$ theories, is that, as pointed out in ref. [49], the spectrum of the BPS states of the ten dimensional heterotic string compactified on T_6 should be identical to the spectrum of the same states for type II strings compactified on $K_3 \times T_2$, since the full $N = 4$ BPS spectrum, invariant under $Sl(2;\mathbb{Z}) \times SO(6, n-6;\mathbb{Z})$ is completely fixed by supersymmetry. This has the striking effect that at the non-perturbative level the type II theory should exhibit enhanced gauge symmetries equivalent to the $N = 4$ heterotic string*.

N	G	repr.
3	$SU(3, n-3)$	(n_c)
4	$SU(1,1) \times SO(6, n-6)$	$(2, n)$
5	$SU(5,1)$	(20)
6	$SO^*(12)$	(32)
8	$E_{7(7)}$	(56)

Table: Representations of G for $\left(F^{-\Lambda}, G^-_\Lambda\right)_{\Lambda=1,\dots,n}$ in extended supergravities

13.5 5 On monodromies in string effective field theories

13.5.1 5.1 Classical and quantum monodromies

We have just seen that the tree-level values of the symplectic sections $\left(X^\Lambda(z), F_\Lambda(z)\right)$ are given by

$$X^\Lambda \equiv X^\Lambda_{\text{tree}} \quad , F_\Lambda = S\eta_{\Lambda\Sigma} X^\Sigma_{\text{tree}} \tag{5.1}$$

The target space duality group $O(2, r;\mathbb{Z})$ acts non-trivially on them

$$\Gamma_{\text{cl}} : \begin{pmatrix} X^\Lambda \\ F_\Lambda \end{pmatrix}_{\text{tree}} \to \begin{pmatrix} A & 0 \\ 0 & \eta A \eta \end{pmatrix} \begin{pmatrix} X^\Lambda \\ F_\Lambda \end{pmatrix}_{\text{tree}} \tag{5.2}$$

generalizing the action of the Weyl group of the rigid case [2].

At the one loop level, one expects that F_Λ^{tree} is changed to 46

$$F_\Lambda^{\text{tree}} \to SX^\Sigma \eta_{\Lambda\Sigma} + f_\Lambda(X) \tag{5.3}$$

where $f_\Lambda(X)$ is a modular covariant structure.

The associated perturbative monodromy can be obtained assuming, according to ref. [1], that the rigid perturbative monodromy does not affect the gravitational sector X^0, X^1, F_0, F_1. Thus the perturbative lower triangular monodromy matrix is $\Gamma_{\text{cl}}T$, where[1] [2]

$$T = \begin{pmatrix} \mathbb{1} & 0 \\ C & \mathbb{1} \end{pmatrix} \tag{5.4}$$

and C is an $(r+2)\times(r+2)$ symmetric matrix with non-vanishing entries on the $r \times r$ block

$$C = \begin{pmatrix} 0 \; 0 \; \dots \; 0 \\ 0 \; 0 \; \dots \; 0 \\ 0 \; 0 & \\ \vdots \; \vdots \; C_{ij} & \\ 0 \; 0 & \end{pmatrix} \qquad i,j = 1,\dots,r \tag{5.5}$$

Indeed, we may think of decomposing $Sp(4+2r)$ into $Sp(4)\times Sp(2r)$ and simply assume that the rigid monodromy $\Gamma_r \in Sp(2r)$ commute with the gravitational $Sp(4)$ sector. This argument should at least apply when the vectors of the Cartan subalgebra of the enhanced gauge symmetry belong to the compact $O(r)$ in $O(2,r)$.

In string theory, the classical stringy moduli space corresponds to the broken phase $U(1)^r$ of several gauge groups with the same rank. For instance, for $r = 2$, $O(2,2;\mathbb{Z})$ interpolates between $SU(2)\times U(1), SU(2)\times SU(2)$ and $SU(3)$51. In the $N = 4$ theory the $O(6;22)$ moduli space corresponds to broken phases of several gauge groups of rank 22 such as, $U(1)^6\times E_8\times E_8$ or $SO(32)\times U(1)^6$ or $SO(44)$ which are not subgroups one of the other [39].

It is obvious that generically this means that the one loop β-function term [19] 20] should have non-trivial monodromies at the points where some higher symmetry is restored. For instance, for $r = 2$ we may expect non trivial monodromies around $t = u$ ($SU(2)\times U(1)$ symmetry restored) and $t = u = i, t = u = e^{2i\pi/3}$($SU(2)\times SU(2)$ or $SU(3)$ symmetry restored), t, u being the parameters defined below.

This means that in supergravity theories derived from strings, because of target space T-duality, the enhanced symmetry points are richer than in the rigid case. Since different enhancement points are consequence of $O(2,r;\mathbb{Z})$ duality, we expect that a modular invariant treatment of quantum monodromies will automatically ensure non trivial monodromy at the enhanced symmetry points.

In the sequel we shall discuss in some more detail the classical and perturbative monodromies in the $r = 1$ case ($O(2,1;\mathbb{Z})$) and the classical monodromies for $r = 2$ ($O(2,2;\mathbb{Z})$)

Consider the tree level prepotential F in the so-called cubic form [3] for $\frac{SU(1,1)}{U(1)} \times \frac{O(2,1)}{O(2)}$:

$$F = \frac{1}{2}\left(X^0\right)^2 st^2 \tag{5.6}$$

where $s = \frac{X^1}{X^0}$ is the dilaton coordinate and $t = \frac{X^2}{X^0}$ is the single modulus of the classical target space duality. We parametrize the $O(2,1;\mathbb{Z})$ vector as follows

$$
\begin{aligned}
X^0 &= \frac{1}{2}\left(1-t^2\right) \\
X^1 &= -t \\
X^2 &= -\frac{1}{2}\left(1+t^2\right) \\
&\left(X^0\right)^2+\left(X^1\right)^2-\left(X^2\right)^2=0
\end{aligned}
\tag{5.7}
$$

The symplectic transformation relating $\left(X^\Lambda, F_\Lambda\right), (\Lambda = 0,1,2)$ to the $\left(\widehat{X}^\Lambda, \widehat{F}_\Lambda\right)$ where $O(2,1)$ is linearly realized is easily found to be

$$
\begin{pmatrix} \widehat{X}^\Lambda \\ \widehat{F}_\Lambda \end{pmatrix} = \begin{pmatrix} P & -2R \\ R & P' \end{pmatrix} \begin{pmatrix} X^\Lambda \\ F_\Lambda \end{pmatrix}
\tag{5.8}
$$

where

$$
P = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 0 & -1 \\ -\frac{1}{2} & 0 & 0 \end{pmatrix}; \quad P' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{pmatrix}; \quad R = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix}
\tag{5.9}
$$

Let us now implement the t-modulus $Sl(2,\mathbb{Z})$ transformations $t \to -\frac{1}{t}, t \to t+n$ (note that while $t \to -\frac{1}{t}$ corresponds to the $SU(2)$ Weyl transformation of the rigid theory, $t \to t+n$ has no counterpart in the rigid case, being of stringy nature). Using the parametrization (5.7) we find

$$
\begin{aligned}
t \to -\tfrac{1}{t}: \quad & \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \equiv -\eta \in O(2,1;\mathbb{Z}) \\
t \to t+n: \quad & \begin{pmatrix} 1-\frac{n^2}{2} & n & \frac{n^2}{2} \\ -n & 1 & n \\ -\frac{n^2}{2} & n & 1+\frac{n^2}{2} \end{pmatrix} \equiv V(n) \in O(2,1;\mathbb{Z})
\end{aligned}
\tag{5.10}
$$

Note that (5.10) implies $n \in 2\mathbb{Z}$, i.e. the subgroup $\Gamma_{(0)}(2)$ of $SL(2,\mathbb{Z})$. Actually this gives a projective representation in the subgroup in $O(2,1;\mathbb{Z})$ of the matrices congruent to the identity mod 2.

It follows that Γ_{cl} is generated by (Γ_1, Γ_2) where

$$
\begin{aligned}
\Gamma_1 &= \begin{pmatrix} -\eta & 0 \\ 0 & -\eta \end{pmatrix} \in Sp(6,\mathbb{Z}) \\
\Gamma_2 &= \begin{pmatrix} V(2) & 0 \\ 0 & \eta V(2) \eta \end{pmatrix} \in Sp(6,\mathbb{Z})
\end{aligned}
\tag{5.11}
$$

On the other hand it is possible to go to a stringy basis with a new metric $X_0^2 + X_1^2 - X_2^2 = \widetilde{X}_1^2 + 2XY$ such that $SL(2,\mathbb{Z})$ is integral valued in $O(2,1;\mathbb{Z})$.

The $O(2,1;\mathbb{Z})$ generators corresponding to translation and inversion are respectively given by:

$$
\begin{pmatrix} 1 & -2n & 0 \\ 0 & 1 & 0 \\ n & -n^2 & 1 \end{pmatrix}; \quad \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}
\tag{5.12}
$$

To make contact with the rigid theory it is convenient to define the inversion generator in $O(2,1;\mathbb{Z})$ with the opposite sign with respect to the previous definition.

Let us now examine the perturbative monodromy matrices T. If we assume as before that the $t \to -\frac{1}{t}$ pertaining to the rigid theory does not affect the gravitational sector (X^0, X^1, F_0, F_1), then we have

$$T = \begin{pmatrix} \eta & 0 \\ C & \eta \end{pmatrix} \quad , \quad C = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} \tag{5.13}$$

corresponding to the embedding of the $Sp(2,\mathbb{Z})$ rigid transformations acting on the rigid section (X^2, F_2) in $Sp(6,\mathbb{Z})$. Furthermore, considering the transformation of the $\mathcal{N}_{\Lambda\Sigma}$ matrix and setting $D = A = \eta, B = 0$ we find

$$\widehat{\mathcal{N}}_{22} = -2 + \mathcal{N}_{22} \tag{5.14}$$

for all other entries $\widehat{\mathcal{N}}_{\Lambda\Sigma} = \mathcal{N}^{\Lambda\Sigma}$. This is exactly the rigid result [1]. However conjugating the T matrix with Γ_2 one gets

$$C_{\Lambda\Sigma} = \begin{pmatrix} 8 & -8 & -12 \\ -8 & 8 & 12 \\ -12 & 12 & 18 \end{pmatrix} \tag{5.15}$$

which shows that $O(2,1;\mathbb{Z})$ introduces non-trivial perturbative monodromies for all couplings. The other perturbative lower diagonal monodromy is the dilaton shift (4.14) which commutes with $O(2,1;\mathbb{Z})$.

Analogous considerations hold for $O(2,n;\mathbb{Z}), n > 1$. We limit ourselves to write down the generators of Γ_{cl} for the $O(2,2;\mathbb{Z})$ case. We use the parametrization of $O(2,2)/O(2) \times O(2)$ given by

$$\begin{aligned} X^0 &= \frac{1}{2}(1 - tu) \\ X^1 &= -\frac{1}{2}(t + u) \\ X^2 &= -\frac{1}{2}(1 + tu) \\ X^3 &= \frac{1}{2}(t - u) \quad \left(X^0\right)^2 + \left(X^1\right)^2 - \left(X^2\right)^2 - \left(X^3\right)^2 = 0 \end{aligned} \tag{5.16}$$

where t, u are the moduli appearing in the F function $F = \left(X^0\right)^2$ stu. In the same way as for the $r = 1$ case it is easy to find the symplectic transformations relating the sections of the cubic parametrization to the X^Λ defined in (5.16). They are given by

$$\begin{pmatrix} X \\ F \end{pmatrix} \to \begin{pmatrix} A & B \\ -B & A \end{pmatrix} \begin{pmatrix} X \\ F \end{pmatrix} \tag{5.17}$$

with

$$\begin{aligned} &X = \left(X^0, X^1, X^2, X^3\right)^T \quad , \quad F = (F_0, F_1, F_2, F_3)^T \\ &A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \quad , \quad B = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} . \end{aligned} \tag{5.18}$$

It is convenient to use the string basis where the metric η takes the form 12]

$$\eta = \begin{pmatrix} 0 & \mathbb{1}_{2\times 2} \\ \mathbb{1}_{2\times 2} & 0 \end{pmatrix} \tag{5.19}$$

corresponding to the basis $\frac{1}{\sqrt{2}}\left(X^0 \mp X^2\right), \frac{1}{\sqrt{2}}\left(X^1 \mp X^3\right)$. Then one finds the following $O(2,2;\mathbb{Z})$ representation

$$\begin{aligned}
ut &\to +\frac{1}{ut} : \begin{pmatrix} 0 & -\mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix} = \gamma_{ut} \\
t &\to -\frac{1}{t} : \begin{pmatrix} \epsilon & 0 \\ 0 & \epsilon \end{pmatrix} = \gamma_t \\
u &\to -\frac{1}{u} : \begin{pmatrix} 0 & \epsilon \\ \epsilon & 0 \end{pmatrix} = \gamma_u \\
t &\to t+n : \begin{pmatrix} \mathbf{N}^t(-n) & 0 \\ 0 & \mathbf{N}(n) \end{pmatrix} = \gamma_n \\
t &\to u : \begin{pmatrix} a & b \\ b & a \end{pmatrix} = \gamma; \quad a = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}; \quad b = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}
\end{aligned} \tag{5.20}$$

where $\epsilon = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $\mathbf{N}(n) = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$.

Γ_{cl} is then generated by the matrices:

$$\Gamma_{ut} = \begin{pmatrix} \gamma_{ut} & 0 \\ 0 & \gamma_{ut} \end{pmatrix}; \Gamma_t = \begin{pmatrix} \gamma_t & 0 \\ 0 & \gamma_t \end{pmatrix}; \Gamma_u = \begin{pmatrix} \gamma_u & 0 \\ 0 & \gamma_u \end{pmatrix}; \Gamma_n = \begin{pmatrix} \gamma_n & 0 \\ 0 & \gamma_{-n}^T \end{pmatrix} \tag{5.21}$$

We note that the points $t = u; t = u = i; t = u = e^{\frac{2\pi i}{3}}$ are enhanced symmetry points corresponding to $SU(2) \times U(1), SU(2) \times SU(2)$, and $SU(3)$ respectively [51]. Therefore we expect non-trivial quantum monodromies at these points according to the previous discussion.

13.5.2 5.2 The BPS mass formula

The classical and one loop monodromies are of course reflected in symmetries of the electrically charged massive states belonging to $O(2,n;\mathbb{Z})$ lorentzian lattice [39]. The BPS mass formula [55] in the gravitational case is

$$M = |Z| = \left| n_\Lambda^{(e)} L^\Lambda - n_{(m)}^\Lambda M_\Lambda \right| = e^{K/2} \left| n_\Lambda^{(e)} X^\Lambda - n_{(m)}^\Lambda F_\Lambda \right| \tag{5.22}$$

Note that the central charge Z has definite $U(1)$ weight

$$Z \to e^{(\bar{f}-f)/2} Z \tag{5.23}$$

while the mass M is Kähler invariant. The symplectic invariance of M also implies that $\left(n_{(m)}^\Lambda, n_\Lambda^{(e)}\right)$ transforms as $\left(X^\Lambda, F_\Lambda\right)$

$$\begin{pmatrix} n_{(m)}^\Lambda \\ n_\Lambda^{(e)} \end{pmatrix} \to \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} n_{(m)}^\Lambda \\ n_\Lambda^{(e)} \end{pmatrix} \tag{5.24}$$

where according to our previous discussion the perturbative symmetries have $B = 0$. Note that $n_{(m)}^\Lambda, n_\Lambda^{(e)}$ must satisfy a lattice condition. In the tree level approximation we may write

$$M = \left| \left(n_\Lambda^{(e)} - n_{(m)}^\Sigma \eta_{\Lambda\Sigma} S \right) X^\Lambda \right| e^{K/2} \tag{5.25}$$

which is invariant under the tree level symmetry $S \to S+1$, but also under the non-perturbative inversion $S \to -\frac{1}{S}$ [34] [13] [12] [14] [15] taking into account that

$$K = -\log i(\bar{S}-S) - \log \frac{X^\Lambda \bar{X}^\Sigma}{M_{Pl}^2}\eta_{\Lambda\Sigma} \tag{5.26}$$

Formula (5.25) is therefore invariant under the $S-T$ duality symmetry $Sl(2;\mathbb{Z})\times O(2,r;\mathbb{Z}) \subset Sp(2r+4;\mathbb{Z})$

The electric mass spectrum can be written as

$$M_{(e)}^2 = |Z|^2 = \frac{M_{Pl}^2}{2i(\bar{S}-S)}\boldsymbol{Q}^{\Lambda\Sigma} n_\Lambda^{(e)} n_\Sigma^{(e)} \tag{5.27}$$

where $i(\bar{S}-S) = \frac{8\pi}{g^2} > 0$ and $\boldsymbol{Q}^{\Lambda\Sigma} = \Phi^\Lambda\bar{\Phi}^\Sigma + \bar{\Phi}^\Lambda\Phi^\Sigma$. Formula (5.27) has exactly the same form as the analogous one obtained in $N=4$ (see ref 12]). When also magnetic charges are present, then

$$\begin{aligned} M^2 &= \frac{M_{Pl}^2}{i(\bar{S}-S)}\left(n_\Lambda^e - S n_\Lambda^m\right)\left(\frac{1}{2}\boldsymbol{Q}_{\Lambda\Sigma} - \frac{i}{2}\widehat{\boldsymbol{Q}}_{\Lambda\Sigma}\right)\left(n_\Sigma^e - \bar{S} n_\Sigma^m\right) \\ &= \frac{M_{Pl}^2}{4}\left(n_m, n_e\right)\left(\boldsymbol{\mathcal{M}}\boldsymbol{Q} + \boldsymbol{\mathcal{L}}\widehat{\boldsymbol{Q}}\right)\begin{pmatrix} n_m \\ n_e \end{pmatrix} \end{aligned} \tag{5.28}$$

where $\boldsymbol{\mathcal{M}} = \frac{1}{\operatorname{Im} S}\begin{pmatrix} S\bar{S} & -\operatorname{Re} S \\ -\operatorname{Re} S & 1 \end{pmatrix}$, $\boldsymbol{\mathcal{L}} = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}$ and $\widehat{\boldsymbol{Q}} = i\left(\Phi^\Lambda\bar{\Phi}^\Sigma - \bar{\Phi}^\Lambda\Phi^\Sigma\right)$.

Recalling that $\boldsymbol{Q}^{\Lambda\Sigma} = \frac{1}{2}\left(\eta^{\Lambda\Sigma} + \frac{\operatorname{Im}\boldsymbol{\mathcal{N}}^{\Lambda\Sigma}}{\operatorname{Im} S}\right)$, this becomes

$$\begin{aligned} M^2 &= \frac{M_{Pl}^2}{i(\bar{S}-S)}\left(n_\Lambda^e - S n_\Lambda^m\right)\left[\frac{1}{4}\left(\frac{\operatorname{Im}\boldsymbol{\mathcal{N}}_{\Lambda\Sigma}}{\operatorname{Im} S} + \eta_{\Lambda\Sigma}\right) - \frac{i}{2}\widehat{\boldsymbol{Q}}_{\Lambda\Sigma}\right]\left(n_\Sigma^e - \bar{S} n_\Sigma^m\right) \\ &= \frac{1}{4}M_{Pl}^2\left(n_m, n_e\right)\left[\frac{1}{2}\boldsymbol{\mathcal{M}}\left(\frac{\operatorname{Im}\boldsymbol{\mathcal{N}}}{\operatorname{Im} S} + \eta\right) + \boldsymbol{\mathcal{L}}\widehat{\boldsymbol{Q}}\right]\begin{pmatrix} n_m \\ n_e \end{pmatrix} \end{aligned} \tag{5.29}$$

From this expression one can see that the antisymmetric term $\widehat{\boldsymbol{Q}}$ vanishes if

$$n_\Lambda^{(e)} = m_1 n_\Lambda \quad , \quad n_{(m)}^\Lambda = m_2 n_\Sigma \eta^{\Lambda\Sigma} \tag{5.30}$$

or, as it happens for the perturbative string, if no magnetic states are present ($n_\Lambda^m = 0, n_\Lambda^e \equiv n_\Lambda$). In such case eq. (5.28) becomes

$$M^2 = \frac{M_{Pl}^2}{8\operatorname{Im} S}\left|m_1 - S m_2\right|^2\left[n_\Lambda n_\Sigma\left(2\boldsymbol{Q}^{\Lambda\Sigma} - \eta^{\Lambda\Sigma}\right) + n_\Lambda n_\Sigma \eta^{\Lambda\Sigma}\right] \tag{5.31}$$

and since $\operatorname{Im}\boldsymbol{\mathcal{N}}^{\Lambda\Sigma}$, being the vector kinetic matrix, is always positive definite,

$$M^2 = 0 \Longleftrightarrow n^\Lambda n^\Sigma \eta_{\Lambda\Sigma} < 0 \quad \left(n^\Lambda \neq 0\right) \tag{5.32}$$

As an example, take $O(2,2;\mathbb{Z})$ and look for solutions of (5.32) corresponding to the string condition $n^\Lambda n_\Lambda = -2$. Using the parametrization (5.16) we have

$$\begin{aligned} n_\Lambda X^\Lambda &= n_0 X^0 + n_1 X^1 - n_2 X^2 - n_3 X^3 \\ &= \frac{1}{2}\left[\left(n_0 + n_2\right) - \left(n_1 + n_3\right)t - \left(n_1 - n_3\right)u - \left(n_0 - n_2\right)tu\right] \end{aligned} \tag{5.33}$$

Setting

$$\begin{aligned} n^0 + n^2 &= -p_2\sqrt{2} \\ n^1 + n^3 &= q_1\sqrt{2} \\ n^1 - n^3 &= -p_1\sqrt{2} \\ n^0 - n^2 &= q_2\sqrt{2} \\ n^\Lambda n_\Lambda &= \left(n^0 + n^2\right)\left(n^0 - n^2\right) + \left(n^1 + n^3\right)\left(n^1 - n^3\right) = -2\left(p_2 q_2 + p_1 q_1\right) = -2 \\ &\rightarrow p_2 q_2 + p_1 q_1 = 1 \end{aligned} \tag{5.34}$$

we have

$$n_\Lambda X^\Lambda = \frac{1}{\sqrt{2}}\left(-p_2 - q_1 t + p_1 u - q_2 t u\right) \tag{5.35}$$

Let us verify that at the three enhancement points we get the correct number of massless states. If we take $t = u\,(X^2 = 0)$ we find

$$\begin{aligned} n_\Lambda X^\Lambda(t = u) &= \frac{1}{\sqrt{2}}\left[-p_2 - \left(q_1 - p_1\right)t - q_2 t^2\right] \\ &\rightarrow q_2 = p_2 = 0 \quad q_1 = p_1 = \pm 1 \end{aligned}$$

yielding the two massless states $(q_1, q_2) = (\pm 1, 0)$. In particular, for $t = u = i$ we have the solutions

$$\begin{aligned} &n_\Lambda X^\Lambda(t = u = i) = \frac{1}{\sqrt{2}}\left[-p_2 + q_2 - \left(q_1 - p_1\right)i\right] \\ &\rightarrow p_2 = q_2 \quad , \quad q_1 = p_1, \quad q_1^2 + q_2^2 = 1 \end{aligned} \tag{5.36}$$

yielding the four states $(q_1, q_2) = (\pm 1, 0), (0, \pm 1)$. Taking instead $t = u = e^{2\pi i/3}$ (such that $t^2 = \bar{t}$), we get

$$\begin{aligned} &n_\Lambda X^\Lambda\left(t = u = e^{2\pi i/3}\right) = 0 \\ &\rightarrow +\frac{1}{2}\left(q_1 + q_2 - p_1\right) - p_2 = 0, q_1 - q_2 - p_1 = 0 \\ &\rightarrow p_1 = q_1 - q_2, p_2 = q_2 \rightarrow q_1^2 + q_2^2 - q_1 q_2 = 1 \end{aligned} \tag{5.37}$$

yielding the six states $(q_1, q_2) = (\pm 1, \pm 1), (\pm 1, 0), (0, \pm 1)$. As expected, these massless states together with the two original $(0, 0)$ states, fill the adjoint representation of $SU(2) \otimes U(1)(t = u), SU(2) \otimes SU(2)(t = u = i), SU(3)\left(t = u = e^{2\pi i/3}\right)$.

Unlike in $N = 4$ theories, in $N = 2$ theories the quantum spectrum will not coincide with the classical spectrum. It will be found by substituting $F_{\Lambda\ \mathrm{tree}} \equiv S\eta_{\Lambda\Sigma}X^\Sigma \rightarrow F_{\Lambda\ \mathrm{tree}} +$ quantum corrections in (5.22).

13.6 6 Conclusions

In this paper we have formulated electromagnetic duality transformations in generic $D = 4, N = 2$ supergravities theories in a form suitable to investigate non-perturbative phaenomena. Our formulation is manifestly duality covariant for the full Lagrangian, including fermionic terms, which unlike the rigid case, cannot be retrieved from the $N = 1$ formulation, nor from the

$N = 2$ tensor calculus approach. Particular attention has been given to classical T-duality symmetries which actually occur in string compactifications and whose linear action on the gauge potential fields do not allow for the existence of a prepotential function F for the $N = 2$ special geometry. As examples we described the "classical" electric and monopole spectrum for T-duality symmetries of the type $O(2, r; \mathbb{Z})$, with particular details for the $r = 1, 2$ cases, by using the $N = 2$ formalism.

For "classical" monodromies this spectrum is of course related to the spectrum of $N = 4$ theories studied by Sen and Schwarz [12]. Possible extensions of duality symmetries to type II strings have been conjectured by Hull and Townsend 49] and also discussed in [2]. In the present context of $N = 2$ heterotic strings the corresponding type II theories, having $N = 2$ space-time supersymmetry would correspond to $(2, 2)$ superconformal field theories, i.e. quantum Calabi-Yau manifolds.

Due to the non-compact symmetries the BPS saturated states with nonvanishing central charges have a spectrum quite different from the rigid case. Indeed in rigid theories the "classical" central charge $Z_{(cl)}$ vanishes at the enhanced symmetry points where the original gauge group is restored since there is no dimensional scale other than the Higgs v.e.v.. On the contrary, in the supergravity theory the BPS spectrum at these particular points corresponds in general to electrically and magnetically charged states with Planckian mass (black holes, gravitational monopoles and dyons) 53, 12, 54 57. The only charged states which become massless at the enhanced symmetric point are those with $\eta^{\Lambda\Sigma} n_\Lambda^{(e)} n_\Sigma^{(e)} < 0$.

We also discussed perturbative monodromies and their possible relations with the rigid case. Non perturbative duality symmetries are more difficult to guess, but it is tempting to conjecture that a quantum monodromy consistent with positivity of the metric and special geometry may be originated by a 3-dimensional Calabi-Yau manifold or its mirror image. If this is the case this manifold should embed in some sense the class of Riemann surfaces studied [1] [2] in connection with the moduli space of $N = 2$ rigid supersymmetric Yang-Mills theories.

14 Other methods and theories with duality

15 5d/4d U-dualities and $\mathcal{N}=8$ black holes Anna Ceresole, Ferrara, Gnecchi

We use the connection between the U-duality groups in $d = 5$ and $d = 4$ to derive properties of the $\mathcal{N} = 8$ black hole potential and its critical points (attractors). This approach allows to study and compare the supersymmetry features of different solutions.

15.0.1 Introduction

The $\mathcal{N} = 8$ supergravity theory in $d = 4$ [1] and $d = 5$ [2] dimensions is a remarkable theory which unifies the gravitational fields with other lower spin particles in a rather unique way, due to the high constraints of local $\mathcal{N} = 8$ supersymmetry, the maximal one realized in a 4d Lagrangian field theory. These theories, particularly in four dimensions, are supposed to enjoy exceptional ultraviolet properties. For this reason, 4d supergravity has been advocated not only as the simplest quantum field theory [3] but also as a potential candidate for a finite theory of quantum gravity, even without its completion into a larger theory [4]. Maximal supergravity in highest dimensions has a large number of classical solutions [5] which may survive at the quantum level. Among them, there are black p-branes of several types[52] and interestingly, 4d black holes of different nature.

On the other hand, theories with lower supersymmetries (such as $\mathcal{N} = 2$) emerging from Calabi-Yau compactifications of M-theory or superstring theory, admit extremal black hole solutions that have been the subject of intense study, because of their wide range of classical and quantum aspects. For asymptotically flat, stationary and spherically symmetric extremal black holes, the attractor behaviour [26, 8] has played an important role not only in determining universal features of fields flows toward the horizon, but also to explore dynamical properties such as wall crossing[9] and split attractor flows[10], the connections with string topological partition functions[11] and relations with microstates counting[12]. Therefore, it has become natural to study the properties of extremal black holes not only in the context of $\mathcal{N} = 2$, but also in theories with higher supersymmetries, up to $\mathcal{N} = 8$[13]-[22].

In $\mathcal{N} = 8$ supergravity, in the Einsteinian approximation, there is a nice relation between the classification of large black holes which undergo the attractor flow and charge orbits which classify, in a duality invariant manner, the properties of the dyonic vector of electric and magnetic charges $Q = (p^\Lambda, q_\Lambda)$ ($\Lambda = 0, ..., 27$ in $d = 4$) [23, 24]. The attractor points are given by extrema of the $4d$ black hole potential, which is given by [16, 17]

$$V_{BH} = \frac{1}{2} Z_{AB} Z^{*\,AB} = \langle Q, V_{AB} \rangle \, \langle Q, \overline{V}^{AB} \rangle \ , \tag{15.1}$$

where the central charge is the antisymmetric matrix ($A, B = 1, ..., 8$)

$$Z_{AB} = \langle Q, V_{AB} \rangle = Q^T \, \Omega \, V_{AB} = f^\Lambda{}_{AB} \, q_\Lambda - h_{\Lambda\, AB} \, p^\Lambda \ , \tag{15.2}$$

the symplectic sections are

$$V_{AB} = (f^\Lambda{}_{AB}, h_{\Lambda\, AB}) \ , \tag{15.3}$$

and Ω is the symplectic invariant metric.

An important role is played by the Cartan quartic invariant I_4[25, 1] in that it only depends on Q and not on the asymptotic values of the 70 scalar fields φ. This means that if we construct I_4 as a combination of quartic powers of the central charge matrix $Z_{AB}(q, p, \varphi)$ [26], the φ dependence drops out from the final expression

$$\frac{\partial}{\partial \varphi} I_4(Z_{AB}) = 0. \tag{15.4}$$

Analogue (cubic) invariants I_3 exist for black holes and/or (black) strings in $d = 5$[8, 23]. These are given by

$$I_3(p^I) = \frac{1}{3!} d_{IJK} p^I p^J p^K \ , \tag{15.5}$$

$$I_3(q_I) = \frac{1}{3!} d^{IJK} q_I q_J q_K \ , \tag{15.6}$$

where d_{IJK}, d^{IJK} are the $(27)^3$ $E_{6(6)}$ invariants. Consequently, the $d = 4$ $E_{7(7)}$ quartic invariant takes the form

$$I_4(Q) = -(p^0 q_0 + p^I q_I)^2 + 4 \left[-p^0 I_3(q) + q_0 I_3(p) + \frac{\partial I_3(q)}{\partial q_I} \frac{\partial I_3(p)}{\partial p^I} \right] . \tag{15.7}$$

On the other hand, in terms of the central charge matrices $Z_{ab}(\phi, q)$ (in $d = 5$ this is the **27** representation of $USp(8)$) and $Z_{AB}(\phi, p, q)$ (in $d = 4$ this is the **28** of $SU(8)$), their expression is

$$I_3(q) = Z_{ab} \Omega^{bc} Z_{cd} \Omega^{dq} Z_{qp} \Omega^{pa} \ , \qquad Z_{ab} \Omega^{ab} = 0 \ , \tag{15.8}$$

$$I_4(p, q) = \frac{1}{4} \left[4\, Tr(Z Z^\dagger Z Z^\dagger) - (Tr\, Z Z^\dagger)^2 + 32 \, (Pf \, Z_{AB}) \right] \ , \tag{15.9}$$

where $ZZ^\dagger = Z_{AB}\bar{Z}^{CB}$, Ω^{ab} is the $5d$ symplectic invariant metric, and the Pfaffian of the central charge is [1]

$$Pf\,(Z_{AB}) = \frac{1}{2^4 4!}\epsilon^{ABCDEFGH} Z_{AB}Z_{CD}Z_{EF}Z_{GH}. \tag{15.10}$$

In fact, these are simply the (totally symmetric) invariants which characterize the 27 dimensional representation of $E_{6(6)}$ and the 56 dimensional representation of $E_{7(7)}$, which are the U-duality [27] symmetries of $\mathcal{N} = 8$ supergravity in $d = 5$ and $d = 4$, respectively.

When charges are chosen such that I_4 and I_3 are not vanishing, one has large black holes and in the extremal case the attractor behaviour may occur. However, while at $d = 5$ there is a unique ($\frac{1}{8}$-BPS) attractor orbit with $I_3 \neq 0$, associated to the space[24, 28]

$$\mathcal{O}_{d=5} = \frac{E_{6(6)}}{F_{4(4)}}\ , \tag{15.11}$$

at $d = 4$ two orbits emerge, the BPS one

$$\mathcal{O}_{d=4,\,BPS} = \frac{E_{7(7)}}{E_{6(2)}}\ , \tag{15.12}$$

and the non BPS one with different stabilizer

$$\mathcal{O}_{d=4,\,non-BPS} = \frac{E_{7(7)}}{E_{6(6)}}. \tag{15.13}$$

Such orbits have further ramifications in theories with lower supersymmetry , but it is the aim of this paper to confine our attention to the $\mathcal{N} = 8$ theory.

In this paper, extending a previous result for $\mathcal{N} = 2$ theories [29], we elucidate the connection between these configurations and we relate the critical points of the $\mathcal{N} = 8$ black hole potential of the $5d$ and $4d$ theories. To achieve this goal we use a formulation of $4d$ supergravity in a $E_{6(6)}$ duality covariant basis [30], which is appropriate to discuss a $4d/5d$ correspondence. This is not the same as the Cremmer-Julia[1] or de Wit-Nicolai[31] manifest $SO(8)$ (and $SL(8,\bar{b}R)$) covariant formulation, but it is rather related to the Sezgin-Van Nieuwenhuizen $5d/4d$ dimensional reduction[32]. These two formulations are related to one another by dualizing several of the vector fields and therefore they interchange electric and magnetic charges of some of the 28 vector fields of the final theory. The precise relation between these theories was recently discussed in [33].

The paper is organized as follows. In sec. 15.0.2 we rewrite the $4d$ black hole potential in terms of central charges. This is essential in order to discuss the supersymmetry properties of the solutions. In fact, in the specific solutions we consider in sec. 15.0.3 and 15.0.4, BPS and non-BPS critical points are simply obtained by some charges sign flip. This will manifest in completely different symmetry properties of the central charge matrix, in the normal frame, at the fixed point. These properties reflect the different character of the BPS and non BPS charge orbits.

The solutions of the critical point equations are particularly simple in the “axion free” case, discussed in sec. 15.0.3 and 15.0.4, which only occur for some chosen charge configurations. In sec. 15.0.3 we derive critical point equations that are completely general and that may be used to study any solution.

The formula for the $\mathcal{N} = 8$ potential given in sec. 15.0.2 was obtained in an earlier work [33], and it is identical to the $\mathcal{N} = 2$ case [29]. The only difference relies in the kinetic matrix a_{IJ} which, in $\mathcal{N} = 2$ is given by real special geometry while in $\mathcal{N} = 8$ is given in terms of the $E_{6(6)}$ coset representatives [32, 16]. However, in the normal frame, when we suitably restrict to two moduli, this matrix does indeed become an $\mathcal{N}=2$ matrix, although the interpretation in terms of central charges is completely different.

The supersymmetry properties of the solutions in the $\mathcal{N}=8$ and $\mathcal{N}=2$ theories are compared in subsection 15.0.4. We will see that in the $\mathcal{N}=2$ interpretation, depending on the sign of the charges, both a BPS and a non-BPS branch exist in $d=5$ while two non BPS branches exist in the $d=4$ theory. In $\mathcal{N}=8$, the occurrence of one less branch in both dimensions is due to the fact that the central and matter charges of the $\mathcal{N}=2$ theory are all embedded in the central charge matrix of the $\mathcal{N}=8$ theory. The higher number of attractive orbits can also be explained by the different form of the relevant non compact groups and their stabilizers for the moduli space of solutions.

15.0.2 4d/5d relations for the $\mathcal{N}=8$ extremal black hole potential

In this section we remind the reader how the $\mathcal{N}=8$ potential was derived in a basis that illustrates the relation between 4 and 5 dimensions [33].

Using known identities [17, 34], the black hole potential can be written as a quadratic form in terms of the charge vector Q and the symplectic 56×56 matrix $\mathcal{M}(\mathcal{N})$, related to the $4d$ vector kinetic matrix $\mathcal{N}_{\Lambda\Sigma}$

$$V_{BH} = -\frac{1}{2} Q^T \mathcal{M}(\mathcal{N}) Q\,, \tag{15.14}$$

where $\mathcal{M}$ is

$$\mathcal{M}(\mathcal{N}) = \begin{pmatrix} \mathrm{Im}\,\mathcal{N} + \mathrm{Re}\mathcal{N}(\mathrm{Im}\,N)^{-1}\mathrm{Re}\mathcal{N} & -\mathrm{Re}\mathcal{N}(\mathrm{Im}\,\mathcal{N})^{-1} \\ -(\mathrm{Im}\,\mathcal{N})^{-1}\mathrm{Re}\mathcal{N} & (\mathrm{Im}\,\mathcal{N})^{-1} \end{pmatrix}. \tag{15.15}$$

The indices $\Lambda\,,\Sigma$ of $\mathcal{N}_{\Lambda\Sigma}$ are now split as $(0,I)$, according to the decomposition of $4d$ charges with respect to $5d$ ones, thus $\mathcal{N}_{\Lambda\Sigma}$ assumes the block form

$$\mathcal{N}_{\Lambda\,\Sigma} = \left(\begin{array}{c|c} \mathcal{N}_{00} & \mathcal{N}_{0\,J} \\ \hline \mathcal{N}_{I\,0} & \mathcal{N}_{I\,J} \end{array}\right)\,, \tag{15.16}$$

The kinetic matrix depends on the 70 scalars of the $\mathcal{N}=8$ theory, which are given, in the $5d/4d$ KK reduction, by the 42 scalars of the $5d$ theory (encoded in the $5d$ vector kinetic matrix $a_{IJ}=a_{JI}$), by the 27 axions a^I and the dilaton field e^{ϕ}. In a normalization that is suitable for comparison to $\mathcal{N}=2$, it has the form

$$\mathcal{N}_{\Lambda\Sigma} = \left(\begin{array}{c|c} \frac{1}{3}d - i\left(e^{2\phi}a_{IJ}a^I a^J + e^{6\phi}\right) & -\frac{1}{2}d_J + ie^{2\phi}a_{KJ}a^K \\ \hline -\frac{1}{2}d_I + ie^{2\phi}a_{IK}a^K & d_{IJ} - ie^{2\phi}a_{IJ} \end{array}\right)\,, \tag{15.17}$$

where

$$d \equiv d_{IJK}a^I a^J a^K \quad , \quad d_I \equiv d_{IJK}a^J a^K \quad , \quad d_{IJ} \equiv d_{IJK}a^K\,. \tag{15.18}$$

The black hole potential of [33], computed from (15.14) using the above formulas, can be rearranged as

$$\begin{aligned}
V_{BH} = & \frac{1}{2}\left(p^0 e^{\phi} a^I\right) a_{IJ}\left(p^0 e^{\phi} a^J\right) + \frac{1}{2}\left(p^0 e^{3\phi}\right)^2 + \frac{1}{2}\left(\frac{d}{6} p^0 e^{-3\phi}\right)^2 + \\
& + \frac{1}{2}\left(\frac{1}{2} e^{-\phi} p^0 d_I\right) a^{IJ}\left(\frac{1}{2} e^{-\phi} p^0 d_J\right) + \frac{1}{2} \times 2\left(-p^0 e^{\phi} a_I\right) a_{IJ}\left(p^J e^{\phi}\right) + \\
& + \frac{1}{2} \times 2\left(\frac{d}{6} p^0 e^{-3\phi}\right)\left(-\frac{1}{2} p^I d_I e^{-3\phi}\right) - \frac{1}{2} \times 2\left(\frac{1}{2} p^0 e^{-\phi} d_I\right) a^{IJ}\left(p^K d_{KJ} e^{-\phi}\right) + \\
& + \frac{1}{2}\left(e^{\phi} p^I\right) a_{IJ}\left(e^{\phi} p^J\right) + \frac{1}{2}\left(\frac{1}{2} e^{-3\phi} p^K d_K\right)^2 + \\
& + \frac{1}{2}\left(e^{-\phi} p^K d_{KI}\right) a^{IJ}\left(e^{-\phi} p^L d_{JL}\right) + \frac{1}{2} \times 2\left(q_0 e^{-3\phi}\right)\left(\frac{d}{6} p^0 e^{-3\phi}\right) + \\
& + \frac{1}{2} \times 2\left(q_I a^I e^{-3\phi}\right)\left(\frac{d}{6} p^0 e^{-3\phi}\right) + \frac{1}{2} \times 2\left(q_I e^{-\phi}\right) a^{IJ}\left(\frac{1}{2} p^0 d_J e^{-\phi}\right) + \\
& - \frac{1}{2} \times 2\left(q_0 e^{-3\phi}\right)\left(\frac{1}{2} p^I d_I e^{-3\phi}\right) - \frac{1}{2} \times 2\left(q_I a^I e^{-3\phi}\right)\left(\frac{1}{2} p^J d_J e^{-3\phi}\right) + \\
& - \frac{1}{2} \times 2\left(q_I e^{-\phi}\right) a^{IJ}\left(p^K d_{KJ} e^{-\phi}\right) + \frac{1}{2}\left(q_0 e^{-3\phi}\right)^2 + \frac{1}{2} \times 2\left(q_0 e^{-3\phi}\right)\left(q_I a^I e^{-3\phi}\right) + \\
& + \frac{1}{2}\left(q_I a^I e^{-3\phi}\right)^2 + \frac{1}{2}\left(q_I e^{-\phi}\right) a^{IJ}\left(q_J e^{-\phi}\right) \ ,
\end{aligned} \tag{15.19}$$

with $a^{IJ} = a^{-1}_{IJ}$. This form shows that it can be written in terms of squares of electric and magnetic components as

$$V_{BH} = \frac{1}{2}(Z_0^e)^2 + \frac{1}{2}\left(Z_m^0\right)^2 + \frac{1}{2} Z_I^e a^{IJ} Z_J^e + \frac{1}{2} Z_m^I a_{IJ} Z_m^J \ , \tag{15.20}$$

provided one defines,

$$\begin{aligned}
Z_0^e &= e^{-3\phi} q_0 + e^{-3\phi} q_I a^I + e^{-3\phi} \frac{d}{6} p^0 - \frac{1}{2} e^{-3\phi} p^I d_I \ , \\
Z_m^0 &= e^{3\phi} p^0 \ , \\
Z_I^e &= \frac{1}{2} e^{-\phi} p^0 d_I - p^J d_{IJ} e^{-\phi} + q_I e^{-\phi} \ , \\
Z_m^I &= e^{\phi} p^I - e^{\phi} p^0 a^I .
\end{aligned} \tag{15.21}$$

In order to get the symplectic embedding of the four dimensional theory, we still need to complexify the central charges. To this end, we define the two complex vectors

$$\begin{aligned}
Z_0 &\equiv \frac{1}{\sqrt{2}}(Z_0^e + i Z_m^0) \ , \\
Z_a &\equiv \frac{1}{\sqrt{2}}(Z_a^e + i Z_m^a) \ ,
\end{aligned} \tag{15.22}$$

where

$$Z_a^e = Z_I^e (a^{-1/2})^I_a \ , \qquad Z_m^a = Z_m^I (a^{1/2})^a_I \tag{15.23}$$

such that

$$V_{BH} = |Z_0|^2 + Z_a \bar{Z}_a \ , \tag{15.24}$$

where now $a = 1, ..., 27$ is a flat index, which can be regarded as a $USp(8)$ antisymmetric traceless matrix.

The potential at the critical point gives the black hole entropy corresponding to the given solution, which in $d=4$ reads

$$\frac{S_{BH}}{\pi} = \sqrt{|I_4|} = V_{BH}^{crit.} \;, \tag{15.25}$$

while in $d=5$ it is [38]

$$\frac{S_{BH}}{\pi} = 3^{3/2}|I_3|^{1/2} = \left(3\,V_5^{crit}\right)^{3/4} \;, \tag{15.26}$$

where I_4 and I_3 are the invariants of the $\mathcal{N}=8$ theory in $d=4$ and $d=5$ respectively.

Symplectic sections

In virtue of the previous discussion, we can trade the central charge (15.2)for the 28-component vector

$$Z_A = f^{\Lambda}{}_{A} q_{\Lambda} - h_{\Lambda A} p^{\Lambda} \;, \tag{15.27}$$

where f and h are symplectic sections satisfying the following properties [40, 41]

a) $\mathcal{N}_{\Lambda\Sigma} = h_{\Lambda A}(f^{-1})^{A}{}_{\Sigma}$,

b) $i(\mathbf{f}^{\dagger}\mathbf{h} - \mathbf{h}^{\dagger}\mathbf{f}) = \mathbf{Id}$,

c) $\mathbf{f}^{T}\mathbf{h} - \mathbf{h}^{T}\mathbf{f} = 0.$

Notice that one still has the freedom of a further transformation

$$\begin{aligned} h &\to hM \;, \\ f &\to fM \;, \end{aligned} \tag{15.28}$$

as it leaves invariant the vector kinetic matrix $\mathcal{N}$, as well as relations $a)-c)$, when M is a unitary matrix

$$MM^{\dagger} = 1. \tag{15.29}$$

Indeed, when the central charge transforms as

$$\begin{aligned} Z &\to ZM \;, \\ ZZ^{\dagger} &\to ZMM^{\dagger}Z^{\dagger} = ZZ^{\dagger} \;, \end{aligned} \tag{15.30}$$

the black hole potential

$$V_{BH} \equiv ZZ^{\dagger} \tag{15.31}$$

is left invariant. In our case, we rearrange the 28 indices into a single complex vector index, to be identified, for a suitable choice of M, with the two-fold antisymmetric representation of $SU(8)$, according to the decomposition $\mathbf{28} \to \mathbf{27} + \mathbf{1}$ of $SU(8) \to USp(8)$; we thus have

$$\begin{aligned} Z_0 &= f^{\Lambda}{}_{0} q_{\Lambda} - h_{\Lambda\,0} p^{\Lambda} = \\ &= f^{0}{}_{0} q_0 + f^{J}{}_{0} q_J - h_{0\,0} p^0 - h_{J\,0} p^J \;, \\ Z_a &= f^{\Lambda}{}_{a} q_{\Lambda} - h_{\Lambda\,a} p^{\Lambda} = \\ &= f^{0}{}_{a} q_0 + f^{J}{}_{a} q_J - h_{0\,a} p^0 - h_{J\,a} p^J \;; \end{aligned} \tag{15.32}$$

which, from the definition in (15.22) yields

$$\begin{aligned} Z_0 &= \frac{1}{\sqrt{2}}\left[e^{-3\phi}q_0 + e^{-3\phi}a^I q_I + \left(e^{-3\phi}\frac{d}{6} + ie^{3\phi}\right)p^0 - \frac{1}{2}\left(e^{-3\phi}d_I\right)p^I\right] , \\ Z_a &= \frac{1}{\sqrt{2}}\left[e^{-\phi}q_I(a^{-1/2})^I{}_a + \left(\frac{1}{2}e^{-\phi}d_I(a^{-1/2})^I{}_a - ie^{\phi}a^J(a^{1/2})_J{}^a\right)p^0 + \right. \\ &\quad \left. - \left(e^{-\phi}d_{IJ}(a^{-1/2})^I{}_a - ie^{\phi}(a^{1/2})_J{}^a\right)p^J\right] . \end{aligned} \tag{15.33}$$

Thus we consider

$$f^\Lambda{}_A = \frac{1}{\sqrt{2}}\left(\begin{array}{c|c} e^{-3\phi} & 0 \\ \hline e^{-3\phi}a^I & e^{-\phi}(a^{-1/2})^I{}_a \end{array}\right) , \tag{15.34}$$

$$h_{\Lambda A} = \frac{1}{\sqrt{2}}\left(\begin{array}{c|c} -e^{-3\phi}\frac{d}{6} - ie^{3\phi} & -\frac{1}{2}e^{-\phi}d_K(a^{-1/2})^K{}_a + ie^{\phi}a^K(a^{1/2})_K{}^a \\ \hline \frac{1}{2}e^{-3\phi}d_I & e^{-\phi}d_{IJ}(a^{-1/2})^J{}_a - ie^{\phi}(a^{1/2})_I{}^a \end{array}\right) . \tag{15.35}$$

From $\mathbf{f}^{-1}$

$$(f^{-1})_\Lambda{}^A = \sqrt{2}\left(\begin{array}{c|c} e^{3\phi} & 0 \\ \hline -e^{\phi}a^I(a^{1/2})_I{}^a & e^{\phi}(a^{1/2})_I{}^a \end{array}\right) , \tag{15.36}$$

by matrix multiplication, we find that relations $a)$ $b)$ and $c)$ are fulfilled by $\mathbf{f}$ and $\mathbf{h}$, that we now recognize to be the symplectic sections.

We finally perform the transformation $f' = fM$ (where $M = f^{-1}f' = h^{-1}h'$), with M unitary matrix, in virtue of identities $a)$, $b)$ and $c)$, valid for both (f, h) and (f', h'). A model independent formula for M valid for any $\mathcal{N}=2$ d-geometry (in particular, for any truncation of $\mathcal{N}=8$ to an $\mathcal{N}=2$ geometry, such as the models treated in this paper) is given by the matrix [42]

$$M = A^{1/2}\hat{M}G^{-1/2} , \tag{15.37}$$

with

$$A = \left(\begin{array}{c|c} 1 & 0...0 \\ \hline \begin{matrix}0\\ \vdots \\ 0\end{matrix} & a_{IJ} \end{array}\right) , \qquad G = \left(\begin{array}{c|c} 1 & 0...0 \\ \hline \begin{matrix}0\\ \vdots \\ 0\end{matrix} & g_{IJ} \end{array}\right) , \qquad g_{IJ} = \frac{1}{4}e^{-4\phi}a_{IJ} , \tag{15.38}$$

where $\hat{M}$ is given by

$$\hat{M} = \frac{1}{2}\begin{pmatrix} 1 & \partial_{\bar{J}}K \\ -i\lambda^I e^{-2\phi} & e^{-2\phi}\delta^I_{\bar{J}} + ie^{-2\phi}\lambda^I\partial_{\bar{J}}K \end{pmatrix} , \tag{15.39}$$

where "$-\lambda^I$" are the imaginary parts of the complex moduli $z^I = a^I - i\lambda^I$, and K is the Kähler potential $K = -\ln(8\mathcal{V})$, with $\mathcal{V} = \frac{1}{3!}d_{IJK}\lambda^I\lambda^J\lambda^K$; the matrix $\hat{M}$ satisfies the properties

$$\begin{aligned} A\hat{M}G^{-1}\hat{M}^\dagger &= Id \ , \\ G^{-1}\hat{M}^\dagger A\hat{M} &= Id. \end{aligned} \tag{15.40}$$

For the models considered below, this matrix M does indeed reproduce, for the given special configurations, the formula in eq. (15.59).

Note that $\hat{M}$ performs the change of basis between the central charges defined as

$$\begin{aligned} Z_0 &= \frac{1}{\sqrt{2}}(Z_0^e + iZ_m^0) \ , \\ Z_I &= \frac{1}{\sqrt{2}}(Z_I^e + ia_{IJ}Z_m^J) \ , \end{aligned} \tag{15.41}$$

and the special geometry charges $(Z, \mathcal{D}_{\bar{I}}\overline{Z})$, that is the charges in "curved" rather than the "flat" indices.

15.0.3 Attractors in the 5 dimensional theory

It was shown in [23] that the cubic invariant of the five dimensions can be written as

$$I_3 = Z_1^5 Z_2^5 Z_3^5 \ , \tag{15.42}$$

where Z_a^5's are related to the skew eigenvalues of the $USp(8)$ central charge matrix in the normal frame

$$e_{ab} = \begin{pmatrix} Z_1^5 + Z_2^5 - Z_3^5 & 0 & 0 & 0 \\ 0 & Z_1^5 + Z_3^5 - Z_2^5 & 0 & 0 \\ 0 & 0 & Z_2^5 + Z_3^5 - Z_1^5 & 0 \\ 0 & 0 & 0 & -(Z_1^5 + Z_2^5 + Z_3^5) \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} . \tag{15.43}$$

We consider a configuration of only three non-vanishing electric charges (q_1, q_2, q_3), that we can take all non-negative. We further confine to two moduli λ_1, λ_2, describing a geodesic submanifold $SO(1,1)^2 \in E_{6(6)}/USp(8)$ whose special geometry is determined by the constraint

$$\frac{1}{3!}d_{IJK}\hat{\lambda}^I\hat{\lambda}^J\hat{\lambda}^K = \hat{\lambda}^1\hat{\lambda}^2\hat{\lambda}^3 = 1 \ , \tag{15.44}$$

where $\hat{\lambda}^I = \mathcal{V}^{-1/3}\lambda^I$, defining the $stu-$model [29].

The metric a_{IJ}, restricted to this surface, takes the diagonal form

$$a_{IJ} = -\frac{\partial^2}{\partial\hat{\lambda}^I\partial\hat{\lambda}^J}\log\mathcal{V}\big|_{\mathcal{V}=1} = \begin{pmatrix} \frac{1}{\hat{\lambda}_1^2} & 0 & 0 \\ 0 & \frac{1}{\hat{\lambda}_2^2} & 0 \\ 0 & 0 & \frac{1}{\hat{\lambda}_3^2} = \hat{\lambda}_1^2\hat{\lambda}_2^2 \end{pmatrix} , \tag{15.45}$$

and the five dimensional black hole potential for electric charges is*

$$V_5^e = q_I a^{IJ} q_J = \sum_{a=1}^{3} Z_a^5(q) Z_a^5(q) \ , \tag{15.46}$$

*In an analogous way, the black hole potential for magnetic charges, $V_5^m = \sum_{a=1}^3 Z_a^5(p)Z_a^5(p)$, is obtained by replacing $q_I \to p^I$ and $a^{IJ} \to a_{IJ}$ [29, 38], with $Z_a^5(p) = p^I(a^{1/2})_I{}^a$.

with $Z^5_a(q) = (a^{-1/2})^I{}_a\, q_I$; the moduli at the attractor point of the 5-dimensional solution are (see eq. 4.4 and 4.7 of [29])

$$\hat{\lambda}^I_{crit} = \frac{I_3^{1/3}}{q^I} \ , \tag{15.47}$$

and

$$\begin{aligned} V_5^{crit} &= 3|q_1 q_2 q_3|^{2/3} = 3 I_3^{2/3} \ , \\ a^{IJ}_{crit} &= \frac{I_3^{2/3}}{q_I^2} \delta^{IJ} \end{aligned} \tag{15.48}$$

with no sum over repeated indices. We find

$$Z_a^{5,\,crit} = I_3^{1/3} \ , \qquad I_3 = Z_1^5 Z_2^5 Z_3^5 . \tag{15.49}$$

These relations also allow to connect the potential in (15.46)

$$V_5 = (Z_1^5)^2 + (Z_2^5)^2 + (Z_3^5)^2 \ , \tag{15.50}$$

with the form given in terms of the central charges [38], where it is the trace of the square matrix

$$V_5 = \frac{1}{2} Z^5_{ab} Z^{5\,ab} . \tag{15.51}$$

The eigenvalues of Z^5_{ab} are written in (15.43) in terms of Z^5_1, Z^5_2, Z^5_3. The 5d central charge matrix in the normal frame at the attractor point thus becomes

$$e_{ab} = \begin{pmatrix} I_3^{1/3}\epsilon & 0 & 0 & 0 \\ 0 & I_3^{1/3}\epsilon & 0 & 0 \\ 0 & 0 & I_3^{1/3}\epsilon & 0 \\ 0 & 0 & 0 & -3I_3^{1/3}\epsilon \end{pmatrix} \ , \tag{15.52}$$

which shows the breaking $USp(8) \to USp(6) \times USp(2)$.

15.0.4 Attractors in the 4 dimensional theory

In this section we reconsider the attractor solutions found in [33, 29]and we reformulate them in terms of the present formalism based on central charges. We separately examine the three “axion free" configurations.

Electric solution $Q = (p^0\,, q_i)$

Let us first compute the 4dim central charge for the electric charge configuration with vanishing axions; using (15.33) we find

$$Z_0 = \frac{i}{\sqrt{2}} e^{3\phi} p^0 \ , \qquad Z_a = \frac{1}{\sqrt{2}} e^{-\phi} q_I (a^{-1/2})^I{}_a . \tag{15.53}$$

The 4-dim potential is

$$V_{BH} = \frac{1}{2} e^{-2\phi} V_5^e + \frac{1}{2} e^{6\phi} (p^0)^2 \ , \tag{15.54}$$

(where ϕ is connected to the volume used in ref.[29] by the formula $\mathcal{V} = e^{6\phi}$) and has the same critical points of the 5 dimensional potential, since

$$\frac{\partial V_{BH}}{\partial \lambda^I} = 0 \iff \frac{\partial V_5^e}{\partial \hat{\lambda}^I} = 0 \ , \qquad \forall\, I = 1, 2\,. \tag{15.55}$$

The attractor values of $\hat{\lambda}^I$ are still given by (15.47), while the ϕ field at the critical point is [29]

$$e^{8\phi}|_{crit.} = I_3^{2/3}(p^0)^{-2}. \tag{15.56}$$

This fixes the central charges at the attractor point to be

$$\begin{aligned} Z_0{}^{attr} &= \frac{i}{\sqrt{2}}|p^0 q_1 q_2 q_3|^{1/4} sign(p^0) = \frac{i}{2}|I_4|^{1/4} sign(p^0)\ , \\ Z_a{}^{attr} &= \frac{1}{\sqrt{2}} I_3^{-1/12}(p^0)^{1/4} q_I \frac{I_3^{1/3}}{q_I} = \frac{1}{2}|I_4|^{1/4}\ , \end{aligned} \tag{15.57}$$

where the quartic invariant is $I_4 = -4\,p^0 q_1 q_2 q_3$. So we find

$$Z_1^{crit} = Z_2^{crit} = Z_3^{crit} = \frac{1}{2}|I_4|^{1/4} \equiv Z\ , \qquad Z_0^{crit} = \frac{i}{2}|I_4|^{1/4} sign(p^0) \equiv i Z_0. \tag{15.58}$$

Let us define the 4d central charge matrix as

$$2 Z_{AB} = e_{AB} - i Z^0 \Omega\ , \tag{15.59}$$

where e_{AB} is the matrix in (15.43) in which, instead of Z_1^5, Z_2^5, Z_3^5 of the 5d theory, we now write the 4d Z_a's defined in (15.33). it can be readily seen that for axion free solutions eq. (15.59) correctly gives

$$V_{BH} = \sum_i |z_i|^2 = |Z_0|^2 + \sum_a |Z_a|^2 \tag{15.60}$$

where z_i's, for $i = 1, .., 4$, are the (complex skew-diagonal) elements of Z_{AB}. We then have

$$\begin{aligned} 2Z_{AB} &= \begin{pmatrix} Z\epsilon & 0 & 0 & 0 \\ 0 & Z\epsilon & 0 & 0 \\ 0 & 0 & Z\epsilon & 0 \\ 0 & 0 & 0 & -3Z\epsilon \end{pmatrix} + \begin{pmatrix} Z_0\epsilon & 0 & 0 & 0 \\ 0 & Z_0\epsilon & 0 & 0 \\ 0 & 0 & Z_0\epsilon & 0 \\ 0 & 0 & 0 & Z_0\epsilon \end{pmatrix} = \\ &= \begin{pmatrix} (Z+Z_0)\epsilon & 0 & 0 & 0 \\ 0 & (Z+Z_0)\epsilon & 0 & 0 \\ 0 & 0 & (Z+Z_0)\epsilon & 0 \\ 0 & 0 & 0 & (-3Z+Z_0)\epsilon \end{pmatrix}. \end{aligned} \tag{15.61}$$

Since (15.57) and (15.58) yield that $Z = |Z_0|$, depending on the choice $p^0 > 0$ or $p^0 < 0$, two different solutions arise. In fact,

$$Z + Z_0 = 0 \quad \to \quad Z_{AB} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2Z_0 \end{pmatrix} \otimes \epsilon\ , \tag{15.62}$$

gives the $\frac{1}{8}$-BPS solution when $p^0 < 0$ and shows $SU(6) \times SU(2)$ symmetry. Conversely,

$$Z = Z_0 \quad \rightarrow \quad Z_{AB} = \begin{pmatrix} Z_0 & 0 & 0 & 0 \\ 0 & Z_0 & 0 & 0 \\ 0 & 0 & Z_0 & 0 \\ 0 & 0 & 0 & -Z_0 \end{pmatrix} \otimes \epsilon \ , \tag{15.63}$$

is the non-BPS solution that corresponds to the choice $p^0 > 0$, with residual $USp(8)$ symmetry.

Magnetic solution $Q = (p_i \, , q^0)$

This case is symmetric to the electric solution of Section 15.0.4. If we take all positive magnetic charges, then the cubic invariant is $I_3 = p^1 p^2 p^3$, the quartic invariant is $I_4 = 4\, q_0\, p^1 p^2 p^3$ and the values of the critical 5d moduli are now (see eq. (5.3) of [29])

$$\hat{\lambda}^I = \frac{p^I}{I_3^{1/3}}. \tag{15.64}$$

The central charges for this configuration are, from (15.33),

$$Z_0 = \frac{1}{\sqrt{2}} e^{-3\phi} q_0 \ , \qquad\qquad Z_a = \frac{i}{\sqrt{2}} e^{\phi} p^I (a^{1/2})_I{}^a \ , \tag{15.65}$$

and the black hole potential is

$$V_{BH} = \frac{1}{2} e^{2\phi} V_5^m + \frac{1}{2} e^{-6\phi} (q_0)^2. \tag{15.66}$$

This gives the attractor value of the ϕ field as

$$e^{8\phi}|_{crit.} = I_3^{-2/3} (q_0)^2. \tag{15.67}$$

At the attractor point $(a^{1/2}_{crit.})_{IJ} = (\hat{\lambda}^I)^{-1} \delta_{IJ}$, and the magnetic central charges are

$$Z_a^{crit} = \frac{i}{\sqrt{2}} (I_3)^{1/4} |q_0|^{1/4} = \frac{i}{2} |I_4|^{1/4} \equiv iZ \ , \qquad a = 1, 2, 3. \tag{15.68}$$

We can then write the central charge matrix corresponding to the **27** representation in the normal frame as

$$e_{AB} = \begin{pmatrix} Z\epsilon & 0 & 0 & 0 \\ 0 & Z\epsilon & 0 & 0 \\ 0 & 0 & Z\epsilon & 0 \\ 0 & 0 & 0 & -3Z\epsilon \end{pmatrix}. \tag{15.69}$$

To describe the four dimensional solution we need the electric central charge, that at the attractor point is

$$Z_0^{crit} = \frac{1}{\sqrt{2}} (I_3)^{1/4} |q_0|^{1/4}\, sign(q_0) = \frac{1}{2} |I_4|^{1/4}\, sign(q_0) \equiv Z_0.$$

Then, using the definition(15.59) the complete 4d central charge matrix is

$$2Z_{AB} = i\begin{pmatrix} Z\epsilon & 0 & 0 & 0 \\ 0 & Z\epsilon & 0 & 0 \\ 0 & 0 & Z\epsilon & 0 \\ 0 & 0 & 0 & -3Z\epsilon \end{pmatrix} - i\begin{pmatrix} Z_0\epsilon & 0 & 0 & 0 \\ 0 & Z_0\epsilon & 0 & 0 \\ 0 & 0 & Z_0\epsilon & 0 \\ 0 & 0 & 0 & Z_0\epsilon \end{pmatrix} =$$

$$= e^{i\pi/2}\begin{pmatrix} (Z-Z_0)\epsilon & 0 & 0 & 0 \\ 0 & (Z-Z_0)\epsilon & 0 & 0 \\ 0 & 0 & (Z-Z_0)\epsilon & 0 \\ 0 & 0 & 0 & (-3Z-Z_0)\epsilon \end{pmatrix} . \tag{15.70}$$

The $sign(q_0)$ determines whether the solution is supersymmetric or not. We may have

$$q_0 > 0 \quad \to \quad Z = Z_0 \ ,$$

$$Z_{AB} = e^{i\pi/2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2Z_0 \end{pmatrix} \otimes \epsilon \tag{15.71}$$

which is a magnetic $\frac{1}{8}$-BPS solutions with $SU(6)\times SU(2)$ symmetry, or

$$q_0 < 0 \quad \to \quad Z = -Z_0 \ ,$$

$$Z_{AB} = e^{i\pi/2}\begin{pmatrix} -Z_0 & 0 & 0 & 0 \\ 0 & -Z_0 & 0 & 0 \\ 0 & 0 & -Z_0 & 0 \\ 0 & 0 & 0 & Z_0 \end{pmatrix} \otimes \epsilon \tag{15.72}$$

which is the non-BPS solution with $USp(8)$ symmetry. These solutions have the same Z_0 as the electric ones, but now the choice of positive q_0 charge leads to the supersymmetric solution while the negative q_0 charge gives the non-supersymmetric one, in contrast with what happened for the choice of p^0 in the electric case in eq. (15.62) and (15.63).

KK dyonic solution $Q = (p^0\,, q_0)$

This charge configuration also has vanishing axions, and the only non-zero charges give

$$\begin{gathered} Z_0^e = e^{-3\phi}q_0 \ , \quad Z_m^0 = e^{3\phi}p^0 \ , \\ \Downarrow \\ Z_0 = \tfrac{1}{\sqrt{2}}(e^{-3\phi}q_0 + ie^{3\phi}p^0). \end{gathered} \tag{15.73}$$

Since none of the 5 dimensional charges are turned on, the four dimensional black hole potential is

$$V_{BH} = \frac{1}{2}\left[e^{-6\phi}q_0^2 + e^{6\phi}(p^0)^2\right] \ , \tag{15.74}$$

which is extremized at the horizon by the value of the ϕ field

$$e^{6\phi}|_{crit.} = \left|\frac{q_0}{p^0}\right| . \tag{15.75}$$

We only focus on the case $p^0 > 0$ and $q_0 > 0$, since all the other choices are related to this by a duality rotation. Evaluating the central charge at the attractor point we find

$$Z_0^{crit} = \sqrt{|p^0 q_0|}\frac{1+i}{\sqrt{2}} = \sqrt{|p^0 q_0|}e^{i\pi/4}.$$

$$\frac{1}{\sqrt{2}}\sqrt{q_0^2 + (p^0)^2}\ e^{i\phi/4}\ , \tag{15.76}$$

Following the prescription in (15.59) we find that at the attractor point

$$\begin{aligned} 2Z_{AB} &= -iZ_0\Omega = \\ &= -ie^{i\pi/4}\begin{pmatrix} \sqrt{|p^0q_0|}\epsilon & 0 & 0 & 0 \\ 0 & \sqrt{|p^0q_0|}\epsilon & 0 & 0 \\ 0 & 0 & \sqrt{|p^0q_0|}\epsilon & 0 \\ 0 & 0 & 0 & \sqrt{|p^0q_0|}\epsilon \end{pmatrix} \end{aligned} \tag{15.77}$$

that gives a non-BPS 4 dimensional black hole with $I_4 = -(p^0q_0)^2$.

Note that eqs. (15.63), (15.72) and (15.77) imply that the sum of the phases of the four complex skew entries is π, as appropriate to a non-BPS $\mathcal{N} = 8$ solution [17]. Also, in all cases, $V_{BH}|_{crit.} = \sqrt{|I_4|}$.

$\mathcal{N} = 8$ and $\mathcal{N} = 2$ attractive orbits at $d = 5$ and $d = 4$

We now compare the different interpretations in the $\mathcal{N} = 8$ and $\mathcal{N} = 2$ theories of the critical points of the very same black hole $4d$ potential, in terms of the axion-free electric solution (sec. 15.0.4) as discussed in this paper and in ref. [29].

Since the "normal frame" solution is common to all symmetric spaces (with rank three), it can be regarded as the generating solution of any model. So we confine our attention to the exceptional $\mathcal{N} = 2$ (octonionic) $E_{7(-25)}$ model [39] which has a charge vector in $5d$ and $4d$ of the same dimension as in $\mathcal{N} = 8$ supergravity. At $d = 5$ the duality group is $E_{6(-26)}$, with moduli space of vector multiplets $E_{6(-26)}/F_4$.

It is known [24, 35] that in $d = 5$ there are two different charge orbits,

$$O^{\mathcal{N}=2}_{d=5,\,BPS} = \frac{E_{6(-26)}}{F_4}\ , \tag{15.78}$$

the BPS one, and the non BPS one

$$O^{\mathcal{N}=2}_{d=5,\,non-BPS} = \frac{E_{6(-26)}}{F_{4(-20)}}\ , \tag{15.79}$$

The latter one precisely corresponds to the non supersymmetric solution and to $(+ + \ -)$, $(- - \ +)$ signs of the q_1, q_2, q_3, charges (implying $\partial Z \neq 0$). For charges of the same sign $(+ + +)$, $(- - -)$ one has the $\frac{1}{8}$BPS solution ($\partial Z = 0$), as discussed in [29].

It is easy to see that in the $\mathcal{N} = 8$ theory all these solutions just interchange Z_1, Z_2, Z_3 and $Z_4 = -3Z_3$ but always give a normal frame matrix of the form

$$Z_{ab} = \begin{pmatrix} Z\epsilon & 0 & 0 & 0 \\ 0 & Z\epsilon & 0 & 0 \\ 0 & 0 & Z\epsilon & 0 \\ 0 & 0 & 0 & -3Z\epsilon \end{pmatrix}\ , \tag{15.80}$$

which has $USp(6) \times USp(2) \in F_{4(4)}$ as maximal symmetry. Another related observation is that while $E_{6(-26)}$ contains both F_4 and $F_{4(-20)}$, so that one expects two orbits and two classes of solution, in the $\mathcal{N} = 8$ case $E_{6(6)}$ contains only the non compact $F_{4(4)}$, thus only one class of solutions is possible.

These orbits and critical points at $d = 5$ have a further story when used to study the $d = 4$ critical points with axion free solutions as it is the case for the electric (p^0, q_1, q_2, q_3) configuration. Since in this case $I_4 = -4p^0 q_1 q_2 q_3$, in the $\mathcal{N} = 8$ case, once one choose q_1, q_2, $q_3 >$ 0, the $I_4 > 0$, $p^0 < 0$ solution is BPS, while the $I_4 < 0$, $p^0 > 0$ is non BPS.

Things again change in $\mathcal{N} = 2$ [37], when now we consider the solution embedded in the Octonionic model with $4d$ moduli space $E_{7(-25)}/E_6 \times U(1)$. A new non BPS orbit in $d = 4$ is generated, corresponding to $Z = 0$ ($\partial Z \neq 0$) solution, so three $4d$ orbits exist in this case depending whether the $(+++)$ and $(++-)$ solutions are combined with $-p^0 \lesseqgtr 0$. So

$$(+,+++) \quad \text{is BPS with } I_4 > 0\ , \quad \mathcal{O} = \frac{E_{7(-25)}}{E_6}\ , \tag{15.81}$$

$$(-,-++) \quad \text{is non BPS with } I_4 > 0\ , \quad \mathcal{O} = \frac{E_{7(-25)}}{E_{6(-14)}}\ , \tag{15.82}$$

$$(+,-++) \ \text{or}\ (-,+++) \quad \text{is non BPS with } I_4 < 0\ , \quad \mathcal{O} = \frac{E_{7(-25)}}{E_{6(-26)}}. \tag{15.83}$$

15.0.5 Maurer-Cartan equations of the four dimensional theory

Let us call Maurer-Cartan equations[16] those which give the derivative of the central charges (coset representatives) with respect to the moduli ϕ, a^I, λ^i. Using (15.21) we have

$$\begin{aligned} \partial_\phi Z^e_0 = -3Z^e_0\ , \qquad \partial_\phi Z^0_m = 3Z^0_m\ , \\ \partial_\phi Z^e_I = -Z^e_I\ , \qquad \partial_\phi Z^I_m = Z^I_m\ , \end{aligned} \tag{15.84}$$

and

$$\begin{aligned} \frac{\partial Z^e_0}{\partial a^I} = e^{-2\phi} Z^e_I\ , \qquad \frac{\partial Z^0_m}{\partial a^I} = 0\ , \\ \frac{\partial Z^I_m}{\partial a^J} = -\delta^I_J e^{-2\phi} Z^0_e\ , \qquad \frac{\partial Z^e_I}{\partial a^J} = -e^{-2\phi} d_{IJK} Z^K_m. \end{aligned} \tag{15.85}$$

In our notation the 5d metric a_{IJ}, $(I, J = 1, .., 27)$ can also be rewritten with a pair of anti-symmetric (traceless) indices

$$a_{\Lambda\Sigma,\Delta\Gamma} = L^{ab}{}_{\Lambda\Sigma} L_{\Delta\Gamma\, ab}\ , \tag{15.86}$$

where $L^{ab}{}_{\Lambda\Sigma}$ is the coset representative; in a fixed gauge (where a, b and Λ, Σ indices are identified)

$$L_I{}^a = (a^{1/2})_I{}^a\ , \qquad (\bar{L}_{Ia} = L^T_{Ia}) \tag{15.87}$$

The object $\bar{b}P_i \equiv a^{1/2}\partial_i a^{-1/2}$ can be regarded as the Maurer-Cartan connection (see reference [32]). In fact, by reminding that $Z^e_a = Z^e_I (a^{-1/2})^I{}_a$, we have $\partial_i Z^e_a = (\partial_i a^{-1/2})^I{}_a Z^e_I$ (since $\partial_i Z^e_I = 0$). Since we can also write

$$\partial_i Z^e_a = (\partial_i a^{-1/2})^I{}_a (a^{1/2})_I{}^b Z^e_b \tag{15.88}$$

we find that $\bar{b}P_{i,a}{}^b$ is such that

$$\partial_i Z^e_a = \bar{b}P_{i,a}{}^b Z^e_b. \tag{15.89}$$

Notice that using $\bar{b}P_{i,a}{}^b = Q_{i,a}{}^b + V_{i,a}{}^b$, we identify a connection which satisfies

$$\nabla_i Z^e_a = V_a{}^b Z^e_b\ , \tag{15.90}$$

with $\nabla_i = \partial_i - Q_i$.

Attractor equations from Maurer-Cartan equations

We can now use this formalism to write the attractor equations for the potential

$$V_{BH} = \frac{1}{2}(Z_0^e)^2 + \frac{1}{2}(Z_m^0)^2 + \frac{1}{2}Z_I^e a^{IJ} Z_J^e + \frac{1}{2}Z_m^I a_{IJ} Z_m^J. \tag{15.91}$$

By differentiating with respect to ϕ, a^I, λ^i, we get

$$\partial_\phi V_{BH} = -3(Z_0^e)^2 + 3(Z_m^0)^2 - Z_I^e a^{IJ} Z_J^e + Z_m^I a_{IJ} Z_m^J = 0\ , \tag{15.92}$$
$$\partial_{a^I} V_{BH} = e^{-2\phi}\left[Z_0^e Z_I^e - Z_J^e a^{JK} d_{IKL} Z_m^L - Z_m^0 a_{IJ} Z_m^J\right] = 0\ , \tag{15.93}$$
$$\partial_{\lambda^i} V_{BH} \equiv \partial_i V_{BH} = \frac{1}{2}Z_I^e\, \partial_i a^{IJ}\, Z_J^e + \frac{1}{2}Z_m^I\, \partial_i a_{IJ}\, Z_m^J = 0. \tag{15.94}$$

From (15.93) we see that a solution with $a^I = 0$ implies

$$\partial_{a^I} V_{BH}\big|_{a^I=0} = 0 = e^{-2\phi}\left[e^{-4\phi} q_0 q_I - q_J a^{JK} d_{IKL} p^L - e^{4\phi} p^0 a_{IJ} p^J\right] = 0\ , \tag{15.95}$$

which is trivially satisfied if we set $\neq 0$ (q_0, p^0) or (q_0, p^I) or (p^0, q_I).

From (15.92) we see that for an axion-free solution, if $Z_0^e, Z_m^I = 0$, we get

$$3(Z_m^0)^2 = Z_I^e a^{IJ} Z_J^e\ , \tag{15.96}$$

and if a_{IJ} is diagonal, $I = J = 1, 2, 3$, we obtain

$$3(Z_m^0)^2 = (Z_1^e)^2 a^{11} + (Z_2^e)^2 a^{22} + (Z_3^e)^2 a^{33}\ , \tag{15.97}$$

which is compatible with $Z_1^e = Z_2^e = Z_3^e = \pm Z_m^0$.

The derivative with respect to the 5d moduli λ^i, $i = 1, .., 42$ for $\mathcal{N} = 8$ theory, only receives contributions from the matrix a_{IJ}. Indeed since Z_I^e, Z_m^I do not depend on the λ^i(see eq.15.21), one finds

$$\partial_i V_4 = 0 = Z_I^e\, \partial_i a^{IJ}\, Z_J^e + Z_m^I\, \partial_i a_{IJ}\, Z_m^J. \tag{15.98}$$

By rewriting the charges multiplied by $(a^{-1/2})^I{}_a$ and $(a^{1/2})_I{}^a$ so that

$$Z_a^e \equiv Z_I^e\, (a^{-1/2})^I{}_a\ , \qquad Z_m^a = Z_m^I (a^{1/2})^a{}_I\ , \tag{15.99}$$

we have

$$\begin{aligned} \partial_i Z_a^e &= \bar{b}P_{i,a}{}^b Z_b^e\ , &\qquad \bar{b}P_{i,a}{}^b &= \partial_i (a^{-1/2})^I{}_a (a^{1/2})_I{}^b\ , \\ \partial_i Z_m^a &= \bar{b}P_i{}^a{}_b Z_m^b\ , &\qquad \bar{b}P_i{}^a{}_b &= \partial_i (a^{1/2})_I{}^a (a^{-1/2})^I{}_b\ , \end{aligned} \tag{15.100}$$

where $\bar{b}P_i^a{}_b = -\bar{b}P_{i\,b}{}^a$ since $\partial_i(Z_a^e Z_m^a) = 0$. Then we also have

$$\begin{aligned} \partial_i (Z_a^e Z_a^e) &= Z_a^e (\bar{b}P_{ia}{}^b) Z_b^e = \\ &= Z_a^e \bar{b}P_{i,ab} Z_b^e = \\ &= Z_a^e \bar{b}P_{i\,(ab)} Z_b^e = 0\ , \end{aligned} \tag{15.101}$$

and if we split $\bar{b}P_{i,ab} = Q_{i\,[ab]} + V_{i\,(ab)}$, with

$$\begin{aligned} \bar{b}P_i{}^a{}_b &= Q_i{}^a{}_b + V_i{}^a{}_b\ , \\ \bar{b}P_{i,a}{}^b &= Q_{i,a}{}^b - V_{i,a}{}^b\ , \end{aligned} \tag{15.102}$$

the critical condition implies

$$\partial_i(Z^e Z^e) = Z^e_a V_{i\,(ab)} Z^e_b = 0 \ , \tag{15.103}$$

and the analogue equation for magnetic charges

$$\partial_i(Z^m Z^m) = Z^a_m V_{i\,(ab)} Z^b_m = 0 \ , \tag{15.104}$$

so that only the vielbein $V_{i\,,ab}$ enters in the equations of motion.

The criticality condition on the potential of eq. (15.98) now gives

$$\partial_i V_{BH} = 0 \quad \to \quad Z^e_a V_i{}^{ab} Z^e_b + Z^a_m V_{i,\,ab} Z^b_m = 0 \ , \tag{15.105}$$

thus, for electric configurations ($Z^b_m = 0$) with $a^I = 0$,

$$Z^e_a V_i{}^{ab} Z^e_b = 0. \tag{15.106}$$

Comparing results of [38] with our formulæ we see that $V_1,\ V_2,\ V_3$, with $V_1 + V_2 + V_3 = 0$, in the case where the metric a_{IJ} is diagonal, correspond to

$$(a^{-1/2})^I{}_a \partial_i (a^{1/2})_J{}^a = (a^{-1/2})^I \partial_i (a^{1/2})_I = \bar{b}P_i{}^I{}_I = V_i{}^I{}_I \equiv V_i{}^I \ , \tag{15.107}$$

where $(a^{-1/2})^I{}_I \equiv (a^{-1/2})^I$, $(a^{1/2})_I{}^I \equiv (a^{1/2})_I$, $I = 1, 2, 3$, and using (15.45) we find

$$\begin{aligned} V^I_1 &= \left(\frac{1}{\hat{\lambda}_1} \,, 0 \,, -\frac{1}{\hat{\lambda}_1} \right) \ , \\ V^I_2 &= \left(0 \,, \frac{1}{\hat{\lambda}_2} \,, -\frac{1}{\hat{\lambda}_2} \right) . \end{aligned} \tag{15.108}$$

Indeed,

$$\sum_{i=1,2,3} V^I_i = 0 \ , \tag{15.109}$$

so, by using eq. (2.31)-(2.33) of ref. [38] one gets the desired result. In fact, using the definitions of $\bar{b}P^I_1$ and $\bar{b}P^I_2$ we get from the $\hat{\lambda}^i$ equations of motion

$$\sum_I Z^e_I V^I_i Z^e_I = 0 \ , \tag{15.110}$$

which explicitly gives

$$\begin{aligned} Z^e_1 Z^e_1 - Z^e_3 Z^e_3 &= 0 \ , \\ Z^e_2 Z^e_2 - Z^e_3 Z^e_3 &= 0 \ , \end{aligned} \tag{15.111}$$

whose solution, combined with eq. (15.97), gives

$$\begin{gathered} (Z^e_1)^2 = (Z^e_2)^2 = (Z^e_3)^2 \ = (Z^0_m)^2 \ , \\ \Downarrow \\ Z^e_1 = Z^e_2 = Z^e_3 = \pm Z^0_m \ , \end{gathered} \tag{15.112}$$

all the other sign choices being equivalent in the 5d theory.

15.1 Self-Duality in Nonlinear Electromagnetism by Gaillard Zumino

15.1.1 Duality rotations in four dimensions

The invariance of Maxwell's equations under "duality rotations" has been known for a long time. In relativistic notation these are rotations of the electromagnetic field strength $F_{\mu\nu}$ into its dual, which is defined by

$$\tilde{F}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\lambda\sigma}F^{\lambda\sigma}, \quad \tilde{\tilde{F}}_{\mu\nu} = -F_{\mu\nu}. \tag{15.113}$$

This invariance can be extended to electromagnetic fields in interaction with the gravitational field, which does not transform under duality. It is present in ungauged extended supergravity theories, in which case it generalizes to a nonabelian group [1]. In [2, 3] we studied the most general situation in which duality invariance of this type can occur. More recently [4] the duality invariance of the Born-Infeld theory, suitably coupled to the dilaton and axion [5], has been studied in considerable detail. In the present note we will show that most of the results of [4, 5] follow quite easily from our earlier general discussion. We shall also present some new results that were not made explicit in [2, 3], especially some properties of the scalar fields.

We begin by recalling and completing some basic results of our paper [2, 3]. Consider a Lagrangian which is a function of n real field strengths $F^a_{\mu\nu}$ and of some other fields χ^i and their derivatives $\chi^i_\mu = \partial_\mu \chi^i$:

$$L = L\left(F^a, \chi^i, \chi^i_\mu\right). \tag{15.114}$$

Since

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu, \tag{15.115}$$

we have the Bianchi identities

$$\partial^\mu \tilde{F}^a_{\mu\nu} = 0. \tag{15.116}$$

On the other hand, if we define

$$\tilde{G}^a_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\lambda\sigma}G^{a\lambda\sigma} \equiv 2\frac{\partial L}{\partial F^{\mu\nu}_a}, \tag{15.117}$$

we have the equations of motion

$$\partial^\mu \tilde{G}^a_{\mu\nu} = 0. \tag{15.118}$$

We consider an infinitesimal transformation of the form

$$\delta \begin{pmatrix} F \\ G \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} F \\ G \end{pmatrix}, \tag{15.119}$$

$$\delta\chi^i = \xi^i(\chi), \tag{15.120}$$

where A, B, C, D are real $n \times n$ constant infinitesimal matrices and $\xi^i(\chi)$ functions of the fields χ^i (but not of their derivatives), and ask under what circumstances the system of the equations of motion (15.181) and (15.183), as well as the equation of motion for the fields χ^i are invariant. The analysis of [2] shows that this is true if the matrices satisfy

$$A^T = -D, \quad B^T = B, \quad C^T = C, \tag{15.121}$$

(where the superscript T denotes the transposed matrix) and in addition the Lagrangian changes under (15.184) and (15.185) as

$$\delta L = \frac{1}{4}\left(FC\tilde{F} + GB\tilde{G}\right). \tag{15.122}$$

The relations (15.186) show that (15.184) is an infinitesimal transformation of the real non-compact symplectic group $Sp(2n, R)$ which has $U(n)$ as maximal compact subgroup. The finite form is

$$\begin{pmatrix} F' \\ G' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} F \\ G \end{pmatrix}, \tag{15.123}$$

where the $n \times n$ real submatrices satisfy

$$c^T a = a^T c, \quad b^T d = d^T b, \quad d^T a - b^T c = 1. \tag{15.124}$$

Notice that the Lagrangian is not invariant. In [2] we showed, however, that the derivative of the Lagrangian with respect to an invariant parameter *is* invariant. The invariant parameter could be a coupling constant or an external background field, such as the gravitational field, which does not change under duality rotations. It follows that the energy-momentum tensor, which can be obtained as the variational derivative of the Lagrangian with respect to the gravitational field, is invariant under duality rotations. No explicit check of its invariance, as was done in [4]–[8], is necessary.

The symplectic transformation (15.188) can be written in a complex basis as

$$\begin{pmatrix} F' + iG' \\ F' - iG' \end{pmatrix} = \begin{pmatrix} \phi_0 & \phi_1^* \phi_1 & \phi_0^* \end{pmatrix} \begin{pmatrix} F + iG \\ F - iG \end{pmatrix}, \tag{15.125}$$

where $*$ means complex conjugation and the submatrices satisfy

$$\phi_0^T \phi_1 = \phi_1^T \phi_0, \quad \phi_0^\dagger \phi_0 - \phi_1^\dagger \phi_1 = 1. \tag{15.126}$$

The relation between the real and the complex basis is

$$\begin{aligned} 2a &= \phi_0 + \phi_0^* + \phi_1 + \phi_1^*, \quad & -2ib &= \phi_0 - \phi_0^* + \phi_1 - \phi_1^*, \\ 2ic &= \phi_0 - \phi_0^* - \phi_1 + \phi_1^*, \quad & 2d &= \phi_0 + \phi_0^* - \phi_1 - \phi_1^*. \end{aligned} \tag{15.127}$$

In [2, 3] we also described scalar fields valued in the quotient space $Sp(2n, R)/U(n)$. The quotient space can be parameterized by a complex symmetric $n \times n$ matrix $K = K^T$ whose real part has positive eigenvalues, or equivalently by a complex symmetric matrix $Z = Z^T$ such that $Z^\dagger Z$ has eigenvalues smaller than 1. They are related by

$$K = \frac{1 - Z^*}{1 + Z^*}, \quad Z = \frac{1 - K^*}{1 + K^*}. \tag{15.128}$$

These formulae are the generalization of the well-known map between the Lobachevskii unit disk and the Poincaré upper half-plane: Z corresponds to the single complex variable parameterizing the unit disk; iK to the one parameterizing the upper half plane.

Under $Sp(2n, R)$

$$K \to K' = (-ic + dK)(a + ibK)^{-1}, \quad Z \to Z' = (\phi_1 + \phi_0^* Z)(\phi_0 + \phi_1^* Z)^{-1}, \tag{15.129}$$

or, infinitesimally,

$$\delta K = -iC + DK - KA - iKBK, \quad \delta Z = V + T^* Z - ZT - iZV^* Z, \tag{15.130}$$

where

$$T = -T^\dagger, \quad V = V^T. \tag{15.131}$$

The invariant nonlinear kinetic term for the scalar fields can be obtained from the Kähler metric [9]

$$\mathrm{Tr}\left(dK^* \frac{1}{K + K^*} dK \frac{1}{K + K^*} \right) = \mathrm{Tr}\left(dZ \frac{1}{1 - Z^* Z} dZ^* \frac{1}{1 - Z Z^*} \right) \tag{15.132}$$

which follows from the Kähler potential

$$\operatorname{Tr}\ln\left(1-ZZ^{*}\right) \quad \text{or} \quad \operatorname{Tr}\ln(K+K^{*}), \tag{15.133}$$

which are equivalent up to a Kähler transformation. It is not hard to show that the metric (15.200) is positive definite. Throughout this paper we assume a flat background space-time metric; the generalization to a nonvanishing gravitational field is straightforward [2]–[5].

15.1.2 Born-Infeld theory

As a particularly simple example we consider the case when there is only one tensor $F_{\mu\nu}$ and no additional fields. Our equations become

$$\tilde{G} = 2\frac{\partial L}{\partial F}, \tag{15.134}$$

$$\delta F = \lambda G, \quad \delta G = -\lambda F \tag{15.135}$$

and

$$\delta L = \frac{1}{4}\lambda\left(G\tilde{G} - F\tilde{F}\right). \tag{15.136}$$

We have restricted the duality transformation to the compact subgroup $U(1) \cong SO(2)$, as appropriate when no additional fields are present. So $A = D = 0,\ B = -C = \lambda$.

Since L is a function of F alone, we can also write

$$\delta L = \delta F\frac{\partial L}{\partial F} = \lambda G\frac{1}{2}\tilde{G}. \tag{15.137}$$

Comparing (15.204) and (15.205), which must agree, we find

$$G\tilde{G} + F\tilde{F} = 0. \tag{15.138}$$

Together with (15.202), this is a partial differential equation for $L(F)$, which is the condition for the theory to be duality invariant. If we introduce the complex field

$$M = F - iG, \tag{15.139}$$

(15.206) can also be written as

$$M\widetilde{M^{*}} = 0. \tag{15.140}$$

Clearly, Maxwell's theory in vacuum satisfies (15.206), or (15.208), as expected. A more interesting example is the Born-Infeld theory [7], given by the Lagrangian

$$L = \frac{1}{g^2}\left(-\Delta^{\frac{1}{2}} + 1\right), \tag{15.141}$$

where

$$\Delta = -\det\left(\eta_{\mu\nu} + gF_{\mu\nu}\right) = 1 + \frac{1}{2}g^2F^2 - g^4\left(\frac{1}{4}F\tilde{F}\right)^2. \tag{15.142}$$

For small values of the coupling constant g (or for weak fields) L approaches the Maxwell Lagrangian. We shall use the abbreviation

$$\beta = \frac{1}{4}F\tilde{F}. \tag{15.143}$$

Then

$$\frac{\partial\Delta}{\partial F} = g^2F - \beta g^4\tilde{F}, \tag{15.144}$$

$$\tilde{G} = 2\frac{\partial L}{\partial F} = -\Delta^{-\frac{1}{2}}\left(F - \beta g^2 \tilde{F}\right), \tag{15.145}$$

and

$$G = \Delta^{-\frac{1}{2}}\left(\tilde{F} + \beta g^2 F\right). \tag{15.146}$$

Using (15.213) and (15.214), it is very easy to check that $G\tilde{G} = -F\tilde{F}$: the Born-Infeld theory is duality invariant. It is also not too difficult to check that $\partial L/\partial g^2$ is actually *invariant* under (15.203) and the same applies to $L - \frac{1}{4}F\tilde{G}$ (which in this case turns out to be equal to $-g^2\partial L/\partial g^2$). These invariances are expected from our general theory.

It is natural to ask oneself whether the Born-Infeld theory is the most general physically acceptable solution of (15.206). This was investigated in [4] where a negative result was reached: more general Lagrangians satisfy (15.206), the arbitrariness depending on a function of one variable.

15.1.3 Schrödinger's formulation of Born's theory

Schrödinger [8] noticed that, for the Born-Infeld theory (15.209), F and G satisfy not only (15.206) [or (15.208)], but also the more restrictive relation

$$M\left(M\widetilde{M}\right) - \widetilde{M}M^2 = \frac{g^2}{8}\widetilde{M}^*\left(M\widetilde{M}\right)^2. \tag{15.147}$$

We have verified this by an explicit, although lengthy, calculation using (15.207), (15.213), (15.214) and (15.210). Schrödinger did not give the details of the calculation, presenting instead convincing arguments based on particular choices of reference systems. One can write (15.215) as

$$\frac{\partial \mathcal{L}}{\partial M} = g^2\widetilde{M}^*, \tag{15.148}$$

where

$$\mathcal{L} = 4\frac{M^2}{\left(M\widetilde{M}\right)}, \tag{15.149}$$

and Schrödinger proposed $\mathcal{L}$ as the Lagrangian of the theory, instead of (15.209). Of course, $\mathcal{L}$ is a Lagrangian in a different sense than L, which is a field Lagrangian in the usual sense. Multiplying (15.215) by M and saturating the unwritten indices $\mu\nu$, the left hand side vanishes, so that (15.208) follows. Using (15.215) it is easy to see that $\mathcal{L}$ is pure imaginary: $\mathcal{L} = -\mathcal{L}^*$. Schrödinger also pointed out that, if we introduce a map

$$\frac{1}{g^2}\frac{\partial \mathcal{L}}{\partial M} = f(M), \tag{15.150}$$

so that (15.215) or (15.216) can be written as

$$f(M) = \widetilde{M}^*, \tag{15.151}$$

the square of the map is the identity map

$$f\left(f(M)\right) = M. \tag{15.152}$$

This, together with the properties

$$f(\widetilde{M}) = -\tilde{f}(M), \quad f(M^*) = f(M)^*, \tag{15.153}$$

ensures the consistency of (15.215). Schrödinger used the Lagrangian (15.217) to construct a conserved, symmetric energy-momentum tensor. We have checked that, when suitably normalized, his energy-momentum tensor agrees with that of Born and Infeld up to an additive term proportional to $\eta_{\mu\nu}$.

Schrödinger's formulation is very clever and elegant and it has the advantage of being *manifestly* covariant under the duality rotation $M \to Me^{i\lambda}$ which is the finite form of (15.203). It is also likely that, as he seems to imply, his formulation is fully equivalent to the Born-Infeld theory (15.209), which would mean that the more restrictive equation (15.215) eliminates the remaining ambiguity in the solutions of (15.208). This virtue could actually be a weakness if one is looking for more general duality invariant theories.

15.1.4 Axion, dilaton and $SL(2,R)$

It is well known that, if there are additional scalar fields which transform nonlinearly, the compact group duality invariance can be enhanced to a duality invariance under a larger noncompact group (see, *e.g.*, [2] and references therein). In the case of the Born-Infeld theory, just as for Maxwell's theory, one complex scalar field suffices to enhance the $U(1) \cong SO(2)$ invariance to the $SU(1,1) \cong SL(2,R)$ noncompact duality invariance. This is pointed out in [5], but it also follows the considerations of our paper [2]. We shall use the letter S instead of K for the scalar field, which, in the example under consideration, is a single complex field, not an $n \times n$ matrix. In today's more standard notation

$$S = S_1 - iS_2 = e^{-\phi} - ia, \quad S_1 > 0, \tag{15.154}$$

where ϕ is the dilaton and a is the axion. For $SL(2,R) \cong Sp(2,R)$, the matrices A, B, C, D are real numbers and $A = -D$, B and C are independent. Then the infinitesimal $SL(2,R)$ transformation is

$$\delta S = -2AS - iBS^2 - iC. \tag{15.155}$$

For the $SO(2) \cong U(1)$ subgroup, $A = 0,\ B = -C = \lambda$,

$$\delta S = -i\lambda S^2 + i\lambda. \tag{15.156}$$

The scalar kinetic term, proportional to

$$\frac{\partial_\mu S^* \partial^\mu S}{(S + S^*)^2}, \tag{15.157}$$

is invariant under the nonlinear transformation (15.247) which, in terms of S_1, S_2, takes the form

$$\delta S_1 = -2AS_1 - iBS_1S_2, \quad \delta S_2 = -2AS_2 + B\left(S_1^2 - S_2^2\right) + C. \tag{15.158}$$

The full noncompact duality transformation on $F_{\mu\nu}$ is now

$$\delta F = AF + BG, \quad \delta G = DF + DG, \quad D = -A, \tag{15.159}$$

and we are seeking a Lagrangian $\hat{L}(F,S)$ which satisfies

$$\delta\hat{L} = \frac{1}{4}\left(FC\tilde{F} + GB\tilde{G}\right), \tag{15.160}$$

where

$$\delta\hat{L} = \delta F\frac{\partial\hat{L}}{\partial F} + \delta S_1\frac{\partial\hat{L}}{\partial S_1} + \delta S_2\frac{\partial\hat{L}}{\partial S_2}, \tag{15.161}$$

and now

$$\tilde{G} = 2\frac{\partial \hat{L}}{\partial F}. \tag{15.162}$$

Equating (15.253) and (15.254) we see that $\hat{L}$ must satisfy

$$\frac{1}{4}\left(BG\tilde{G} - CF\tilde{F}\right) + \frac{1}{2}AF\tilde{G} + \delta S_1 \frac{\partial \hat{L}}{\partial S_1} + \delta S_2 \frac{\partial \hat{L}}{\partial S_2} = 0. \tag{15.163}$$

This equation can be solved as follows. Assume that $L(\mathcal{F})$ satisfies (15.202) and (15.206), *i.e.*

$$\mathcal{G}\tilde{\mathcal{G}} + \mathcal{F}\tilde{\mathcal{F}} = 0, \tag{15.164}$$

where

$$\tilde{\mathcal{G}} = 2\frac{\partial \mathcal{L}}{\partial \mathcal{F}}. \tag{15.165}$$

For instance, the Born-Infeld Lagrangian $L(\mathcal{F})$ does this. Then

$$\hat{L}(S, F) = L(S_1^{\frac{1}{2}} F) + \frac{1}{4} S_2 F\tilde{F} \tag{15.166}$$

satisfies (15.256). Indeed

$$\frac{\partial \hat{L}(S, F)}{\partial F} = \frac{\partial L}{\partial \mathcal{F}} S_1^{\frac{1}{2}} + \frac{1}{2} S_2 \tilde{F}. \tag{15.167}$$

So

$$\tilde{G} = \tilde{\mathcal{G}} S_1^{\frac{1}{2}} + S_2 \tilde{F}, \tag{15.168}$$

$$G = \mathcal{G} S_1^{\frac{1}{2}} + S_2 F, \tag{15.169}$$

where we have defined

$$\mathcal{F} = S_1^{\frac{1}{2}} F, \tag{15.170}$$

and $\tilde{\mathcal{G}}$ is given by (15.258). Now

$$G\tilde{G} = \mathcal{G}\tilde{\mathcal{G}} S_1 + S_2^2 F\tilde{F} + 2S_2 \mathcal{F}\tilde{\mathcal{G}}. \tag{15.171}$$

Using (15.257) in this equation we find

$$G\tilde{G} = \left(S_2^2 - S_1^2\right) F\tilde{F} + 2S_2 \mathcal{F}\tilde{\mathcal{G}}. \tag{15.172}$$

We also have

$$F\tilde{G} = \mathcal{F}\tilde{\mathcal{G}} + S_2 F\tilde{F}. \tag{15.173}$$

On the other hand, since

$$\frac{\partial L}{\partial S_1^{\frac{1}{2}}} = \frac{\partial \mathcal{L}}{\partial \mathcal{F}} F = \frac{1}{2}\tilde{\mathcal{G}} F, \tag{15.174}$$

we obtain

$$\frac{\partial \hat{L}}{\partial S_1} = \frac{\partial L}{\partial S_1^{\frac{1}{2}}} \frac{1}{2} S_1^{-\frac{1}{2}} = \frac{1}{4}\tilde{\mathcal{G}} S_1^{-\frac{1}{2}} F = \frac{1}{4}\tilde{\mathcal{G}}\mathcal{F} S_1^{-1}. \tag{15.175}$$

In addition

$$\frac{\partial \hat{L}}{\partial S_2} = \frac{1}{4} F\tilde{F}. \tag{15.176}$$

Using (15.265), (15.266), (15.268) and (15.269), together with (15.251), we see that (15.256) is satisfied. It is easy to check that the scale invariant combinations $\mathcal{F}$ and $\mathcal{G}$, given by (15.263) and (15.258) have the very simple transformation law

$$\delta\mathcal{F} = S_1 B\mathcal{G}, \quad \delta\mathcal{G} = -S_1 B\mathcal{F}, \tag{15.177}$$

i.e., they transform according to the $U(1) \cong SO(2)$ compact subgroup just as F and G in (15.203), but with the parameter λ replaced by $S_1 B$. If $L(\mathcal{F})$ is the Born-Infeld Lagrangian, the theory with scalar fields given by $\hat{L}$ in (15.259) can also be reformulated à la Schrödinger. From (15.262) and (15.263) solve for $\mathcal{F}$ and $\mathcal{G}$ in terms of F, G, S_1 and S_2. Then $\mathcal{M} = \mathcal{F} - i\mathcal{G}$ must satisfy the same equation (15.215) that M does when no scalar fields are present.

15.1.5 Connections to string theory

The duality rotations considered here are relevant to effective field theories from superstrings. The supersymmetric extension [17] of the Lagrangian (15.259) with $L(\mathcal{F}) = -\frac{1}{4}\mathcal{F}^2$ describes the dilaton plus Yang-Mills sector of effective $N = 1$ supergravity theories obtained from superstrings in the weak coupling ($S_1 \to \infty$) limit. The $SL(2, Z)$ subgroup of $SL(2, R)$ that is generated by the elements $4\pi S \to 1/4\pi S$ and $S \to S - i/4\pi$ relates different string theories [12] to one another. The generalization of [2] to two dimensional theories [19] has been used to derive the Kähler potential for moduli and matter fields in effective field theories from superstrings. In this case the scalars are valued on a coset space $\mathcal{K}/\mathcal{H}$, $\mathcal{K} \in SO(n, n)$, $\mathcal{H} \in SO(n) \times SO(n)$. The kinetic energy is invariant under $\mathcal{K}$, and the full classical theory is invariant under a subgroup of $\mathcal{K}$. String loop corrections reduces the invariance to a discrete subgroup that contains the $SL(2, Z)$ group generated by $T \to 1/T$, $T \to T - i$, where T is the squared radius of compactification in string units.

15.2 Nonlinear Electromagnetic Self-Duality and Legendre Transformations by Gaillard Zumino

Abstract

We discuss continuous duality transformations and the properties of classical theories with invariant interactions between electromagnetic fields and matter. The case of scalar fields is treated in some detail. Special discrete elements of the continuous group are shown to be related to the Legendre transformation with respect to the field strengths.

15.2.1 Duality rotations in four dimensions

The invariance of Maxwell's equations under "duality rotations" has been known for a long time. In relativistic notation these are rotations of the electromagnetic field strength $F_{\mu\nu}$ into its dual, which is defined by

$$\tilde{F}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\lambda\sigma}F^{\lambda\sigma}, \quad \tilde{\tilde{F}}_{\mu\nu} = -F_{\mu\nu}. \tag{15.178}$$

This invariance can be extended to electromagnetic fields in interaction with the gravitational field, which does not transform under duality. It is present in ungauged extended supergravity theories, in which case it generalizes to a nonabelian group [1]. In [2, 3] we studied the most general situation in which classical duality invariance of this type can occur. More recently [4] the duality invariance of the Born-Infeld theory, suitably coupled to the dilaton and axion [5], has been studied in considerable detail. In the present note we will show that most of the results of [4, 5] follow quite easily from our earlier general discussion. We shall also present some new results.

We begin by recalling and completing some basic results of [2, 3, 6]. Consider a Lagrangian which is a function of n real field strengths $F^a_{\mu\nu}$ and of some other fields χ^i and their derivatives $\chi^i_\mu = \partial_\mu \chi^i$:

$$L = L\left(F^a, \chi^i, \chi^i_\mu\right). \tag{15.179}$$

Since

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu, \tag{15.180}$$

we have the Bianchi identities

$$\partial^\mu \tilde{F}^a_{\mu\nu} = 0. \tag{15.181}$$

On the other hand, if we define

$$\tilde{G}^a_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\lambda\sigma} G^{a\lambda\sigma} \equiv 2\frac{\partial L}{\partial F^{\mu\nu}_a}, \tag{15.182}$$

we have the equations of motion

$$\partial^\mu \tilde{G}^a_{\mu\nu} = 0. \tag{15.183}$$

We consider an infinitesimal transformation of the form

$$\delta \begin{pmatrix} G \\ F \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} G \\ F \end{pmatrix}, \tag{15.184}$$

$$\delta\chi^i = \xi^i(\chi), \tag{15.185}$$

where A, B, C, D are real $n \times n$ constant infinitesimal matrices and $\xi^i(\chi)$ functions of the fields χ^i (but not of their derivatives), and ask under what circumstances the system of the equations of motion (15.181) and (15.183), as well as the equation of motion for the fields χ^i are invariant. The analysis of [2] shows that this is true if the matrices satisfy

$$A^T = -D, \quad B^T = B, \quad C^T = C, \tag{15.186}$$

(where the superscript T denotes the transposed matrix) and in addition the Lagrangian changes under (15.184) and (15.185) as

$$\delta L = \frac{1}{4}\left(FB\tilde{F} + GC\tilde{G}\right). \tag{15.187}$$

The relations (15.186) show that (15.184) is an infinitesimal transformation of the real non-compact symplectic group $Sp(2n, R)$ which has $U(n)$ as maximal compact subgroup. The finite form is

$$\begin{pmatrix} G' \\ F' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} G \\ F \end{pmatrix}, \tag{15.188}$$

where the $n \times n$ real submatrices satisfy

$$c^T a = a^T c, \quad b^T d = d^T b, \quad d^T a - b^T c = 1. \tag{15.189}$$

For the $U(n)$ subgroup, one has in addition

$$A = D, \quad B = -C, \tag{15.190}$$

or, in finite form,

$$a = d, \quad b = -c. \tag{15.191}$$

Notice that the Lagrangian is not invariant. In [2] we showed, however, that the derivative of the Lagrangian with respect to an invariant parameter *is* invariant. The invariant parameter could be a coupling constant or an external background field, such as the gravitational field, which does not change under duality rotations. It follows that the energy-momentum tensor, which can be obtained as the variational derivative of the Lagrangian with respect to the gravitational field, is invariant under duality rotations. No explicit check of its invariance, as was done in [4, 5, 7, 8], is necessary. Using (15.184) and (15.186) it is easy to verify that

$$\delta\left(L - \frac{1}{4}F\tilde{G}\right) = \delta L - \frac{1}{4}\left(FB\tilde{F} + GC\tilde{G}\right), \tag{15.192}$$

so (15.187) is *equivalent* to the invariance of $L - \frac{1}{4}F\tilde{G}$.

The symplectic transformation (15.188) can be written in a complex basis as

$$\begin{pmatrix} F' + iG' \\ F' - iG' \end{pmatrix} = \begin{pmatrix} \phi_0 & \phi_1^* \phi_1 & \phi_0^* \end{pmatrix} \begin{pmatrix} F + iG \\ F - iG \end{pmatrix}, \tag{15.193}$$

where $*$ means complex conjugation and the submatrices satisfy

$$\phi_0^T \phi_1 = \phi_1^T \phi_0, \quad \phi_0^\dagger \phi_0 - \phi_1^\dagger \phi_1 = 1. \tag{15.194}$$

The relation between the real and the complex basis is

$$\begin{aligned} 2a &= \phi_0 + \phi_0^* - \phi_1 - \phi_1^*, \quad 2ib = \phi_0 - \phi_0^* - \phi_1 + \phi_1^*, \\ -2ic &= \phi_0 - \phi_0^* + \phi_1 - \phi_1^*, \quad 2d = \phi_0 + \phi_0^* + \phi_1 + \phi_1^*. \end{aligned} \tag{15.195}$$

In [2, 3] we also described scalar fields valued in the quotient space $Sp(2n, R)/U(n)$. The quotient space can be parameterized by a complex symmetric $n \times n$ matrix $K = K^T$ whose real part has positive eigenvalues, or equivalently by a complex symmetric matrix $Z = Z^T$ such that $Z^\dagger Z$ has eigenvalues smaller than 1. They are related by

$$K = \frac{1 - Z^*}{1 + Z^*}, \quad Z = \frac{1 - K^*}{1 + K^*}. \tag{15.196}$$

These formulae are the generalization of the well-known map between the Lobachevskii unit disk and the Poincaré upper half-plane: Z corresponds to the single complex variable parameterizing the unit disk, iK to the one parameterizing the upper half plane.

Under $Sp(2n, R)$

$$K \to K' = (-ib + aK)(d + icK)^{-1}, \quad Z \to Z' = (\phi_1 + \phi_0^* Z)(\phi_0 + \phi_1^* Z)^{-1}, \tag{15.197}$$

or, infinitesimally,

$$\delta K = -iB + AK - KD - iKCK, \quad \delta Z = V + T^* Z - ZT - iZV^* Z, \tag{15.198}$$

where

$$T = -T^\dagger, \quad V = V^T. \tag{15.199}$$

The invariant nonlinear kinetic term for the scalar fields can be obtained from the Kähler metric [9]

$$\mathrm{Tr}\left(dK^* \frac{1}{K + K^*} dK \frac{1}{K + K^*}\right) = \mathrm{Tr}\left(dZ \frac{1}{1 - Z^* Z} dZ^* \frac{1}{1 - ZZ^*}\right) \tag{15.200}$$

which follows from the Kähler potential

$$\mathrm{Tr}\ln(1 - ZZ^*) \quad \text{or} \quad \mathrm{Tr}\ln(K + K^*), \tag{15.201}$$

which are equivalent up to a Kähler transformation. It is not hard to show that the metric (15.200) is positive definite. In this section the normalization of the fields $F^a_{\mu\nu}$ has been chosen to be canonical when iK is set equal to the unit matrix, *i.e.*, when the self-duality group reduces to the $U(n)$ subgroup; the full $Sp(2n, R)$ self-duality can be realized when the matrix K is a function of scalar fields. Throughout this paper we assume a flat background space-time metric; the generalization to a nonvanishing gravitational field is straightforward [2]–[5].

15.2.2 Born-Infeld theory

As a particularly simple example we consider the case when there is only one tensor $F_{\mu\nu}$ and no additional fields. Our equations become

$$\tilde{G} = 2\frac{\partial L}{\partial F}, \tag{15.202}$$

$$\delta F = \lambda G, \quad \delta G = -\lambda F \tag{15.203}$$

and

$$\delta L = \frac{1}{4}\lambda\left(G\tilde{G} - F\tilde{F}\right). \tag{15.204}$$

We have restricted the duality transformation to the compact subgroup $U(1) \cong SO(2)$, as appropriate when no additional fields are present. So $A = D = 0$, $C = -B = \lambda$.

Since L is a function of F alone, we can also write

$$\delta L = \delta F\frac{\partial L}{\partial F} = \lambda G\frac{1}{2}\tilde{G}. \tag{15.205}$$

Comparing (15.204) and (15.205), which must agree, we find

$$G\tilde{G} + F\tilde{F} = 0. \tag{15.206}$$

Together with (15.202), this is a partial differential equation for $L(F)$, which is the condition for the theory to be duality invariant. If we introduce the complex field

$$M = F - iG, \tag{15.207}$$

(15.206) can also be written as

$$M\widetilde{M^*} = 0. \tag{15.208}$$

Clearly, Maxwell's theory in vacuum satisfies (15.206), or (15.208), as expected. A more interesting example is the Born-Infeld theory [7], given by the Lagrangian

$$L = \frac{1}{g^2}\left(-\Delta^{\frac{1}{2}} + 1\right), \tag{15.209}$$

where

$$\Delta = -\det\left(\eta_{\mu\nu} + gF_{\mu\nu}\right) = 1 + \frac{1}{2}g^2F^2 - g^4\left(\frac{1}{4}F\tilde{F}\right)^2. \tag{15.210}$$

For small values of the coupling constant g (or for weak fields) L approaches the Maxwell Lagrangian. We shall use the abbreviation

$$\beta = \frac{1}{4}F\tilde{F}. \tag{15.211}$$

Then

$$\frac{\partial\Delta}{\partial F} = g^2F - \beta g^4\tilde{F}, \tag{15.212}$$

$$\tilde{G} = 2\frac{\partial L}{\partial F} = -\Delta^{-\frac{1}{2}}\left(F - \beta g^2\tilde{F}\right), \tag{15.213}$$

and

$$G = \Delta^{-\frac{1}{2}}\left(\tilde{F} + \beta g^2F\right). \tag{15.214}$$

Using (15.213) and (15.214), it is very easy to check that $G\tilde{G} = -F\tilde{F}$: the Born-Infeld theory is duality invariant. It is also not too difficult to check that $\partial L/\partial g^2$ is actually *invariant* under (15.203) and the same applies to $L - \frac{1}{4}F\tilde{G}$ (which in this case turns out to be equal to $-g^2\partial L/\partial g^2$). These invariances are expected from our general theory.

It is natural to ask oneself whether the Born-Infeld theory is the most general physically acceptable solution of (15.206). This was investigated in [4] where a negative result was reached: more general Lagrangians satisfy (15.206), the arbitrariness depending on a function of one variable. We discuss this in detail in Section 6.

15.2.3 Schrödinger's formulation of Born's theory

Schrödinger [8] noticed that, for the Born-Infeld theory (15.209), F and G satisfy not only (15.206) [or (15.208)], but also the more restrictive relation

$$M\left(M\widetilde{M}\right) - \widetilde{M}M^2 = \frac{g^2}{8}\widetilde{M}^*\left(M\widetilde{M}\right)^2. \tag{15.215}$$

We have verified this by an explicit, although lengthy, calculation using (15.207), (15.213), (15.214) and (15.210). Schrödinger did not give the details of the calculation, presenting instead convincing arguments based on particular choices of reference systems. One can write (15.215) as

$$\frac{\partial \mathcal{L}}{\partial M} = g^2\widetilde{M}^*, \tag{15.216}$$

where

$$\mathcal{L} = 4\frac{M^2}{\left(M\widetilde{M}\right)}, \tag{15.217}$$

and Schrödinger proposed $\mathcal{L}$ as the Lagrangian of the theory, instead of (15.209). Of course, $\mathcal{L}$ is a Lagrangian in a different sense than L, which is a field Lagrangian in the usual sense. Multiplying (15.215) by M and saturating the unwritten indices $\mu\nu$, the left hand side vanishes, so that (15.208) follows. Using (15.215) it is easy to see that $\mathcal{L}$ is pure imaginary: $\mathcal{L} = -\mathcal{L}^*$. Schrödinger also pointed out that, if we introduce a map

$$\frac{1}{g^2}\frac{\partial \mathcal{L}}{\partial M} = f(M), \tag{15.218}$$

so that (15.215) or (15.216) can be written as

$$f(M) = \widetilde{M}^*, \tag{15.219}$$

the square of the map is the identity map

$$f\left(f(M)\right) = M. \tag{15.220}$$

This, together with the properties

$$f(\widetilde{M}) = -\tilde{f}(M), \quad f(M^*) = f(M)^*, \tag{15.221}$$

ensures the consistency of (15.215). Schrödinger used the Lagrangian (15.217) to construct a conserved, symmetric energy-momentum tensor. We have checked that, when suitably normalized, his energy-momentum tensor agrees with that of Born and Infeld up to an additive term proportional to $\eta_{\mu\nu}$.

Schrödinger's formulation is very clever and elegant and it has the advantage of being *manifestly* covariant under the duality rotation $M \to Me^{i\lambda}$ which is the finite form of (15.203). It is also likely that, as he seems to imply, his formulation is fully equivalent to the Born-Infeld theory (15.209), which would mean that the more restrictive equation (15.215) eliminates the remaining ambiguity in the solutions of (15.208). This virtue could actually be a weakness if one is looking for more general duality invariant theories.

15.2.4 General solution of the self-duality equation

The self-duality equation (15.206) can be solved in general as follows. Assuming Lorentz invariance in four dimensional space-time, the Lagrangian must be a function of the two invariants

$$\alpha = \frac{1}{4}F^2, \quad \beta = \frac{1}{4}F\tilde{F}, \quad L = L(\alpha, \beta). \tag{15.222}$$

Now

$$\tilde{G} = 2\frac{\partial L}{\partial F} = L_\alpha F + L_\beta \tilde{F}, \quad G = -L_\alpha \tilde{F} + L_\beta F, \tag{15.223}$$

where we have used the standard notation $L_\alpha = \partial L/\partial\alpha,\ L_\beta = \partial L/\partial\beta$. Substituting these expressions in (15.206) we obtain

$$\left[(L_\beta)^2 - (L_\alpha)^2 + 1\right]\beta + 2L_\alpha L_\beta \alpha = 0. \tag{15.224}$$

This partial differential equation for L can be simplified by the change of variables

$$x = \alpha, \quad y = \left(\alpha^2 + \beta^2\right)^{\frac{1}{2}}, \tag{15.225}$$

which gives

$$(L_x)^2 - (L_y)^2 = 1. \tag{15.226}$$

Alternatively one can use the variables

$$p = \frac{1}{2}(x+y), \quad q = \frac{1}{2}(x-y), \tag{15.227}$$

to obtain the form

$$L_p L_q = 1. \tag{15.228}$$

The equation (15.226), or (15.228), has been studied extensively in mathematics and there are several methods to obtain its general solution [10]. (It is interesting that the same equation occurs in a study of 5-dimensional Born-Infeld theory [11].) In our case we must also impose the physical boundary condition that the Lagrangian should approximate the Maxwell Lagrangian

$$L_M = -\alpha = -x = -p - q \tag{15.229}$$

when the field strength F is small.

According to one of the methods given in Courant-Hilbert, the general solution of (15.228) is given by

$$L = \frac{2p}{v'(s)} + v(s), \tag{15.230}$$

$$q = \frac{p}{[v'(s)]^2} + s, \tag{15.231}$$

where the arbitrary function $v(s)$ is determined by the initial values:

$$L(p=0, q) = v(q), \tag{15.232}$$

$$L_p(p=0, q) = \frac{1}{v'(q)}. \tag{15.233}$$

One must solve for $s(p, q)$ from (15.231) and substitute into (15.230). To verify [11] that these equations solve (15.228), differentiate (15.230) and (15.231):

$$dL = \frac{2dp}{v'} + \left(v' - \frac{2p}{[v']^2}v''\right)ds, \tag{15.234}$$

$$dq = \frac{dp}{v'^2} + \left(1 - \frac{2p}{[v']^3}v''\right)ds, \tag{15.235}$$

and eliminate ds between (15.234) and (15.235) to obtain

$$dL = \frac{1}{v'}dp + v'dq, \tag{15.236}$$

i.e.,

$$L_p = \frac{1}{v'}, \quad L_q = v', \quad L_pL_q = 1. \tag{15.237}$$

The condition that L should approach the Maxwell Lagrangian for small field strengths implies that

$$v(s) = L(p = 0, s) \cong -s \tag{15.238}$$

for small s.

It is trivial to check the above procedure for the Maxwell Lagrangian, and we shall not do it here. The Born-Infeld Lagrangian (with $g = 1$ for simplicity) is given by

$$L_{BI} = -\Delta^{\frac{1}{2}} + 1, \tag{15.239}$$

$$\Delta = (1 + 2p)(1 + 2q), \tag{15.240}$$

in terms of the variables p and q. Setting $p = 0$ we see that this corresponds to

$$v(s) = -(1 + 2s)^{\frac{1}{2}} + 1, \tag{15.241}$$

$$v'(s) = -(1 + 2s)^{-\frac{1}{2}}. \tag{15.242}$$

Then (15.231) gives

$$q = p(1 + 2s) + s, \tag{15.243}$$

which is solved by

$$s = \frac{q - p}{1 + 2p}, \quad 1 + 2s = \frac{1 + 2q}{1 + 2p}. \tag{15.244}$$

Using (15.230), we reconstruct the Lagrangian

$$L_{BI} = -2p\left(\frac{1 + 2q}{1 + 2p}\right)^{\frac{1}{2}} - \left(\frac{1 + 2q}{1 + 2p}\right)^{\frac{1}{2}} + 1 = -\left[(1 + 2p)(1 + 2q)\right]^{\frac{1}{2}} + 1. \tag{15.245}$$

Unfortunately, in spite of this elegant method for finding solutions of the self-duality equation, it seems very difficult to find new explicit solutions given in terms of simple functions. The reason is that, even for a simple function $v(s)$, solving the equation (15.231) for s gives complicated functions $s(p,q)$.

15.2.5 Axion, dilaton and $SL(2,R)$

It is well known that, if there are additional scalar fields which transform nonlinearly, the compact group duality invariance can be enhanced to a duality invariance under a larger non-compact group (see, *e.g.*, [2] and references therein). In the case of the Born-Infeld theory, just as for Maxwell's theory, one complex scalar field suffices to enhance the $U(1) \cong SO(2)$ invariance to the $SU(1,1) \cong SL(2,R)$ noncompact duality invariance. This is pointed out in [5], but it also follows from the considerations of our paper [2]. In the example under consideration, K is a single complex field, not an $n \times n$ matrix. In order to agree with today's more standard notation we shall use

$$S = iK = S_1 + iS_2 = a + ie^{-\phi}, \quad S_2 > 0, \tag{15.246}$$

where ϕ is the dilaton and a is the axion. For $SL(2,R) \cong Sp(2,R)$, the matrices A, B, C, D are real numbers and $A = -D$, B and C are independent. Then the infinitesimal $SL(2,R)$ transformation is

$$\delta S = B + 2AS - CS^2, \tag{15.247}$$

and the finite transformation is

$$S' = \frac{aS+b}{cS+d}, \quad ad - bc = 1. \tag{15.248}$$

For the $SO(2) \cong U(1)$ subgroup, $A = 0$, $C = -B = \lambda$,

$$\delta S = -\lambda - \lambda S^2. \tag{15.249}$$

The scalar kinetic term, proportional to

$$\frac{\partial_\mu S^* \partial^\mu S}{(S - S^*)^2}, \tag{15.250}$$

is invariant under the nonlinear transformation (15.247) which, in terms of S_1, S_2, takes the form

$$\delta S_1 = B + 2AS_1 - C\left(S_1^2 - S_2^2\right), \quad \delta S_2 = 2AS_2 - 2CS_1S_2. \tag{15.251}$$

Since the scalar kinetic term is separately invariant, we assume from now on that $\hat{L}(S,F)$ does not depend on the derivatives of S.

The full noncompact duality transformation on $F_{\mu\nu}$ is now

$$\delta G = AG + BF, \quad \delta F = CG + DF, \quad D = -A, \tag{15.252}$$

and we are seeking a Lagrangian $\hat{L}(S,F)$ which satisfies

$$\delta\hat{L} = \frac{1}{4}\left(FB\tilde{F} + GC\tilde{G}\right), \tag{15.253}$$

where

$$\delta\hat{L} = \delta F\frac{\partial\hat{L}}{\partial F} + \delta S_1\frac{\partial\hat{L}}{\partial S_1} + \delta S_2\frac{\partial\hat{L}}{\partial S_2}, \tag{15.254}$$

and now

$$\tilde{G} = 2\frac{\partial\hat{L}}{\partial F}. \tag{15.255}$$

Equating (15.253) and (15.254) we see that $\hat{L}$ must satisfy

$$\frac{1}{4}\left(CG\tilde{G} - BF\tilde{F}\right) - \frac{1}{2}AF\tilde{G} + \delta S_1\frac{\partial\hat{L}}{\partial S_1} + \delta S_2\frac{\partial\hat{L}}{\partial S_2} = 0. \tag{15.256}$$

This equation can be solved as follows. Assume that $L(\mathcal{F})$ satisfies (15.202) and (15.206), *i.e.*

$$\mathcal{G}\tilde{\mathcal{G}} + \mathcal{F}\tilde{\mathcal{F}} = 0, \tag{15.257}$$

where

$$\tilde{\mathcal{G}} = 2\frac{\partial L}{\partial\mathcal{F}}. \tag{15.258}$$

For instance, the Born-Infeld Lagrangian $L(\mathcal{F})$ does this. Then

$$\hat{L}(S,F) = L(S_2^{\frac{1}{2}}F) + \frac{1}{4}S_1F\tilde{F} \tag{15.259}$$

satisfies (15.256). Indeed

$$\frac{\partial \hat{L}(S,F)}{\partial F} = \frac{\partial L}{\partial \mathcal{F}} S_2^{\frac{1}{2}} + \frac{1}{2} S_1 \tilde{F}. \tag{15.260}$$

So

$$\tilde{G} = \tilde{\mathcal{G}} S_2^{\frac{1}{2}} + S_1 \tilde{F}, \tag{15.261}$$

$$G = \mathcal{G} S_2^{\frac{1}{2}} + S_1 F, \tag{15.262}$$

where we have defined

$$\mathcal{F} = S_2^{\frac{1}{2}} F, \tag{15.263}$$

and $\tilde{\mathcal{G}}$ is given by (15.258). Now

$$G\tilde{G} = \mathcal{G}\tilde{\mathcal{G}} S_2 + S_1^2 F\tilde{F} + 2S_1 \mathcal{F}\tilde{\mathcal{G}}. \tag{15.264}$$

Using (15.257) in this equation we find

$$G\tilde{G} = \left(S_1^2 - S_2^2\right) F\tilde{F} + 2S_1 \mathcal{F}\tilde{\mathcal{G}}. \tag{15.265}$$

We also have

$$F\tilde{G} = \mathcal{F}\tilde{\mathcal{G}} + S_1 F\tilde{F}. \tag{15.266}$$

On the other hand, since

$$\frac{\partial L}{\partial S_2^{\frac{1}{2}}} = \frac{\partial L}{\partial \mathcal{F}} F = \frac{1}{2}\tilde{\mathcal{G}} F, \tag{15.267}$$

we obtain

$$\frac{\partial \hat{L}}{\partial S_2} = \frac{\partial L}{\partial S_2^{\frac{1}{2}}} \frac{1}{2} S_2^{-\frac{1}{2}} = \frac{1}{4}\tilde{\mathcal{G}} S_2^{-\frac{1}{2}} F = \frac{1}{4}\tilde{\mathcal{G}}\mathcal{F} S_2^{-1}. \tag{15.268}$$

In addition

$$\frac{\partial \hat{L}}{\partial S_1} = \frac{1}{4} F\tilde{F}. \tag{15.269}$$

Using (15.265), (15.266), (15.268) and (15.269), together with (15.251), we see that (15.256) is satisfied. It is easy to check that the scale invariant combinations $\mathcal{F}$ and $\mathcal{G}$, given by (15.263) and (15.258) have the very simple transformation law

$$\delta\mathcal{F} = S_2 C\mathcal{G}, \quad \delta\mathcal{G} = -S_2 C\mathcal{F}, \tag{15.270}$$

i.e., they transform according to the $U(1) \cong SO(2)$ compact subgroup just as F and G in (15.203), but with the parameter λ replaced by $S_2 C$. If $L(\mathcal{F})$ is the Born-Infeld Lagrangian, the theory with scalar fields given by $\hat{L}$ in (15.259) can also be reformulated à la Schrödinger. >From (15.262) and (15.263) solve for $\mathcal{F}$ and $\mathcal{G}$ in terms of F, G, S_1 and S_2. Then $\mathcal{M} = \mathcal{F} - i\mathcal{G}$ must satisfy the same equation (15.215) that M does when no scalar fields are present.

15.2.6 Duality as a Legendre transformation

We have observed that, even in the general case of $Sp(2n, R)$, although the Lagrangian is not invariant, the combination [see (15.192)]

$$\hat{L} - \frac{1}{4} F\tilde{G} \tag{15.271}$$

is invariant. Here we restrict ourselves to the case of $SL(2,R)$, one tensor $F_{\mu\nu}$ and one complex scalar field $S = S_1 + iS_2$. As in Section 5, we use the notation $\hat{L}$ to denote the part of the

Lagrangian that depends on the scalar fields, as well as on $F_{\mu\nu}$, but not on scalar derivatives. Then

$$\hat{L}(S_1, S_2, F) - \frac{1}{4}F\tilde{G} = \hat{L}(S_1', S_2', F') - \frac{1}{4}F'\tilde{G}', \tag{15.272}$$

where

$$\begin{pmatrix} G' \\ F' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} G \\ F \end{pmatrix}, \quad S' = \frac{aS+b}{cS+d}, \quad ad - bd = 1, \tag{15.273}$$

$$\tilde{G} = 2\frac{\partial \hat{L}}{\partial F}. \tag{15.274}$$

There are several interesting special cases of this invariance statement. The first corresponds to $a = d = 1,\ c = 0,\ b$ arbitrary, which gives

$$G' = G + bF, \quad F' = F, \quad S_1' = S_1 + b, \quad S_2' = S_2. \tag{15.275}$$

The second corresponds to $b = c = 0,\ d = 1/a,\ a$ arbitrary, which gives

$$G' = aG, \quad F' = \frac{1}{a}F, \quad S' = a^2 S, \quad S_1' = a^2 S_1, \quad S_2' = a^2 S_2. \tag{15.276}$$

The third corresponds to $a = d = 0,\ b = -1/c,\ c$ arbitrary, which gives

$$G' = -\frac{1}{c}F, \quad F' = cG, \quad S' = -\frac{1}{c^2 S}, \quad S_1' = -\frac{S_1}{c^2|S|^2}, \quad S_2' = \frac{S_2}{c^2|S|^2}. \tag{15.277}$$

Using (15.275) in (15.272) we find

$$\hat{L}(S_1, S_2, F) - \frac{1}{4}F\tilde{G} = \hat{L}(S_1 + b, S_2, F) - \frac{1}{4}F\left(\tilde{G} + b\tilde{F}\right). \tag{15.278}$$

Taking $b = -S_1$, we obtain

$$\hat{L}(S_1, S_2, F) = \hat{L}(0, S_2, F) + \frac{1}{4}S_1 F\tilde{F}, \tag{15.279}$$

which gives the dependence of $\hat{L}$ on S_1, in agreement with (15.259). This choice for the constant b is allowed because this part of the Lagrangian, which does not include the kinetic term for the scalar fields, does not contain derivatives of the scalar fields. Using (15.276) in (15.272) we find

$$\hat{L}(S_1, S_2, F) - \frac{1}{4}F\tilde{G} = \hat{L}\left(a^2 S_1, a^2 S_2, \frac{1}{a}F\right) - \frac{1}{4}F\tilde{G}, \tag{15.280}$$

i.e.,

$$\hat{L}(S_1, S_2, F) = \hat{L}\left(a^2 S_1, a^2 S_2, \frac{1}{a}F\right). \tag{15.281}$$

Setting $S_2 = 0$ in this equation, we see that $\hat{L}(S_1, 0, F)$ is a function of $S_1^{\frac{1}{2}}F$, in agreement with the more precise statement (15.279). Setting instead $S_1 = 0$, we find that $\hat{L}(0, S_2, F)$ is a function of $S_2^{\frac{1}{2}}F$, in agreement with (15.259).

Using (15.277) in (15.272) we find

$$\hat{L}(S_1, S_2, F) - \frac{1}{4}F\tilde{G} = \hat{L}\left(-\frac{S_1}{c^2|S|^2}, \frac{S_2}{c^2|S|^2}, cG\right) + \frac{1}{4}G\tilde{F}, \tag{15.282}$$

i.e.,

$$\hat{L}\left(-\frac{S_1}{c^2|S|^2}, \frac{S_2}{c^2|S|^2}, cG\right) = \hat{L}(S_1, S_2, F) - \frac{1}{2}F\tilde{G}, \tag{15.283}$$

or

$$\hat{L}\left(-\frac{1}{c^2 S}, cG\right) = \hat{L}(S,F) - \frac{1}{2}F\tilde{G}. \tag{15.284}$$

We have shown that the Ansatz (15.259) of Section 5 is a natural consequence of the invariance of $\hat{L} - \frac{1}{4}F\tilde{G}$. Equation (15.284) with (15.274) can be interpreted as a Legendre transformation. Given a Lagrangian $\hat{L}(S,F)$, *define* the dual Lagrangian $\hat{L}_D(S,F_D)$, a function of the dual field F_D, by

$$\hat{L}_D(S,F_D) + \hat{L}(S,F) = \frac{1}{2}FF_D, \tag{15.285}$$

$$F_D = 2\frac{\partial \hat{L}}{\partial F}, \quad F = 2\frac{\partial \hat{L}_D}{\partial F_D}. \tag{15.286}$$

With these definitions, the dual of the dual of a function equals the original function.* In general, the dual Lagrangian is a very different function from the original Lagrangian. For a self-dual theory, if we set

$$F_D = \tilde{G}, \quad \tilde{F}_D = -G, \tag{15.287}$$

we see from (15.284) that

$$-\hat{L}\left(-\frac{1}{c^2 S}, cG\right) = \hat{L}_D(S,\tilde{G}), \tag{15.288}$$

which must be independent of c, since G is.

The above argument can be inverted. Let the Legendre transformation (15.285) produce a dual Lagrangian given by (15.288) with $c = 1$, or

$$\hat{L}_D(S,F_D) = -\hat{L}\left(-\frac{1}{S}, -\tilde{F}_D\right) = -\hat{L}\left(-\frac{1}{S}, G\right). \tag{15.289}$$

It then follows that $\hat{L} - \frac{1}{4}F\tilde{G}$ is invariant under (15.277) with $c = 1$, *i.e.*,

$$G' = -F, \quad F' = G, \quad S' = -\frac{1}{S}. \tag{15.290}$$

If we now assume that it is also invariant under (15.275) with arbitrary b, it follows that it is invariant under the entire group $SL(2,R)$. Indeed, if we call t_b the transformation (15.275) and s the transformation (15.290), the product $t_b s t_{b'} s t_{b''}$ gives the most general transformation of $SL(2,R)$.

If we normalize the scalar field differently, taking *e.g.*, instead of S,

$$\tau = cS, \quad L'(\tau,F) = \hat{L}(S,F), \tag{15.291}$$

$$L'_D(\tau,F'_D) + L'(\tau,F) = \frac{1}{2c}FF'_D, \tag{15.292}$$

and write the Legendre transformation as

$$2\frac{\partial L'(\tau,F)}{\partial F} = \frac{1}{c}F'_D, \quad 2\frac{\partial L'_D(\tau,F'_D)}{\partial F'_D} = \frac{1}{c}F, \tag{15.293}$$

we see that

$$F'_D = cF_D = c\tilde{G}, \tag{15.294}$$

*The unconventional factor 1/2 on the right hand side of (15.285) is introduced to avoid overcounting when summing over the indices of the antisymmetric tensors F and F_D.

and

$$L'_D(\tau, F'_D) = \hat{L}_D(S, F_D) = -\hat{L}\left(-\frac{1}{c^2 S}, -c\tilde{F}_D\right)$$
$$= -L'\left(-\frac{1}{cS}, -c\tilde{F}_D\right) = -L'\left(-\frac{1}{\tau}, -\tilde{F}'_D\right), \tag{15.295}$$

for a self-dual theory.

A standard normalization [12, 13] is $c = 4\pi$, in which case the expectation value of the field τ is

$$< \tau >= \frac{\theta}{2\pi} + i\frac{4\pi}{g^2}. \tag{15.296}$$

In the presence of magnetically charged particles and dyons (both electrically and magnetically charged) the invariance of the charge lattice restricts [14] the $SL(2,R)$ group to the $SL(2,Z)$ subgroup generated by

$$\tau \to -\frac{1}{\tau}, \quad \tau \to \tau + 1. \tag{15.297}$$

At the quantum level the Legendre transformation corresponds to the integration over the field F in the functional integral, after adding to the Lagrangian $\hat{L}$ a term $-\frac{1}{2}FF_D$.

15.2.7 Concluding remarks

Nonlinear electromagnetic Lagrangians, like the Born-Infeld Lagrangian, can be supersymmetrized [15, 16] by means of the four-dimensional $N = 1$ superfield formalism, and this can be done even in the presence of supergravity. When the Lagrangian is self-dual, it is natural to ask whether its supersymmetric extension possesses a self-duality property that can be formulated in a supersymmetric way. We were not able to do this in the nonlinear case. When the Lagrangian is quadratic in the fields $F^a_{\mu\nu}$, the problem as been solved in [17], where the combined requirements of supersymmetry and self-duality were used to constrain the form or the weak coupling ($S_2 \to \infty$) limit of the effective Lagrangian from string theory, in which one neglects the nonabelian nature of the gauge fields.

The $SL(2,Z)$ subgroup of $SL(2,R)$ that is generated by the elements $4\pi S \to -1/4\pi S$ and $S \to S + 1/4\pi$ relates different string theories [18] to one another.

The generalization of [2] to two dimensional theories [19] has been used to derive the Kähler potential for moduli and matter fields in effective field theories from superstrings. In this case the scalars are valued on a coset space $\mathcal{K}/\mathcal{H}$, $\mathcal{K} \in SO(n,n)$, $\mathcal{H} \in SO(n) \times SO(n)$. The kinetic energy is invariant under $\mathcal{K}$, and the full classical theory is invariant under a subgroup of $\mathcal{K}$. String loop corrections reduces the invariance to a discrete subgroup that contains the $SL(2,Z)$ group generated by $T \to 1/T$, $T \to T - i$, where ReT is the squared radius of compactification in string units.

15.3 Exact electromagnetic duality by David I Olive

15.3.1 Introduction

Electromagnetic duality is a very old idea, possibly predating Maxwell's equations. Although the route that has recently led to a precise and convincing formulation has been long, it has turned out to be of quite surprising interest. This is because it has synthesised many hitherto independent lines of thought, and so intriguingly interelated disparate ideas arising in the quest for a unified theory of particle physics valid in the natural space-time with three space and one time dimension. Despite the progress, final proof is lacking and likely to require further breakthroughs in fundamental mathematics.

Although nature does not seem to display exact electromagnetic duality, realistic theories could well be judiciously broken versions of the exact theory in which sufficient structure survives to explain such long-standing puzzles as quark confinement in the way advocated by Seiberg and Witten [1]. Spectacular support for their arguments comes from applications in pure mathematics where new insight has been gained into the classification of four-manifolds [2], transcending the celebrated work of Donaldson [3].

Here I shall review the developments leading up to the formulation of exact electromagnetic duality, taking the view that an understanding of this must precede that of the symmetry breaking.

The Original Idea

The apparent similarity between the electric and magnetic fields $\underline{E}$ and $\underline{B}$ was confirmed and made more precise by Maxwell's discovery of his equations. In vacuo, they can be written concisely as just two equations [4]:

$$\nabla.(\underline{E}+i\underline{B})=0, \qquad (1a)$$

$$\nabla\wedge(\underline{E}+i\underline{B})=i\frac{\partial}{\partial t}(\underline{E}+i\underline{B}) \qquad (1b)$$

at the expense of introducing a complex vector field $\underline{E}+i\underline{B}$. These equations display several symmetries whose physical importance became clear subsequently. They display Poincaré (rather than Galilean) symmetry, and, beyond that, conformal symmetry (with respect to space-time transformations preserving angles and not just lengths). Unlike the Poincaré symmetry, the conformal symmetry is specific to four space-time dimensions. Even more sensitive to the precise space-time metric is the electromagnetic duality rotation symmetry of Maxwell's equations

$$\underline{E}+i\underline{B}\rightarrow e^{i\phi}(\underline{E}+i\underline{B}) \qquad (2)$$

since only in $3+1$ dimensions do the electric and magnetic fields both constitute vectors so that the complex linear combination $\underline{E}+i\underline{B}$ appearing in (1) and (2) can be formed. It is the extension of the fascinating symmetry (2) of (1) that is the main theme of what follows.

Notice that we can form two real, quadratic expressions invariant with respect to (2) [4]:

$$\frac{1}{2}|\underline{E}+i\underline{B}|^2=\frac{1}{2}(E^2+B^2),$$

$$\frac{1}{2i}(\underline{E}+i\underline{B})^*\wedge(\underline{E}+i\underline{B})=\underline{E}\wedge\underline{B},$$

respectively the energy and momentum densities of the electromagnetic field.

On the other hand, $\frac{1}{2}(\underline{E}+i\underline{B})^2$ is complex with real and imaginary parts given by

$$\frac{1}{2}(E^2-B^2)+i\underline{E}.\underline{B}.$$

As the real part is the Lagrangian density, this shows that it forms a doublet under (2) when combined with $\underline{E}.\underline{B}$ which is a total derivative. Thus the Maxwell action forms a doublet with a "topological quantity" which is proportional to the instanton number in non-abelian theories.

We would like to generalise the electromagnetic duality rotation symmetry (2) to include matter. We could also consider generalisations to non-abelian gauge theories of the type which seem to unify the fundamental interactions. In either case we meet the same difficulty that the gauge potentials enter the equations of motion and that we do not know how to extend the transformation (2) to include them. Eventually we shall find a way of combining the two generalisations, thereby extending the symmetry.

There is another, familiar, difficulty with the equations of motion in non-abelian gauge theories; that they are conformally invariant (in $3 + 1$ dimensions). As a consequence, the gauge particles which are the quanta of the gauge potentials, should be massless. This is fine for the photon, but not for any other gauge particles. This problem seems to be rather general and deep: unified theories chosen according to geometric principles tend to exhibit unwelcome conformal symmetry. This occurs in string theory too, at least as a world sheet symmetry. As a consequence, there is a general problem of understanding the origin of mass through a geometrical mechanism for breaking conformal symmetry. We know of only two possible solutions, at first sight different, but related in what follows.

The first is the idea that mass arises from the vacuum spontaneously breaking some of the gauge symmetry via a "Higgs" scalar field [5,6,7]. The second is a principle due to Zamolodchikov [8] that we now discuss.

15.3.2 Zamolodchikov's Principle and Solitons

The second insight into the origin of mass comes from another area of physics. Yet, as we shall see, it seems connected to the first, in four dimensions at least. The conformal symmetry on the world sheet needed for the internal consistency of string theory hinders the emergence of a physically realistic mass spectrum in an otherwise unified theory. However, when string theory is abstracted to the study of "conformal field theory", and applied to the study of second order phase transitions of two dimensional materials, it is seen that the simple application of heat breaks the conformal symmetry controlling the behaviour at the critical temperature. In specific models it was learnt by Onsager, Baxter and their followers [9] that it is possible to supply heat while maintaining integrability (or solvability). Zamolodchikov [8] has elevated this observation to a principle which has rationalised the theory of solitons and more. In two dimensions, the local conservation laws characteristic of conformal symmetry (or augmented versions such as W-symmetry) are chiral. This means that the densities are either left moving or right moving (at the speed of light) and so can be added, multiplied or differentiated. If conformal symmetry is judiciously broken, a certain (infinite) subset of the chiral densities remain conserved, although no longer chirally so. Their conserved charges, that is their space integrals, generate an infinite dimensional extension of the Poincaré algebra in which the charges carry integer spins. The charges with spin plus or minus 1 are the conventional light cone components of momentum. The sinh-Gordon equation illustrates this nicely. It can be written

$$\frac{\partial^2\phi}{\partial t^2} - \frac{\partial^2\phi}{\partial x^2} + \frac{\mu^2}{2\beta}\left(e^{\beta\phi} - e^{-\beta\phi}\right) = 0. \tag{3}$$

The last term, proportional to $e^{-\beta\phi}$, can be multiplied by a variable coefficient η so that $\eta = 1$ yields (3), while $\eta = 0$ yields the Liouville equation. Liouville exploited the conformal symmetry of his equation in order to solve it completely, long ago.

It is interesting to investigate the behaviour as η varies from 1 to 0. As long as $\eta > 0$, a simple redefinition of the field ϕ by a displacement restores the sinh-Gordon form (3) but with μ replaced by $\mu\eta^{1/4}$. As $\mu\lambda$ is the mass of the particle which is the quantum excitation of ϕ, we see that it is singular as η approaches zero with critical exponent 1/4. So we see how mass arises from the breaking of conformal symmetry. This short discussion was classical but it extends to the quantum regime as envisaged by Zamolodchikov [8].

The sine-Gordon equation is obtained from the sinh-Gordon equation (3) by replacing β by $i\beta$. It then exhibits the symmetry

$$\phi \to \phi + \frac{2\pi}{\beta} \tag{4}$$

and, consequently, possesses an infinite number of vacuum solutions $\phi_n = \frac{2\pi n}{\beta}$, $n \in \mathbb{Z}$, all with the same minimum energy, zero. The particle of mass $\mu\lambda$ describes fluctuations about any of these vacua. But there also exist classical solutions which interpolate two successive vacua and which are stable with respect to fluctuations. These solutions can be motionless, describing a new particle, the soliton, at rest, or can be boosted to any velocity less than that of light. The jump in n, equal to ± 1, can be regarded as a topological quantum number, indicating either a soliton or an antisoliton. What is particularly remarkable is that one can consider a solution with an arbitrary number of solitons and/or antisolitons, initially well separated, but approaching each other, then colliding and finally emerging with velocities unchanged and energy profiles generally unscathed except for time advances relative to uninterrupted trajectories [10]. Thus the solitons persist in their structure despite their collisions and can legitimately be regarded as providing classical models of a particle with a finite mass and a structure of finite extent. This phenomenon is a very special feature of sine-Gordon theory that can be ascribed to the infinite number of conservation laws mentioned previously, themselves relics of conformal symmetry.

This sort of integrable field theory has two “sorts" of particle, the quanta of the fluctuation of the field ϕ (obtained by second quantisation) and the solitons which are classical solutions. Skyrme [11] was the first to ask whether these two “sorts" of particle are intrinsically different and found that they were not. His explanation was that, in the full quantum theory, it is possible to construct a new quantum field whose fluctuations are the solitons. The new field operator is obtained by an exponential expression in the original field ϕ

$$\psi_{\pm}(x) = e^{i\beta(\phi \pm \int_{-\infty}^{x} dx' \frac{\partial \phi}{\partial t})} \tag{5}$$

with two spin components (and a normal ordering understood). Coleman and Mandelstam [12,13] later confirmed that ψ satisfied the equations of motion of the massive Thirring model.

The construction (5) is an example of the vertex operator construction later to be so important in string theory and in the representation theory of infinite dimensional algebras (resembling quantum field theories).

There are multicomponent generalisations of the sine-Gordon equations called the affine Toda theories, likewise illustrating Zamolodchikov's principle in a nice way, and revealing the role of algebraic structures such as affine Kac Moody algebras [14]. Again there are particles created by each field component, now possessing interesting mass and coupling patterns (related to group theory). Remarkably, there are an equal number of soliton species and these display very similar properties [15].

15.3.3 Magnetic Charge and its Quantisation

Let us now return to the question of extending the electromagnetic duality rotation symmetry (2) to matter carrying electric and magnetic charges. Suppose first that matter can be regarded as being composed of classical point particles carrying typical electric and magnetic charges q and g, say. Then it is easy to include the source charges on the right hand side of Maxwell's equations (1) and to supplement (1) by the equations of motion for the individual particles subject to a generalised Lorentz force. This system maintains the symmetry (2) if, in addition,

$$q + ig \to e^{i\phi}(q + ig). \tag{6}$$

The price to be paid for achieving this is the inclusion of unobserved magnetic charge. We must therefore suppose that the failure to observe magnetic charge is either due to an associated very large mass or some other reason.

Turning from the classical to the quantum theory, we immediately find a difficulty, namely that the electromagnetic couplings of the matter wave functions require the introduction of gauge potentials, a procedure which is not straightforward in the presence of magnetic charge.

Nevertheless, in 1931, Dirac overcame this difficulty and showed that the introduction of magnetic charge could be consistent with the quantum theory, provided its allowed values were constrained [16]. His result was that a magnetic charge g_1, carrying no electric charge, could occur in the presence of an electric charge q_2, like the electron carrying no magnetic charge, provided

$$q_2 g_1 = 2\pi n\lambda \qquad\qquad n = 0, \pm 1, \pm 2, \dots \tag{7}$$

As he pointed out, this condition had a stunning consequence: provided g_1 exists somewhere in the universe, even though unobserved, then any electric charge *must* occur in integer multiples of the unit $\frac{2\pi\lambda}{g_1}$, by (7). This quantisation of electric charge is indeed a feature of nature and this explanation is actually the best yet found. Although apparent alternative explanations, evading the necessity for magnetic charge, have appeared, they turn out to be unexpectedly equivalent to the above argument, as we shall see.

There is a problem with the Dirac condition (7), namely that it does not respect the symmetry (6). In fact (7) is not quite right because, although Dirac's argument is impeccable, there is an implicit assumption hidden within the situation considered. It took a surprisingly long time to rectify this and hence restore the symmetry (2) and (6), as we see later.

Given that this difficulty is overcome, we can seek a consistent quantum field theory with both electric and magnetic charges. Then, presumably, the particles carrying magnetic charge would have a structure determined by the theory, and hence a mass dependent on the charges carried. Just as the Maxwell energy density respected the symmetry (2), we would expect this mass formula to respect (6) so that

$$M(q, g) = M(|q + ig|) = M(\sqrt{q^2 + g^2}). \tag{8}$$

We now proceed to find such a theory.

15.3.4 Magnetic Monopoles (and dyons) as Solitons

We can now draw together several clues in the ideas already discussed. One concerns the quantisation of electric charge: since the electric charge operator, Q, is the generator of the $U(1)$ gauge group of Maxwell theory, its quantisation could be explained by supposing that it is actually a generator of a larger, simple gauge group (that could unify it with other interactions). If the larger group were $SU(2)$, for example, Q would then be a generator of an $SU(2)$ Lie algebra, that is, an internal angular momentum algebra. Consequently its eigenvalues would be quantised, thereby providing an alternative explanation of electric charge quantisation which apparently evades the need for magnetic charge.

However, we still have to furnish a mechanism selecting the direction of Q amongst the three $SU(2)$ directions. This can be achieved by a scalar field with three components $(\phi_1(x), \phi_2(x), \phi_3(x))$, like the $SU(2)$ gauge fields. This scalar field has to have the unusual feature of not vanishing in the vacuum, so that it can select the Q direction there. It is therefore a "Higgs" field providing the mechanism whereby the vacuum spontaneously breaks the $SU(2)$ gauge symmetry down to the $U(1)$ subgroup [5,6,7]. As well as this, it also breaks conformal symmetry, introducing mass for two of the gauge particles, leaving the photon massless.

There is a simple formula for the resultant masses of the gauge particles

$$M(q, 0) = a|q|, \tag{9}$$

where q is the eigenvalue of Q specifying the electric charge of a specific mass eigenstate. a constitutes a new fundamental parameter specifying the magnitude of the vacuum expectation value of the scalar Higgs field. Actually the mass formula (9) is much more general. Instead of the gauge group being $SU(2)$, it could have been any simple Lie group, G, say, and (9) holds as long as the Higgs field lies in the adjoint representation of G, like the gauge fields.

In the vacuum, the gauge group G is spontaneously broken to a subgroup

$$U(1)_Q \times K/Z, \tag{10}$$

where Q still generates the invariant $U(1)$ subgroup commuting with K. The denominator Z indicates a finite cyclic group in which the $U(1)_Q$ intersects K, and will not be important for what we have to say.

But this setup, a spontaneously broken gauge theory with a Higgs in the adjoint representation, is very much the analogue in four dimensions of the sine-Gordon theory in two dimensions. Instead of the symmetry relating the degenerate vacua being discrete, (4), it is now continuous, being the gauge symmetry, G, and it is again possible to trap nontrivial topologically stable field configurations of finite energy. Indeed in 1974, 't Hooft and Polyakov [17,18] found a classical soliton solution emitting a $U(1)$ magnetic flux with strength $4\pi\lambda/q$ in the $SU(2)$ theory with heavy gauge particles carrying electric charges $\pm q$. Thus there is a soliton which is a magnetic monopole whose charge indeed satisfies the Dirac condition (7). Thus the desired novelty of this explanation of electric charge quantisation is illusory as it reduces to Dirac's original argument [16]. For a more detailed review of the material in this section see [19].

However what we have done is inadvertently achieve our other aim, that of constructing a theory in which the magnetic monopoles have structure and a definite mass which can be calculated by feeding the field configuration into the energy density and integrating over space. The result is the following inequality, known as the "Bogomolny" bound [20],

$$M(0, g) \geq a|g|. \tag{11}$$

The similarity to the Higgs formula (9) prompts the question as to whether the inequality in (11) can be saturated to give equality. This is possible in the "Prasad-Sommerfield" limit in which the self interactions of the Higgs field vanish [21]. Then the lower bound in (11) is achieved if the fields satisfy certain first order differential equations, known as the "Bogomolny equations" [19]

$$\mathcal{E}_i = 0, \qquad \mathcal{D}_0\phi = 0, \qquad \mathcal{B}_i = \pm\mathcal{D}_i\phi, \tag{12}$$

where $\mathcal{E}_i$ and $\mathcal{B}_i$ denote the nonabelian electric and magnetic fields. Solutions to (12) have zero space momentum and therefore describe a magnetic monopole at rest, with mass $a|g|$ (if $|g|$ has its minimum least positive value).

The sine-Gordon solitons satisfy similar first order differential equations that imply that the mass can also be expressed as a surface term, but there is an important difference. This is that the Bogomolny equations (12) (unlike the first order sine-Gordon equations) can also be solved for higher values of the topological charge, here magnetic charge. When the magnetic charge is m times its least positive value, the space of solutions to (12), called the moduli space, form a manifold of $4m$ dimensions. $3m$ of these dimensions can be interpreted as referring to the space coordinates of m individual magnetic monopoles of like charge.

This means that m like magnetic monopoles can exist in arbitrary configurations of static equilibrium (unlike m sine-Gordon solitons, which must move). So, as they have no inclination to move relatively, like magnetic monopoles at rest must fail to exert forces on each other [22]. (This is reminiscent of the multi-instanton solutions to self-dual gauge theories: indeed the Bogomolny equations (12) can be interpreted as self-dual equations in four Euclidean dimensions).

The remaining m coordinates, one for each monopole, have a more subtle, but nevertheless, important interpretation: they correspond to degrees of freedom conjugate to the electric charge of each monopole. Because of this, it is possible for each magnetic monopole soliton to carry an electric charge, q, say [23]. In this case, they are called "dyons", following the terminology introduced by Schwinger [24]. Then the mass of an individual dyon is given by, [25],

$$M(q, g) = a|q + ig| = a\sqrt{q^2 + g^2}. \tag{13}$$

The first remarkable fact about this formula is that it is universal. It applies equally to the dyon solitons of the theory and to the gauge particles, as it includes the Higgs formula (9). In fact it applies to all the particles of the theory created by the fundamental quantum fields, as it also includes the photon and Higgs particles which are both chargeless and massless. Thus, whatever G, (13) unifies the Higgs and Bogomolny formulae and is therefore democratic in the sense that it does not discriminate as to whether the particle considered arises as a classical soliton or as a quantised field fluctuation [26].

Secondly the mass formula (13) does indeed respect the electromagnetic duality rotation symmetry (6) as it has the structure (8).

15.3.5 Electromagnetic Duality Conjectures

We have seen that spontaneously broken gauge theories with adjoint Higgs (in the Prasad-Sommerfield limit) have remarkable properties, at least according to the naive arguments just outlined. Magnetically neutral particles occur as quantum excitations of the fields present in the action, whereas magnetically charged particles occur as solitons, that is, solutions to the classical equations of motion. Yet, despite this difference in description, all particles enjoy a universal mass formula (13).

Skyrme showed that, in two dimensions, the soliton of sine-Gordon theory could be considered as being created by a new quantum field obeying the equations of motion of the massive Thirring model [11,12,13]. Thus the same quantum field theory can be described by two distinct actions, related by the vertex operator transformation (5). It is natural to ask if something similar can happen in four dimensions, with the theory under consideration. There, the solitons carry magnetic charge with an associated Coulomb magnetic field. This suggests that the hypothetical quantum field operator, creating the monopole solitons, should couple to a "magnetic" gauge group with strength inversely related to the original "electric" gauge coupling because of Dirac's quantisation condition,

$$q_0 \to g_0 = \pm\frac{4\pi\lambda}{q_0}, \tag{14}$$

or possibly half this.

Thinking along these lines, two more specific conjectures were proposed in 1977. First, considering a more general theory, with a simple exact gauge symmetry group H, (i.e. not of the form (10)), Goddard, Nuyts and Olive established a non-abelian version of the Dirac quantisation condition (7) and used it to propose the conjecture that the magnetic, or dual group H^v could be constructed in two steps as follows [27].

(i) The Lie algebra of H^v is specified by saying that its roots α^v are the coroots of H:-

$$\alpha \to \alpha^v = \frac{2\alpha}{\alpha^2}. \tag{15a}$$

(ii) The global structure of the group H^v is specified by constructing its centre $Z(H^v)$ from that of H, $Z(H)$:-

$$Z(H) \to Z(H^v) = \frac{Z(\tilde{H})}{Z(H)}, \tag{15b}$$

where $\tilde{H}$ is the universal covering group of H, that is, the unique simply connected Lie group with the same Lie algebra as H.

This conjecture remains open. Notice the similarity between (15) and (14). In order to make progress, Montonen and Olive sought a more specific proposal in a simpler context, and considered spontaneously broken gauge theories of the type discussed above, but with the gauge group henceforth definitely chosen to be $SU(2)$ [26]. This is broken to $U(1)$ by a triplet Higgs field so that the mass formula (13) holds good.

The possible quantum states of the theory carry values of q and g which form an integer lattice when plotted in the complex $q + ig$ plane (with Cartesian coordinates (q,g)). Ignoring possible dyons, the single particle states correspond to five points of this lattice. The photon and the Higgs particle correspond to the origin (0,0), the heavy gauge particles $W^{\pm}$ to the points $(\pm q_0, 0)$. Thus the particles created by the fundamental fields in the original, "electric" action lie on the real, electric axis. The magnetic monopole solitons $M^{\pm}$ lie on the imaginary, magnetic axis at $(0, \pm g_0)$, while the dyons could lie on the horizontal lines through these two points. Since, at this stage, it is unclear what values of their electric charges are allowed, they will temporarily be omitted, to be restored later.

Now, if we follow the transformation (14) by a rotation through a right angle in the $q + ig$ plane, the five points just described are rearranged. This suggests that the "dual" or magnetic formulation of the theory with $M^{\pm}$ created by fields present in the action will also be a similar spontaneously broken gauge theory, but with the coupling constant altered by (14). In this new formulation it is the $W^{\pm}$ particles that would occur as solitons.

This is the Montonen-Olive electromagnetic duality conjecture in its original form [26]. In principle, it could be proven by finding the analogue of Skyrme's vertex operator construction (5) [11], but, even with present knowledge, this seems impossibly difficult. Notice that the sine-Gordon quantum field theory was described by two quite dissimilar actions whereas in the four dimensional theories the two hypothetical actions have a similar structure but refer to electric and magnetic formulations.

The magnitudes of physical quantities should agree whichever of the two actions is chosen as a starting point for their calculation. The conjecture will immediately pass at least two simple tests of this kind, showing that it is not obviously inconsistent. The first test concerns the mass formula (13) and is passed precisely because of the universal property that has already been emphasised.

A second test concerns the fact that, according to the existence of static solutions to the Bogomolny equations with magnetic charge $2g_0$ discussed earlier, an M^+M^+ pair exert no static forces on each other. This result is according to the electric formulation of the theory and ought to be confirmed in the magnetic formulation. This is equivalent to checking that there is no W^+W^+ force in the electric formulation. In the Born approximation, two Feynman diagrams contribute, photon and Higgs exchange. Using Feynman rules, one finds that photon exchange yields the expected Coulomb repulsion but that the second contribution precisely cancels the first. This can happen because the Higgs is massless in the Prasad-Sommerfield limit [26].

Thus, at the level considered, the conjecture is consistent, but there are more searching questions to be asked. Their answers will lead to a reformulation of the conjecture that passes even more stringent tests.

15.3.6 Catechism concerning the Duality Conjecture

The Montonen Olive electromagnetic duality conjecture immediately provokes the following questions:-
(1) How can the magnetic monopole solitons possess the unit spin necessary for heavy gauge particles?
(2) Will not the quantum corrections to the universal mass formula (13) vitiate it?
(3) Surely the dyon states, properly included, will spoil the picture just described?

Clearly the answers to the first two questions will depend on the choice of quantisation procedure, and presumably the most favourable one should be selected. The idea of what the appropriate choice was, and how it answered the first two questions came almost immediately, though understanding has continued to improve until the present. The answer to the third question remained a mystery until it was decisively answered by Sen in 1994, as we shall describe [28].

The immediate response of D'Adda, Di Vecchia and Horsley was the proposal that the quantisation procedure be supersymmetric [29]. The point is that the theory we have described is begging to be made supersymmetric since this can be achieved without spoiling any of the features we have described. For example, since the scalar and gauge fields lie in the same, adjoint representation of the gauge group they can lie in the same supermultiplet. The vanishing of the Higgs self-interaction implied by the Prasad-Sommerfield limit is then a consequence of supersymmetry. Because the helicity change between scalar and vector is one unit, the supersymmetry is presumably of the "extended" kind, with either $N = 2$ or $N = 4$ possible. Osborn was the first to advocate the second possibility [30].

The reason supersymmetry helps answer the first question is that, given that it holds in the full quantum theory, it must be represented on any set of single particle states carrying the same specific values of the charges and of energy and momentum. This is so, regardless of the nature of the particles, whether they are created by quantum fields or arise as soliton states, that is, whether or not they carry magnetic charge. When the extended supersymmetry algebra with N supercharges acts on massive states, the algebra is isomorphic to a Clifford algebra in a Euclidean space with $4N$ dimensions [31]. This algebra has a unique irreducible representation of 2^{2N} dimensions. This representation includes states whose helicity h varies over a range $\Delta h = N$ with intervals of $1/2$. The limits of this range may differ but should not exceed 1 in magnitude if the states can be created by fields satisfying renormalisable equations of motion, according to the standard wisdom.

So, of necessity, the monopoles carry spin, quite likely unit spin. Secondly, quantum corrections tend to cancel in supersymmetric theories, essentially because the supersymmetric harmonic oscillator has no zero point energy. This is relevant to the second question because the small fluctuations about the soliton profile decompose into such oscillators, with the mass correction equal to the sum of zero point energies.

The structure of the representation theory raises some questions. According to the renormalisability criterion, the maximum range of helicity is $\Delta h = 1 - (-1) = 2$, which is just consistent with $N = 2$ supersymmetry, but apparently forbids $N = 4$ supersymmetry. Another difficulty concerns the understanding of how the Higgs mechanism providing the mass of the gauge particles, works in the presence of supersymmetry. The point is that the expression on the right hand side of the supercharge anticommutator, $\gamma.P$, is a singular matrix when $P^2 = 0$, that is, for massless states. As a consequence, the supersymmetry algebra acting on massless states is isomorphic to an Euclidean algebra in $2N$ dimensions and so now possesses a unique irreducible representation of 2^N dimensions (the square root of the number in the massive case), with helicity range $\Delta h = N/2$. This now accommodates both $N = 2$ and $N = 4$ superalgebras but no more. Indeed it is the reason we cannot envisage a more extended supersymmetry, such as $N = 8$, which requires gravitons of spin 2, whose interactions are not renormalisable.

The question arises of how to reconcile the jump in the dimensions of the representations with the acquisition of mass for a given field content of scalar and gauge fields. The answer, due to Witten and Olive [32], is that something special happens precisely when the Higgs field lies in the adjoint representation, as we have assumed, and so can lie in the same supermultiplet as the gauge field. Then the electric charge, q, occurs as a central charge, providing an additional term on the right hand side of the supercharge anticommutator, thereby altering the structure of the algebra. The condition for a "short" representation, that is, of dimension 2^N, is now $P^2 = a^2q^2$, rather than $P^2 = 0$. Thus, providing the Higgs formula (9) holds, mass can be acquired without altering the dimension of the irreducible representation. Furthermore, magnetic charge can occur as yet another additional central charge with the condition for a "short" representation being simply the universal Bogomolny-Higgs formula (13). In particular, this means that this formula now has an exact quantum status as it follows from the supersymmetry algebra which is presumably an exact, quantum statement (though there may be subtle renormalisation effects) [32].

15.3.7 More on Supersymmetry and $N = 2$ versus $N = 4$

The possibility that, unlike the unextended supersymmetry algebra, the extended ones could be modified by the inclusion of central charges was originally noted by Haag, Lopuszanski and Sohnius [33], while the physical identification of these charges was due to Witten and Olive [32]. The confirmation of the result involved a new matter of principle. Hitherto supersymmetry algebras had been checked via the algebra of transformations of the fields entering the action. But since these will never carry magnetic charge in the electric formulation, this method will not detect the presence of magnetic charge in the algebra. Instead, it is necessary to manipulate all the charges explicitly, treating them as space integrals of local polynomials in the fields and their derivatives.

The supersymmetry algebras possess an automorphism (possibly outer) involving chiral transformations of the supercharges

$$Q^{\alpha}_{L,R} \to e^{\pm i\phi/2} Q^{\alpha}_{L,R} \qquad \alpha = 1, 2 \ldots N \tag{16}$$

where the suffices refer to the handedness. When the central charges are present, this automorphism requires them to simultaneously transform by (6) (at least in the $N = 2$ case). Thus the electromagnetic duality rotation is now seen to relate to a chiral rotation of the supercharges, sometimes known as R-symmetry.

So far, the theory could possess either $N = 2$ or $N = 4$ supersymmetry and it is necessary to determine which, if possible. At first sight, $N = 2$ is simpler as there are precisely two central charges, q and g, as we have described, whereas in $N = 4$ there are more. However the $N = 4$ theory has one very attractive feature, namely that there is precisely one irreducible representation of the supersymmetry algebra fitting the renormalisability criterion $|h| \leq 1$, which, moreover, has to be "short", and therefore satisfy the universal mass formula (13) [30]. It follows that any dyon state must, willynilly, lie in a multiplet isomorphic to the one containing the gauge particles. Correspondingly there is only one supermultiplet of fields and, as a result, the supersymmetric action is unique apart from the values of the coupling constants.

However, there is an even more compelling reason for $N = 4$ supersymmetry which emerged some years later. In a series of papers it became apparent that the Callan-Symanzik β-function vanished identically in the unique $N = 4$ supersymmetric theory [34,35]. This was therefore the first example of a quantum field theory in four dimensions with this property. The vanishing has at least three remarkable consequences favourable to the ideas considered:

(1) As β controls the running of the coupling constant, its vanishing means that the gauge coupling constant does not renormalise. Presumably this applies in both the electric and magnetic formulations and it means that there is no question whether the Dirac quantisation applies to the bare or renormalised coupling constants, as these are the same (Rossi [36]).

(2) The trace of the energy momentum tensor is usually proportional to β times a local quantity and so it should vanish in this theory, indicating that the theory is exactly conformally invariant (if a vanishes). Thus the $N = 4$ supersymmetric gauge theory is the first known example of a quantum conformal field theory in four dimensions, to be compared with the rich spectrum of examples in two dimensions. Furthermore, the Higgs mechanism producing a nonzero value for the vacuum expectation value parameter a presumably provides an integrable deformation realising Zamolodchikov's principle in four dimensions [8]. Notice that the naive idea that conformal field theories should be more numerous in four rather than two dimensions seems to be false despite the fact that the conformal algebra has only fifteen rather than an infinite number of dimensions. Besides the $N = 4$ supersymmetric gauge theory, there are now a few other known conformal field theories in four dimensions, all supersymmetric gauge theories.

(3) Finally, just as the trace anomaly vanishes, so does the axial anomaly. In fact the two properties are related by a supersymmetry transformation. This means that the chiral symmetry (16) can be extended to the fields of the theory and is an exact symmetry for $N = 4$. Thus we have answered an earlier question and seen that, indeed, electromagnetic duality rotations can be extended to include matter, albeit in a very special case.

The second point above, concerning the realisation of Zamolodchikov's principle in four dimensions via a special sort of Higgs mechanism [8], raises questions about the nature of "integrability" in four dimensions. As far as is known, the $N = 4$ supersymmetry algebra is the largest extension of Poincaré symmetry there, but only provides a finite number of conservation laws, unlike the infinite number available in two dimensions. On the other hand, there are, apparently, monopole/dyon solutions with particle-like attributes (certainly if the duality conjecture is to be believed). But a complete and direct proof is lacking, even though the results for like monopoles are encouraging.

For each value of magnetic charge, the moduli space of solutions to the Bogomolny equations (12) forms a manifold whose points correspond to static configurations of distinct monopoles with total energy $a|g|$. The problem of describing their relative motion was answered by Manton [37], at least if it was slow. His idea follows from the analogy with a Newtonian point particle confined to move freely on a Riemannian manifold. It can remain at rest at any point of the manifold, but, if it moves, it follows a geodesic on the manifold determined by the Riemannian metric. He realised that the moduli spaces of the Bogomolny equations must possess such a metric and saw how to derive it from the action. Actually it has a hyperkähler structure which makes it very interesting mathematically. Moreover, Atiyah and Hitchin calculated the metric explicitly for the moduli space with double magnetic charge [38]. This is sufficient to determine the classical scattering of two monopoles at low relative velocity and yielded surprisingly involved behaviour, including a type of incipient breathing motion perpendicular to the scattering plane, visible on a video prepared by IBM.

Despite these beautiful results, there is no idea of how to describe relative motion of monopole solitons with unlike charge. The duality conjecture predicts the possibility of pair annihilation, unlike the sine-Gordon situation. This is why we say the soliton behaviour is incompletely understood. It is certainly more complicated than in two dimensions.

15.3.8 The Schwinger Quantisation Condition and the Charge Lattice

The remaining difficulty, one that has been repeatedly deferred, concerns the dyon spectrum. We know that there exist dyon solutions carrying magnetic charge, but we do not know what values of the electric charge are allowed. The problem is that the Dirac quantisation condition (7) does not determine this, nor does it respect the electromagnetic duality rotation (6) which is apparently so fundamental.

It was Schwinger and Zwanziger who independently resolved the problem [39,24]. They saw that Dirac's assumption that the monopole carried no electric charge was unjustified, and responsible for the difficulties. Instead, they applied Dirac's argument to two dyons, carrying respective charges (q_1, g_1) and (q_2, g_2), and found

$$q_1 g_2 - q_2 g_1 = 2\pi n \lambda, \qquad n = 0, \pm 1, \pm 2, \ldots . \tag{17}$$

This is known (somewhat unfairly) as the Schwinger quantisation condition and it does now respect the duality rotation symmetry (6) applied simultaneously to the two dyons. Notice that it is significant that the group $SO(2)$ has two invariant tensors, the Kronecker delta entering the mass formula (13) and the antisymmetric tensor entering (17).

As mentioned earlier, the values of $q + ig$ realised by localised states composed of particles should lie at the points of a lattice in the complex plane. The origin of this lattice structure are the conservation laws for charge and the TCP theorem. The set of allowed values must be closed under both addition and reversal of sign as these operations can be realised physically by combining states and by TCP conjugation.

Without loss of generality, it can be assumed that there exist a subset of states carrying purely electric charge. As long as magnetic charge exists (17) implies that there is a minimum positive value, q_0, say. Then the allowed values of pure electric charge are nq_0, $n \in \mathbb{Z}$, that is, a discrete one dimensional lattice. Now let us examine the most general values of $q + ig$ allowed by the Schwinger quantisation condition, (17). By it, the smallest allowed positive magnetic charge, g_0 satisfies

$$g_0 = \frac{2\pi n_0 \lambda}{q_0}, \tag{18}$$

where n_0 is a positive integer dependent on the detailed theory considered. Now consider two dyons with magnetic charge g_0 and electric charges q_1 and q_2 respectively. By (17) and (18)

$$q_1 - q_2 = \frac{2\pi n \lambda}{g_0} = \frac{n q_0}{n_0}.$$

However, as there must consequently be a state with pure electric charge $q_1 - q_2$, n must be a multiple of n_0. Hence for any dyon with magnetic charge g_0, its electric charge

$$q = q_0 \left(n + \frac{\theta}{2\pi} \right),$$

where θ is a new parameter of the theory which is, in a sense, angular since increasing it by 2π is equivalent to increasing n by one unit. So

$$q + ig = q_0(n + \tau),$$

where

$$\tau = \frac{\theta}{2\pi} + \frac{2\pi i n_0 \lambda}{q_0^2}. \tag{19}$$

Repeating the argument for more general states with magnetic charge $m g_0$

$$q + ig = q_0(m\tau + n), \qquad m, n \in \mathbb{Z}. \tag{20}$$

This is the charge lattice and it finally breaks the continuous symmetry (2) and (6) in a spontaneous manner [40]. This lattice has periods q_0 and $q_0\tau$ with ratio τ, (19). Notice that τ is a complex variable formed of dimensionless parameters dependent on the detailed theory. Its imaginary part is positive, being essentially the inverse of the fine structure constant.

So far, this part of the argument has been very general, but, given a specific theory, an important question for electromagnetic duality concerns the identification of the subset of the charge lattice that can be realised by single particle states, rather than multiparticle states.

It is easy to show that, if single particle states obey the universal mass formula (13), and are stable with respect to any two-body decay into lighter particles permitted by the conservation of electric and magnetic charge, then they must correspond to points of the charge lattice which are "primitive vectors".

A point P of the charge lattice is a primitive vector if the line OP contains no other points of the lattice strictly between O, the origin, and P. Thus the only primitive vectors on the real axis are $(\pm q_0, 0)$. Equivalently, a primitive vector is a point given by (20) in which the integers m and n are coprime (in saying this we must agree that 0 is divisible by any integer).

The proof of the assertion is simple: it relies on the fact that the mass of a particle at P is proportional to its Euclidean distance OP from the origin, by (13). So, by the triangle inequality, any particle is stable unless its two decay products correspond to points collinear with itself and the origin. This is impossible, providing the original particle corresponds to a primitive vector.

There are an infinite number of primitive vectors on the charge lattice, for example, all the points with $m = \pm 1$ or $n = \pm 1$. The corresponding masses can be indefinitely large. If $m = 2$, every second point is a primitive vector. If $m = 3$, every third point fails to be a primitive vector, and so on.

This result tells us what to expect for the spectrum of dyons, namely that they correspond to the primitive vectors off the real axis. Since the mass formula used in this argument is characteristic of supersymmetric gauge theories as discussed above, it ought to be possible to recover this result from consideration of the Bogomolny moduli spaces governing the static soliton solutions. This is what Sen achieved in 1994, [28], and a simplified explanation follows.

For $m = \pm 1$, the dyons relate to points in the $m = \pm 1$ moduli space since the single monopoles are solutions to the Bogomolny equations. However, as discussed earlier, the points of the $m = \pm 2$ moduli space correspond to configurations of a pair of like monopoles in static equilibrium. Thus the $m = \pm 2$ single particle states cannot be Bogomolny solutions. Instead they must be regarded as quantum mechanical bound states, with zero binding energy (in order to satisfy the mass formula). Remembering Manton's treatment of moving monopoles following geodesics on the moduli space determined by the hyperkähler metric thereon, it is clear that it is crucial to examine the spectrum of the Laplacian determined by this metric, as this is proportional to the quantum mechanical Hamiltonian [41]. In particular, zero modes in the discrete spectrum are sought. There is some subtlety, treated by Sen, concerning the fact that the quantum mechanics possesses $N = 4$ supersymmetry because the metric is hyperkähler, but using the Atiyah-Hitchin metric, Sen was able to solve for the zero modes, and show that only every other permitted value of the electric charge could occur. Thus the dyons with magnetic charge $2g_0$ do indeed correspond precisely to the primitive vectors on the charge lattice. For higher values of $|m|$, the explicit metric is not known, but Hodge's theorem relates the counting of the zero modes of the Laplacian on the moduli space to its cohomology, which can be determined without knowledge of the metric. (This argument is said to be due to Segal, unpublished).

These are the results that finally clear up the dyon problem and leave the electromagnetic duality conjecture in good shape, though a reassessment will be in order. Before discussing this, we ask whether the angle θ, occurring in (19), appears explicitly as a parameter in the action of the spontaneously broken gauge theory. Witten found the answer in 1979 [40]. Because the gauge group is non abelian, $SU(2)$, a term proportional to the instanton number, k can be added to the action, so that the Feynman weighting factor becomes:

$$exp\left(\frac{i\text{Action}}{\lambda}\right) \to exp\left(\frac{i\text{Action}}{\lambda} + \frac{i\tilde{\theta}}{2\pi}k\right). \tag{21}$$

As k is proportional to an integral of $F\tilde{F}$ over space time, it is a surface term which cannot affect the classical equations of motion, but it does affect the quantum theory. Note that, like θ, $\tilde{\theta}$ is an angular variable as the theory is unaffected if it is increased by 2π. In fact the two angles are indeed equal as Witten showed by an elementary calculation of the electric and magnetic charges using Noether's theorem. Thus θ is what is known as the instanton or vacuum angle.

The above result has another consequence, yet again singling out the $N = 4$ supersymmetric theory as the only viable one for exact electromagnetic duality. This is because an application of the chiral rotation (16) to the fermion fields alters the Lagrangian density by an anomalous term proportional to the axial anomaly $\beta F\tilde{F}$. This means that the instanton angle can be altered by a redefinition of the fermion field, and so has no physical meaning, unless β, and hence the axial anomaly, vanishes. This forces us back to the $N = 4$ theory, with the conclusion that only in this theory does the charge lattice really make sense. Finally note that in this theory the integer n_0 occurring in (19) equals 2. This is because the $N = 4$ theory has only one supermultiplet which includes the gauge particle and hence must be an $SU(2)$ triplet. No doublets are allowed in $N = 4$, unlike $N = 2$.

15.3.9 Exact Electromagnetic Duality and the Modular Group

Armed with the new insight that the spectrum of single particle states correspond to the primitive vectors of the charge lattice, augmented by the origin, rather than the five points previously considered, we can see that the original Montonen-Olive conjecture was too modest. Instead of possessing two equivalent choices of action, the $N = 4$ supersymmetric gauge theory apparently possesses an infinite number of them, all with an isomorphic structure, but with different values of the parameters [28].

Roughly speaking, the reason is that it is the charge lattice that describes the physical reality. Choices of action correspond to choices of basis in the lattice, that is a pair of non collinear primitive vectors (or, a pair of periods). As the charge lattice is two dimensional, these choices are related by the action of the modular group, an infinite discrete group containing the previous transformation (14).

Let us choose a primitive vector in the charge lattice, represented by a complex number, q_0', say. Then we may ascribe short $N = 4$ supermultiplets of quantum fields to each of the three points $\pm q_0'$ and 0. The particles corresponding to the origin are massless and neutral whereas the particles corresponding to $\pm q_0'$ possess complex charge $\pm q_0'$ and mass $a|q_0'|$. We may form an $N = 4$ supersymmetric action with these fields. It is unique, given the coupling $|q_0'|$, apart from the vacuum angle whose specification requires a second primitive vector, $q_0'\tau'$, say, non-collinear with q_0'. The remaining single particle states are expected to arise as monopole solitons or as quantum bound states of them as discussed above.

Since the two non-collinear primitive vectors q_0' and $q_0'\tau'$ form an alternative basis for the charge lattice, they can be expressed as integer linear combinations of the original basis, q_0 and $q_0\tau$:

$$q_0'\tau' = aq_o\tau + bq_0, \tag{22a}$$

$$q_0' = cq_0\tau + dq_0, \tag{22b}$$

where

$$a, b, c, d \in \mathbb{Z}. \tag{22c}$$

Equally, $q_0\tau$ and q_0 can be expressed as integer linear combinations of $q_0'\tau'$ and q_0'. This requires that the matrix of coefficients in (22a) and (22b) has determinant equal to ± 1,

$$ad - bc = \pm 1. \tag{23}$$

By changing a sign we can take this to be plus one. Then the matrices

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

form a group, $SL(2, \mathbb{Z})$, whose quotient by its centre is called the modular group. Equation (22a) divided by (22b) yields

$$\tau' = \frac{a\tau + b}{c\tau + d}.$$

These transformations form the modular group and preserve the sign of the imaginary part of τ. This gives the relation between the values of the dimensionless parameters in the two choices of action corresponding to the two choices of basis. It is customary to think of the modular group as being generated by elements T and S where

$$T : \tau \to \tau + 1 \qquad S : \tau \to -\frac{1}{\tau}$$

According to (19), T increases the vacuum angle by 2π. This is obviously a symmetry of (21). If the vacuum angle vanishes, S precisely yields the transformation (14) previously considered.

Proof of the quantum equivalence of all the actions associated with each choice of basis in the charge lattice would presumably require a generalised vertex operator transformation relating the corresponding quantum fields. Since these transformations would represent the modular group the prospect is challenging.

Meanwhile it has been possible to evaluate the partition function of the theory on certain space-time manifolds, provided that the theory is simplified by a “twisting procedure" that renders it “topological". Vafa and Witten verified that the results indeed possessed modular symmetry [42].

15.3.10 Conclusion

According to the new results reviewed above, it now appears increasingly plausible that electromagnetic duality is realised exactly in the $N = 4$ supersymmetric $SU(2)$ gauge theory in which the Higgs field acquires a non-zero vacuum expectation value. This theory is a deformation of one of the very few exact conformal field theories in Minkowski space time. The supporting analysis involves an array of almost all the previously advanced ideas particular to quantum field theories in four dimensions, but awaits definitive proof.

Despite its remarkable quantum symmetry this theory is apparently not physical unless further deformed. Seiberg and Witten have proposed deformations such that enough structure remains as to offer an explanation of quark confinement, perhaps the outstanding riddle in quantum field theory [1,43].

More generally, the potential validity of exact electromagnetic duality in at least one theory means that quantum field theory in four dimensions is much richer than the sum of its parts, quantum mechanics and classical field theory. This is because the new symmetry is essentially quantum in nature with no classical counterpart. Moreover it relates strong to weak coupling regimes of the theory. Consequently, the new insight opens a veritable Pandora's box whose contents are now subject to urgent study.

REFERENCES from Exact electromagnetic duality by David I Olive

1. N Seiberg and E Witten, *Nucl Phys* **B426** (94) 19-52, *Erratum* **B430** (94) 485-486, "Electromagnetic duality, monopole condensation, and confinement in $N = 2$ supersymmetric Yang-Mills theory".
2. E Witten, hep-th9411102, "Monopoles and four-manifolds"
3. S Donaldson and P Kronheimer, "The geometry of four manifolds", (Oxford 1990)
4. L Silberstein, *Ann d Physik* **24** (07) 783-784, "Nachtrag zur Abhandlung über "elektromagnetische Grundgleichungen in bivektorieller Behandlung " "
5. P Higgs, *Phys Rev* **145** (66) 1156-1163, "Spontaneous symmetry breakdown without massless bosons"
6. F Englert and R Brout, *Phys Rev Lett* **13** (64) 321-323, "Broken symmetry and the mass of gauge vector bosons"
7. TWB Kibble, *Phys Rev* **155** (67) 1554-1561, "Symmetry breaking in non-abelian gauge theories"
8. AB Zamolodchikov, *Advanced Studies in Pure Mathematics* **19** (89) 642-674, "Integrable Field Theory from Conformal field Theory"
9. RJ Baxter, "Exactly solved models in statistical mechanics", (Academic Press 1982)
10. JK Perring and THR Skyrme, *Nucl Phys* **31** (62) 550-555, "A model unified field equation"
11. THR Skyrme, *Proc Roy Soc* **A262** (61) 237-245, "Particle states of a quantized meson field"
12. S Coleman, *Phys Rev* **D11** (75) 2088-2097, "Quantum sine-Gordon equation as the massive Thirring model"
13. S Mandelstam, *Phys Rev* **D11** (75) 3026-3030, "Soliton operators for the quantized sine-Gordon equation"
14. HW Braden, EF Corrigan, PE Dorey and R Sasaki, *Nucl Phys* **B338** (90) 689-746, "Affine Toda field theory and exact S-matrices"
15. DI Olive, N Turok and JWR Underwood, *Nucl Phys* **B401** (93) 663-697, "Solitons and the energy-momentum tensor for affine Toda theory"
16. PAM Dirac, *Proc Roy Soc* **A33** (31) 60-72, "Quantised singularities in the electromagnetic field"
17. G 't Hooft, *Nucl Phys* **B79** (74) 276-284, "Magnetic monopoles in unified gauge theories"
18. AM Polyakov, *JETP Lett* **20** (74) 194-195, "Particle spectrum in quantum field theory"
19. P Goddard and DI Olive, *Rep on Prog in Phys* **41** (78) 1357-1437, "Magnetic monopoles in gauge field theories"
20. EB Bogomolny, *Sov J Nucl Phys* **24** (76) 449-454, "The stability of classical solutions"
21. MK Prasad and CM Sommerfield, *Phys Rev Lett* **35** (75) 760-762, "Exact classical solution for the 't Hooft monopole and the Julia-Zee dyon"
22. N Manton, *Nucl Phys* **B126** (77) 525-541, "The force between 't Hooft-Polyakov monopoles"
23. B Julia and A Zee, *Phys Rev* **D11** (75) 2227-2232, "Poles with both magnetic and electric charges in non-Abelian gauge theory"
24. J Schwinger, *Science* **165** (69) 757-761, "A magnetic model of matter"
25. S Coleman, S Parke, A Neveu, and CM Sommerfield, *Phys Rev* **D15** (77) 544-545, "Can one dent a dyon?"

26. C Montonen and D Olive, *Phys Lett* **72B** (77) 117-120, "Magnetic monopoles as gauge particles?"
27. P Goddard, J Nuyts and D Olive, *Nucl Phys* **B125** (77) 1-28, "Gauge theories and magnetic charge"
28. A Sen, *Phys Lett* **329B** (94) 217-221, "Dyon-monopole bound states, self-dual harmonic forms on the multi-monopole moduli space, and $SL(2,\mathbb{Z})$ invariance in string theory"
29. A D'Adda, R Horsley and P Di Vecchia, *Phys Lett* **76B** (78) 298-302 "Supersymmetric monopoles and dyons"
30. H Osborn, *Phys Lett* **83B** (79) 321-326, "Topological charges for $N=4$ supersymmetric gauge theories and monopoles of spin 1"
31. W Nahm, *Nucl Phys* **B135** (78) 149-166, "Supersymmetries and their representations"
32. E Witten and D Olive, *Phys Lett* **78B** (78) 97-101, "Supersymmetry algebras that include topological charges"
33. R Haag, JT Lopuszanski and M Sohnius, *Nucl Phys* **B88** (75) 257-274, "All possible generators of supersymmetry of the S-matrix"
34. S Mandelstam, *Nucl Phys* **B213** (83) 149-168, "Light-cone superspace and the ultraviolet finiteness of the $N=4$ model"
35. L Brink, O Lindgren and BEW Nilsson, *Phys Lett* **123B** (83) 323-328 "The ultra-violet finiteness of the $N=4$ Yang-Mills theory"
36. P Rossi, *Phys Lett* **99B** (81) 229-231, "$N=4$ supersymmetric monopoles and the vanishing of the β function"
37. N Manton, *Phys Lett* **110B** (82) 54-56, "A remark on the scattering of BPS monopoles"
38. MF Atiyah and NJ Hitchin *Phys Lett* **107A** (85) 21-25, "Low energy scattering of non-abelian monopoles"
39. D Zwanziger, *Phys Rev* **176** (68) 1489-1495, "Quantum field theory of particles with both electric and magnetic charges"
40. E Witten *Phys Lett* **86B** (79) 283-287, "Dyons of charge $e\theta/2\pi$"
41. G Gibbons and N Manton, *Nucl Phys* **B274** (86) 183-224, "Classical and quantum dynamics of monopoles"
42. C Vafa and E Witten, *Nucl Phys* **B431** (94) 3-77, "A strong coupling test of S-duality"
43. N Seiberg and E Witten, *Nucl Phys* **B431** (94) 484-550, "Monopoles, duality and chiral symmetry breaking in $N=2$ supersymmetric QCD"

15.4 Duality in Quantum Field Theory and String Theory by Álvarez-Gaumé Zamora

These lectures give an introduction to duality in Quantum Field Theory. We discuss the phases of gauge theories and the implications of the electric-magnetic duality transformation to describe the mechanism of confinement. We review the exact results of $N=1$ supersymmetric QCD and the Seiberg-Witten solution of $N=2$ super Yang-Mills. Some of its extensions to String Theory are also briefly discussed.

15.4.1 The duality symmetry.

From a historical point of view we can say that many of the fundamental concepts of twentieth century Physics have Maxwell's equations at its origin. In particular some of the symmetries that have led to our understanding of the fundamental interactions in terms of relativistic quantum field theories have their roots in the equations describing electromagnetism. As we will now describe, the most basic form of the duality symmetry also appears in the source free

Maxwell equations:

$$\nabla \cdot (\mathbf{E} + i\,\mathbf{B}) = 0\,,$$
$$\frac{\partial}{\partial t}(\mathbf{E} + i\,\mathbf{B}) + i\,\nabla \times (\mathbf{E} + i\,\mathbf{B}) = 0. \tag{15.298}$$

These equations are invariant under Lorentz transformations, and making all of Physics compatible with these symmetries led Einstein to formulate the Theory of Relativity. Other important symmetries of (15.298) are conformal and gauge invariance, which have later played important roles in our understanding of phase transitions and critical phenomena, and in the formulation of the fundamental interactions in terms of gauge theories. In these lectures however we will study the implications of yet another symmetry hidden in (15.298): duality. The simplest form of duality is the invariance of (15.298) under the interchange of electric and magnetic fields:

$$\mathbf{B} \to \mathbf{E}\,,$$
$$\mathbf{E} \to -\mathbf{B}\,. \tag{15.299}$$

In fact, the vacuum Maxwell equations (15.298) admit a continuous $SO(2)$ transformation symmetry *

$$(\mathbf{E} + i\,\mathbf{B}) \to e^{i\phi}(\mathbf{E} + i\,\mathbf{B})\,. \tag{15.300}$$

If we include ordinary electric sources the equations (1.1) become:

$$\nabla \cdot (\mathbf{E} + i\,\mathbf{B}) = q\,,$$
$$\frac{\partial}{\partial t}(\mathbf{E} + i\,\mathbf{B}) + i\,\nabla \times (\mathbf{E} + i\,\mathbf{B}) = \mathbf{j}_e\,. \tag{15.301}$$

In presence of matter, the duality symmetry is not valid. To keep it, magnetic sources have to be introduced:

$$\nabla \cdot (\mathbf{E} + i\,\mathbf{B}) =)q + ig)\,,$$
$$\frac{\partial}{\partial t}(\mathbf{E} + i\,\mathbf{B}) + i\nabla \times (\mathbf{E} + i\,\mathbf{B}) = (\mathbf{j}_e + i\,\mathbf{j}_m)\,. \tag{15.302}$$

Now the duality symmetry is restored if at the same time we also rotate the electric and magnetic charges

$$(q + ig) \to e^{i\phi}(q + ig)\,. \tag{15.303}$$

The complete physical meaning of the duality symmetry is still not clear, but a lot of work has been dedicated in recent years to understand the implications of this type of symmetry. We will focus mainly on the applications to Quantum Field Theory. In the final sections, we will briefly review some of the applications to String Theory, where duality make striking an profound predictions.

15.4.2 Dirac's charge quantization.

From the classical point of view the inclusion of magnetic charges is not particularly problematic. Since the Maxwell equations, and the Lorentz equations of motion for electric and magnetic charges only involve the electric and magnetic field, the classical theory can accommodate any values for the electric and magnetic charges.

However, when we try to make a consistent quantum theory including monopoles, deep consequences are obtained. Dirac obtained his celebrated quantization condition precisely by

*Notice that the duality transformations are not a symmetry of the electromagnetic action. Concerning this issue see [1].

studying the consistency conditions for a quantum theory in the presence of electric and magnetic charges [2]. We derive it here by the quantization of the angular momentum, since it allows to extend it to the case of dyons, *i.e.*, particles that carry both electric and magnetic charges.

Consider a non-relativistic charge q in the vicinity of a magnetic monopole of strength g, situated at the origin. The charge q experiences a force $m\ddot{\vec{r}} = q\dot{\vec{r}} \times \vec{B}$, where $\vec{B}$ is the monopole field given by $\vec{B} = g\vec{r}/4\pi r^3$. The change in the orbital angular momentum of the electric charge under the effect of this force is given by

$$\begin{aligned} \frac{d}{dt}\left(m\vec{r}\times\dot{\vec{r}}\right) &= m\vec{r}\times\ddot{\vec{r}} \\ &= \frac{qg}{4\pi r^3}\,\vec{r}\times\left(\dot{\vec{r}}\times\vec{r}\right) = \frac{d}{dt}\left(\frac{qg}{4\pi}\,\frac{\vec{r}}{r}\right). \end{aligned} \tag{15.304}$$

Hence, the total conserved angular momentum of the system is

$$\vec{J} = \vec{r}\times m\dot{\vec{r}} - \frac{qg}{4\pi}\,\frac{\vec{r}}{r}\,. \tag{15.305}$$

The second term on the right hand side (henceforth denoted by $\vec{J}_{em}$) is the contribution coming from the electromagnetic field. This term can be directly computed by using the fact that the momentum density of an electromagnetic field is given by its Poynting vector, $\vec{E}\times\vec{B}$, and hence its contribution to the angular momentum is given by

$$\vec{J}_{em} = \int d^3x\,\vec{r}\times(\vec{E}\times\vec{B}) = \frac{g}{4\pi}\int d^3x\,\vec{r}\times\left(\vec{E}\times\frac{\vec{r}}{r^3}\right).$$

In components,

$$\begin{aligned} J^i_{em} &= \frac{g}{4\pi}\int d^3x E^j\partial_j(\hat{x}^i) \\ &= \frac{g}{4\pi}\int_{S^2}\hat{x}^i\vec{E}\cdot d\vec{s} - \frac{g}{4\pi}\int d^3x(\vec{\nabla}\cdot\vec{E})\,\hat{x}^i\,. \end{aligned} \tag{15.306}$$

When the separation between the electric and magnetic charges is negligible compared to their distance from the boundary S^2, the contribution of the first integral to $\vec{J}_{em}$ vanishes by spherical symmetry. We are therefore left with

$$\vec{J}_{em} = -\frac{gq}{4\pi}\hat{r}\,. \tag{15.307}$$

Returning to equation (15.305), if we assume that orbital angular momentum is quantized. Then it follows that

$$\frac{qg}{4\pi} = \frac{1}{2}n\,, \tag{15.308}$$

where n is an integer. Equation (15.308) is the Dirac's charge quantization condition. It implies that if there exists a magnetic monopole of charge g somewhere in the universe, then all electric charges are quantized in units of $2\pi/g$. If we have a number of purely electric charges q_i and purely magnetic charges g_j, then any pair of them will satisfy a quantization condition:

$$q_i g_j = 2\pi n_{ij}\,. \tag{15.309}$$

Thus, any electric charge is an integral multiple of $2\pi/g_j$. For a given g_j, let these charges have n_{0j} as the highest common factor. Then, all the electric charges are multiples of $q_0 = n_{0j}2\pi/g_j$. Similar considerations apply to the quantization of the magnetic charge.

Till now, we have only dealt with particles that carry either an electric or a magnetic charge. Consider now two dyons of charges (q_1, g_1) and (q_2, g_2). For this system, we can repeat the calculation of $\vec{J}_{em}$ by following the steps in (15.306) where now the electromagnetic fields are split as $\vec{E} = \vec{E}_1 + \vec{E}_2$ and $\vec{B} = \vec{B}_1 + \vec{B}_2$. The answer is easily found to be

$$\vec{J}_{em} = -\frac{1}{4\pi}(q_1 g_2 - q_2 g_1)\,\hat{r} \tag{15.310}$$

The charge quantization condition is thus generalized to

$$\frac{q_1 g_2 - q_2 g_1}{4\pi} = \frac{1}{2} n_{12} \tag{15.311}$$

This is referred to as the Dirac-Schwinger-Zwanziger condition [3].

15.4.3 A charge lattice and the $SL(2, \mathbf{Z})$ group.

In the previous section we derived the quantization of the electric charge of particles without magnetic charge, in terms of some smallest electric charge q_0. For a dyon (q_n, g_n), this gives $q_0 g_n = 2\pi n$. Thus, the smallest magnetic charge the dyon can have is $g_0 = 2\pi m_0/q_0$, with m_0 a positive integer dependent on the detailed theory considered. For two dyons of the same magnetic charge g_0 and electric charges q_1 and q_2, the quantization condition implies $q_1 - q_2 = nq_0$, with n a multiple of m_0. Therefore, although the difference of electric charges is quantized, the individual charges are still arbitrary. It introduces a new parameter θ that contributes to the electric charge of any dyon with magnetic charge g_0 by

$$q = q_0\left(n_e + \frac{\theta}{2\pi}\right). \tag{15.312}$$

Observe that the parameter $\theta + 2\pi$ gives the same electric charges that the parameter θ by shifting $n_e \to n_e + 1$. Thus, we look at the parameter θ as an angular variable.

This arbitrariness in the electric charge of dyons through the θ parameter can be fixed if the theory is CP invariant. Under a CP transformation $(q, g) \to (-q, g)$. If the theory is CP invariant, the existence of a state (q, g_0) necessarily leads to the existence of $(-q, g_0)$. Applying the quantization condition to this pair, we get $2q = q_0 \times \mathit{integer}$. This implies that $q = nq_0$ or $q = (n + \frac{1}{2})q_0$. If $\theta \neq 0, \pi$, the theory is not CP invariant. It indicates that the θ parameter is a source of CP violation. Later on we will identify θ with the instanton angle.

One can see that the general solution of the Dirac-Schwinger-Zwanziger condition (15.311) is

$$q = q_0\left(n_e + \frac{\theta}{2\pi} n_m\right), \tag{15.313}$$

$$g = n_m g_0\,, \tag{15.314}$$

with n_e and n_m integer numbers These equations can be expressed in terms of the complex number

$$q + ig = q_0(n_e + n_m \tau)\,, \tag{15.315}$$

where

$$\tau \equiv \frac{\theta}{2\pi} + \frac{2\pi i m_0}{q_0^2}\,. \tag{15.316}$$

Observe that this definition only includes intrinsic parameters of the theory, and that the imaginary part of τ is positive definite. This complex parameter will play an important role in supersymmetric gauge theories. Thus, physical states with electric and magnetic charges (q, g)

are located on a discrete two dimensional lattice with periods q_0 and $q_0\tau$, and are represented by the corresponding vector (n_m, n_e) (see fig. 1).

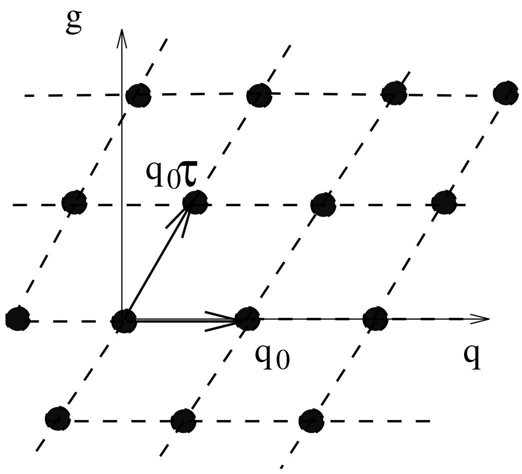


Figure 1: The charge lattice with periods q_0 and $q_0\tau$. The physical states are located on the points of the lattice.

Notice that the lattice of charges obtained from the quantization condition breaks the classical duality symmetry group $SO(2)$ that rotated the electric and magnetic charges (15.303). But another symmetry group arises at quantum level. Given a lattice as in figure 1 we can describe it in terms of different fundamental cells. Different choices correspond to transforming the electric and magnetic numbers (n_m, n_e) by a two-by-two matrix:

$$(n_m, n_e) \to (n_m, n_e) \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}^{-1}, \tag{15.317}$$

with $\alpha, \beta, \gamma, \delta \in \mathbf{Z}$ satisfying $\alpha\delta - \beta\gamma = 1$. This transformation leaves invariant the Dirac-Schwinger-Zwanziger quantiztion condition (15.311). Hence the duality transformations are elements of the discrete group $SL(2,\mathbf{Z})$. Its action on the charge lattice can be implemented by modular transformations of the parameter τ

$$\tau \to \frac{\alpha\tau + \beta}{\gamma\tau + \delta}. \tag{15.318}$$

This transformations preserve the sign of the imaginary part of τ, and are generated just by the action of two elements:

$$T: \quad \tau \to \tau + 1, \tag{15.319}$$

$$S: \quad \tau \to \frac{-1}{\tau}. \tag{15.320}$$

The effect of T is to shift $\theta \to \theta + 2\pi$. Its action is well understood: it just maps the charge lattice (n_m, n_e) to $(n_m, n_e - n_m)$. As physics is 2π-periodic in θ, it is a symmetry of the theory. Then, if the state $(1,0)$ is in the physical spectrum, the state $(1, n_e)$, with any integer n_e, is also a physical state.

The effect of S is less trivial. If we take $\theta = 0$ just for simplicity, the S action is $q_0 \to g_0$ and sends the lattice vector (n_m, n_e) to the lattice vector $(-n_e, n_m)$. So it interchanges the electric and magnetic roles. In terms of coupling constants, it represents the transformation $\tau \to -1/\tau$, implying the exchange between the weak and strong coupling regimes. In this respect the duality symmetry could provide a new source of information on nonpertubative physics.

If we claim that the S transformation is also a symmetry of the theory we have full $SL(2, \mathbf{Z})$ symmetry. It implies the existence of any state (n_m, n_e) in the physical spectrum, with n_m and n_e relatively to-prime, just from the knowledge that there are the physical states $\pm(0,1)$ and $\pm(1,0)$. There are some examples of theories 'duality invariant', for instance the $SU(2)$ gauge theory with $N = 4$ supersymmetry and the $SU(2)$ gauge theory with $N = 2$ supersymmetry and four flavors [4].

A priori however there is no physical reason to impose S-invariance, in contrast with T-invariance. The stable physical spectrum may not be $SL(2, \mathbf{Z})$ invariant. But if the theory still admits somehow magnetic monopoles, we could apply the S-transformation as a change of variables of the theory, where a magnetic state is mapped to an electric state in terms of the dual variables. It could be convenient for several reasons: Maybe there are some physical phenomena where the magnetic monopoles become relevant degrees of freedom; this is the case for the mechanism of confinement, as we will see below. The other reason could be the difficulty in the computation of some dynamical effects in terms of the original electric variables because of the large value of the electric coupling q_0. The S-transformation sends q_0 to $1/q_0$. In terms of the dual magnetic variables, the physics is weakly coupled.

Just by general arguments we have learned a good deal of information about the duality transformations. Next we have to see where such concepts appear in quantum field theory.

15.4.4 The Higgs Phase

The Higgs mechanism and mass gap.

We start considering that the relevant degrees of freedom at large distances of some theory in 3+1 dimensions are reduced to an Abelian Higgs model:

$$\begin{aligned} \mathcal{L}(\phi^*, \phi, A_\mu) = &-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (D_\mu\phi)^*(D^\mu\phi) \\ &- \frac{\lambda}{2}(\phi^*\phi - M^2)^2 , \end{aligned} \tag{15.321}$$

where

$$\begin{aligned} F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu , \\ D_\mu\phi &= (\partial_\mu + iqA_\mu)\phi , \end{aligned} \tag{15.322}$$

and q is the electric charge of the particle ϕ.

An important physical example of a theory described at large distances by the effective Lagrangian (15.321) (in its nonrelativistic approximation) is a superconductor. Sound waves of a solid material causes complicated deviations from the ideal lattice of the material. Conducting electrons interact with the quantums of those sound waves, called phonons. For electrons near the Fermi surface, their interactions with the phonons create an attractive force. This force can be strong enough to cause bound states of two electrons with opposite spin, called Cooper pairs. The lowest state is a scalar particle with charge $q = -2e$, which is represented by ϕ in (15.321). To understand the basic features of a superconductor we only need to consider its relevant self-interactions and the interaction with the electromagnetic field resulting from its electric charge q. This is the dynamics which is encoded in the effective Lagrangian (15.321). The values of the parameters λ and M^2 depend of the temperature T, and in general contribute to increase the energy of the system. To have an stable ground state, we require $\lambda(T) > 0$ for any value of the temperature. But the function $M^2(T)$ do not need to be negative for all T. In fact, when the temperature T drops below a critical value T_c, the function $M^2(T)$ becomes positive. In such situation, the ground state reaches its minimal energy when the Higgs particle condenses,

$$|\langle\phi\rangle| = M\,. \tag{15.323}$$

If we make perturbation theory around this minima,

$$\phi(x) = M + \varphi(x), \tag{15.324}$$

with vanishing external electromagnetic fields, we find that there is a mass gap between the ground state and the first excited levels. There are particles of spin one with mass square

$$\mathcal{M}_V^2 = 2qM^2, \tag{15.325}$$

which corresponds to the inverse of the penetration depth of static electromagnetic fields in the superconductor. There are also spin zero particles with mass square

$$\mathcal{M}_H^2 = 2\lambda M^2. \tag{15.326}$$

So perturbation theory already shows a quite different behavior of the Higgs theory from the Coulomb theory. There is only one real massive scalar field and the electromagnetic interaction becomes short-ranged, with the photon correlator being exponentially suppressed. This is a distinction that must survive nonperturbatively. But up to now, the above does not yet distinguish a Higgs theory from just any non-gauge theory with massive vector particles. There is yet another new phenomena in the Higgs mode which shows the spontaneous symmetry breaking of the $U(1)$ gauge theory.

Vortex tubes and flux quantization.

We have seen that the Higgs condensation produces the electromagnetic interactions to be short-range. Ignoring boundary effects in the material, the electric and magnetic fields are zero inside the superconductor. This phenomena is called the Meissner effect.

If we turn on an external magnetic field $\mathbf{H}_0$ beyond some critical value, one finds that small regions in the superconductor make a transition to a 'non-superconducting' state. Stable magnetic flux tubes are allowed along the material, with a transverse size of the order of the inverse of the mass gap. Their magnetic flux satisfy a quantization rule that can be understood only by a combination of the spontaneous symmetry breaking of the $U(1)$ gauge symmetry and some topological arguments.

Parameterize the complex Higgs field by

$$\phi(x) = \rho(x)e^{i\chi(x)}, \tag{15.327}$$

and perform fluctuations around the configuration which minimizes the energy. *i.e.*, we consider that $\rho(x) \simeq M$ almost everywhere, but at some points ρ may be zero. At such points χ needs not be well defined and therefore in all the rest of space χ could be multivalued. For instance, if we take a closed contour C around a zero of $\rho(x)$, then following χ around C could give values that run from 0 to $2\pi n$, with n an integer number, instead of coming back to zero. These are exactly the field configurations that produce the quantized magnetic flux tubes [5].

Consider a two-dimensional plane, cut somewhere through a superconducting piece of material, with polar coordinates (r, θ) and work in the time-like $A_0 = 0$. To have a finite energy per unit length static configuration we should demand that

$$\begin{aligned} \phi(x) &\to Me^{i\chi(\theta)}\,, \\ A_i(x) &\to \frac{\text{const}}{r}\,, \end{aligned} \tag{15.328}$$

for $r \to \infty$. Obviously, to keep the fields single valued, we must have

$$\chi(2\pi) = \chi(0) + 2\pi n \,. \tag{15.329}$$

If $n \neq 0$, it is clear that at some point of the two-dimensional plane we should have that the continuous field ϕ vanishes. Such field configurations do not correspond to the ground state.

Solve the field equations with the boundary conditions (15.328) and (15.329) fixed, and minimize the energy. We find stable vortex tubes with non-trivial magnetic flux through the two-dimensional plane. To see this, perform a singular gauge transformation *

$$\begin{aligned} \phi(x) &\to e^{iq\Lambda(x)}\phi(x)\,, \\ A_\mu(x) &\to A_\mu(x) - \partial_\mu \Lambda(x)\,, \end{aligned} \tag{15.330}$$

with $\Lambda = 2\pi n\theta/q$. We compute the magnetic flux in such a gauge and we find

$$\Phi = \oint A_\mu dx^\mu = \Lambda(2\pi) - \Lambda(0) = \frac{2\pi n}{q}\,. \tag{15.331}$$

It is important to realize that such field configurations, called Abrikosov vortices, are stable. The vortex tube cannot break since it cannot have an end point: as the magnetic flux is quantized, we would have be able to deform continuously the singular gauge transformation Λ to zero, something obviously not possible for $n \neq 0$. Physically this is the statement that the magnetic flux is conserved, a consequence of the Maxwell equations. Mathematically it means that for $n \neq 0$ the function $\chi(\theta)$ belongs to a nontrivial homotopy class of the fundamental group $\Pi_1(U(1)) = \mathbf{Z}$.

The existence of these macroscopic stable objects can be used as another characterization of the Higgs phase. They should survive beyond perturbation theory.

Magnetic monopoles and permanent magnetic confinement.

The magnetic flux conservation in the Abelian Higgs model tells us that the theory does not include magnetic monopoles. But it is remarkable that the magnetic flux is precisely a multiple of the quantum of magnetic charge $2\pi/q$ found by Dirac. If we imagine the effective gauge theory (15.321) enriched somehow by magnetic monopoles, they would form end points of the vortex tubes. The energy per unit length, *i.e.*, the string tension σ, of these flux tubes is of the order of the scale of the Higgs condensation,

$$\sigma \sim M^2. \tag{15.332}$$

It implies that the total energy of a system composed of a monopole and an anti-monopole, with a convenient magnetic flux tube attached between them, would be at least proportional to the separation length of the monopoles. In other words: magnetic monopoles in the Higgs phase are permanently confined.

15.4.5 The Georgi-Glashow model and the Coulomb phase.

The Georgi-Glashow model is a Yang-Mills-Higgs system which contains a Higgs multiplet ϕ^a ($a = 1, 2, 3$) transforming as a vector in the adjoint representation of the gauge group $SO(3)$, and the gauge fields $W_\mu = W_\mu^a T^a$. Here, T^a are the hermitian generators of $SO(3)$ satisfying $[T^a, T^b] = if^{abc}T^c$. In the adjoint representation, we have $(T^a)_{bc} = -if^a_{bc}$ and, for $SO(3)$, $f^{abc} = \epsilon^{abc}$. The field strength of W_μ and the covariant derivative on ϕ^a are defined by

$$\begin{aligned} G_{\mu\nu} &= \partial_\mu W_\nu - \partial_\nu W_\mu + ie[W_\mu, W_\nu]\,, \\ D_\mu \phi^a &= \partial_\mu \phi^a - e\epsilon^{abc} W_\mu^b \phi^c\,. \end{aligned} \tag{15.333}$$

*Singular in the sense of being not well defined in all space.

The minimal Lagrangian is then given by

$$\mathcal{L} = -\frac{1}{4}G^a_{\mu\nu}G^{a\mu\nu} + \frac{1}{2}D^\mu\phi^a D_\mu\phi^a - V(\phi)\,, \tag{15.334}$$

where,

$$V(\phi) = \frac{\lambda}{4}\left(\phi^a\phi^a - a^2\right)^2 . \tag{15.335}$$

The equations of motion following from this Lagrangian are

$$(D_\nu G^{\mu\nu})^a = -e\,\epsilon^{abc}\,\phi^b\,(D^\mu\phi)^c, \quad D^\mu D_\mu\phi^a = -\lambda\phi^a(\phi^2 - a^2)\,. \tag{15.336}$$

The gauge field strength also satisfies the Bianchi identity

$$D_\nu\,\widetilde{G}^{\mu\nu a} = 0\,. \tag{15.337}$$

Let us find the vacuum configurations in this theory. Introducing non-Abelian electric and magnetic fields, $G^{0i}_a = -\mathcal{E}^i_a$ and $G^{ij}_a = -\epsilon^{ij}{}_k\mathcal{B}^k_a$, the energy density is written as

$$\theta_{00} = \frac{1}{2}\left((\mathcal{E}^i_a)^2 + (\mathcal{B}^i_a)^2 + \left(D^0\phi_a)^2 +)D^i\phi_a)^2\right) + V(\phi)\,. \tag{15.338}$$

Note that $\theta_{00} \geq 0$, and it vanishes only if

$$G^{\mu\nu}_a = 0, \quad D_\mu\phi = 0, \quad V(\phi) = 0\,. \tag{15.339}$$

The first equation implies that in the vacuum, W^a_μ is pure gauge and the last two equations define the Higgs vacuum. The structure of the space of vacua is determined by $V(\phi) = 0$ which solves to $\phi^a = \phi^a_{vac}$ such that $|\phi_{vac}| = a$. The space of Higgs vacua is therefore a two-sphere (S^2) of radius a in field space. To formulate a perturbation theory, we have to choose one of these vacua and hence, break the gauge symmetry spontaneously The part of the symmetry which keeps this vacuum invariant, still survives and the corresponding unbroken generator is $\phi^c_{vac}T^c/a$. The gauge boson associated with this generator is $A_\mu = \phi^c_{vac}W^c_\mu/a$ and the electric charge operator for this surviving $U(1)$ is given by

$$Q = e\frac{\phi^c_{vac}T^c}{a}\,. \tag{15.340}$$

If the group is compact, this charge is quantized. The perturbative spectrum of the theory can be found by expanding ϕ^a around the chosen vacuum as

$$\phi^a = \phi^a_{vac} + \phi'^a\,.$$

A convenient choice is $\phi^c_{vac} = \delta^{c3}a$. The perturbative spectrum (which becomes manifest after choosing an appropriate unitary gauge) consists of a massive Higgs of spin zero with a square mass

$$\mathcal{M}^2_H = 2\lambda a^2, \tag{15.341}$$

a massless photon, corresponding to the $U(1)$ gauge boson A^3_μ, and two charged massive W-bosons, A^1_μ and A^2_μ, with square mass

$$\mathcal{M}^2_W = e^2a^2. \tag{15.342}$$

This mass spectrum is realistic as long as we are at weak coupling, $e^2 \sim \lambda\mathcal{L}1$. At strong coupling, nonperturbative effects could change significatively eqs. (15.341) and (15.342). But the fact that there is an unbroken subgroup of the gauge symmetry ensures that there is some massless gauge boson, which a long range interaction. This is the characteristic of the Coulomb phase.

15.4.6 The 't Hooft-Polyakov monopoles

Let us look for time-independent, finite energy solutions in the Georgi-Glashow model. Finiteness of energy requires that as $r \to \infty$, the energy density θ_{00} given by (15.338) must approach zero faster than $1/r^3$. This means that as $r \to \infty$, our solution must go over to a Higgs vacuum defined by (15.339). In the following, we will first assume that such a finite energy solution exists and show that it can have a monopole charge related to its soliton number which is, in turn, determined by the associated Higgs vacuum. This result is proven without having to deal with any particular solution explicitly. Next, we will describe the 't Hooft-Polyakov ansatz for explicitly constructing one such monopole solution, where we will also comment on the existence of Dyonic solutions. In the last two subsections we will derive the Bogomol'nyi bound and the Witten effect.

The Topological nature of the magnetic charge.

For convenience, in this subsection we will use the vector notation for the $SO(3)$ gauge group indices and not for the spatial indices.

Let $\vec{\phi}_{vac}$ denote the field $\vec{\phi}$ in a Higgs vacuum. It then satisfies the equations

$$\begin{aligned} \vec{\phi}_{vac} \cdot \vec{\phi}_{vac} &= a^2 \,, \\ \partial_\mu \vec{\phi}_{vac} - e\,\vec{W}_\mu \times \vec{\phi}_{vac} &= 0 \,, \end{aligned} \tag{15.343}$$

which can be solved for $\vec{W}_\mu$. The most general solution is given by

$$\vec{W}_\mu = \frac{1}{ea^2}\,\vec{\phi}_{vac} \times \partial_\mu \vec{\phi}_{vac} + \frac{1}{a}\vec{\phi}_{vac} A_\mu \,. \tag{15.344}$$

To see that this actually solves (15.343), note that $\partial_\mu \vec{\phi}_{vac} \cdot \vec{\phi}_{vac} = 0$, so that

$$\begin{aligned} &\frac{1}{ea^2}(\vec{\phi}_{vac} \times \partial_\mu \vec{\phi}_{vac}) \times \vec{\phi}_{vac} = \\ &\frac{1}{ea^2}\left(\partial_\mu \vec{\phi}_{vac} a^2 - \vec{\phi}_{vac}(\vec{\phi}_{vac} \cdot \partial_\mu \phi_{vac})\right) = \frac{1}{e}\partial_\mu \vec{\phi}_{vac} \,. \end{aligned} \tag{15.345}$$

The first term on the right-hand side of Eq. (15.344) is the particular solution, and $\vec{\phi}_{vac} A_\mu$ is the general solution to the homogeneous equation. Using this solution, we can now compute the field strength tensor $\vec{G}_{\mu\nu}$. The field strength $F_{\mu\nu}$ corresponding to the unbroken part of the gauge group can be identified as

$$\begin{aligned} F_{\mu\nu} = \frac{1}{a}\vec{\phi}_{vac} \cdot \vec{G}_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu \\ &+ \frac{1}{a^3 e}\vec{\phi}_{vac} \cdot (\partial_\mu \vec{\phi}_{vac} \times \partial_\nu \vec{\phi}_{vac}) \,. \end{aligned} \tag{15.346}$$

Using the equations of motion in the Higgs vacuum it follows that

$$\partial_\mu F^{\mu\nu} = 0 \,, \qquad \partial_\mu \widetilde{F}^{\mu\nu} = 0 \,.$$

This confirms that $F_{\mu\nu}$ is a valid $U(1)$ field strength tensor. The magnetic field is given by $B^i = -\frac{1}{2}\epsilon^{ijk} F_{jk}$. Let us now consider a static, finite energy solution and a surface Σ enclosing the core of the solution. We take Σ to be far enough so that, on it, the solution is already in the Higgs vacuum. We can now use the magnetic field in the Higgs vacuum to calculate the magnetic charge g_Σ associated with our solution:

$$\begin{aligned} g_\Sigma &= \int_\Sigma B^i ds^i \\ &= -\frac{1}{2ea^3}\int_\Sigma \epsilon_{ijk}\,\vec{\phi}_{vac} \cdot \left(\partial^j \vec{\phi}_{vac} \times \partial^k \vec{\phi}_{vac}\right) ds^i \,. \end{aligned} \tag{15.347}$$

It turns out that the expression on the right hand side is a topological quantity as we explain below: Since $\phi^2 = a$; the manifold of Higgs vacua ($\mathcal{M}_0$) has the topology of S^2. The field $\vec{\phi}_{vac}$ defines a map from Σ into $\mathcal{M}_0$. Since Σ is also an S^2, the map $\phi_{vac} : \Sigma \to \mathcal{M}_0$ is characterized by its homotopy group $\pi_2(S^2)$. In other words, ϕ_{vac} is characterized by an integer ν (the winding number) which counts the number of times it wraps Σ around $\mathcal{M}_0$. In terms of the map ϕ_{vac}, this integer is given by

$$\nu = \frac{1}{4\pi a^3} \int_\Sigma \frac{1}{2}\epsilon_{ijk}\vec{\phi}_{vac} \cdot \left(\partial^j \vec{\phi}_{vac} \times \partial^k \vec{\phi}_{vac}\right) ds^i \,. \tag{15.348}$$

Comparing this with the expression for magnetic charge, we get the important result

$$g_\Sigma = \frac{-4\pi\nu}{e} \,. \tag{15.349}$$

Hence, the winding number of the soliton determines its monopole charge. Note that the above equation differs from the Dirac quantization condition by a factor of 2. This is because the smallest electric charge which could exist in our model is $e/2$ for an spinorial representation of $SU(2)$, the universal covering group of $SO(3)$. Then, in this model $m_0 = 2$.

The 't Hooft-Polyakov ansatz.

Now we describe an ansatz proposed by 't Hooft [6] and Polyakov [7] for constructing a monopole solution in the Georgi-Glashow model. For a spherically symmetric, parity-invariant, static solution of finite energy, they proposed:

$$\begin{aligned} \phi^a &= \frac{x^a}{er^2} H(aer) \,, \\ W_i^a &= -\epsilon_{ij}^a \frac{x^j}{er^2} \left(1 - K(aer)\right) \,, \\ W_0^a &= 0 \,. \end{aligned} \tag{15.350}$$

For the non-trivial Higgs vacuum at $r \to \infty$, they chose $\phi^c_{vac} = ax^c/r = a\hat{x}^c$. Note that this maps an S^2 at spatial infinity onto the vacuum manifold with a unit winding number. The asymptotic behavior of the functions $H(aer)$ and $K(aer)$ are determined by the Higgs vacuum as $r \to \infty$ and regularity at $r = 0$. Explicitly, defining $\xi = aer$, we have: as $\xi \to \infty$, $H \sim \xi$, $K \to 0$ and as $\xi \to 0$, $H \sim \xi$, $(K-1) \sim \xi$. The mass of this solution can be parameterized as

$$\mathcal{M} = \frac{4\pi a}{e} f\left(\lambda/e^2\right).$$

For this ansatz, the equations of motion reduce to two coupled equations for K and H which have been solved exactly only in certain limits. For $r \to 0$, one gets $H \to ec_1r^2$ and $K = 1 + ec_2r^2$ which shows that the fields are non-singular at $r = 0$. For $r \to \infty$, we get $H \to \xi + c_3 exp(-a\sqrt{2\lambda}r)$ and $K \to c_4\xi exp(-\xi)$ which leads to $W_i^a \approx -\epsilon_{ij}^a x^j/er^2$. Once again, defining $F_{ij} = \phi^c G^c_{ij}/a$, the magnetic field turns out to be $B^i = -x^i/er^3$. The associated monopole charge is $g = -4\pi/e$, as expected from the unit winding number of the solution. It should be mentioned that 't Hooft's definition of the Abelian field strength tensor is slightly different but, at large distances, it reduces to the form given above.

Note that in the above monopole solution, the presence of the Dirac string is not obvious. To extract the Dirac string, we have to perform a singular gauge transformation on this solution which rotates the non-trivial Higgs vacuum $\phi^c_{vac} = a\hat{x}^c$ into the trivial vacuum $\phi^c_{vac} = a\delta^{c3}$. In the process,the gauge field develops a Dirac string singularity which now serves as the source of the magnetic charge [6].

The 't Hooft-Polyakov monopole carries one unit of magnetic charge and no electric charge. The Georgi-Glashow model also admits solutions which carry both magnetic as well as electric charges. An ansatz for constructing such a solution was proposed by Julia and Zee [8]. In this ansatz, ϕ^a and W_i^a have exactly the same form as in the 't Hooft-Polyakov ansatz, but W_0^a is no longer zero: $W_0^a = x^a J(aer)/er^2$. This serves as the source for the electric charge of the dyon. It turns out that the dyon electric charge depends of a continuous parameter and, at the classical level, does not satisfy the quantization condition. However, semiclassical arguments show that, in CP invariant theories, and at the quantum level, the dyon electric charge is quantized as $q = ne$. This can be easily understood if we recognize that a monopole is not invariant under a gauge transformation which is, of course, a symmetry of the equations of motion. To deal with the associated zero-mode properly, the gauge degree of freedom should be regarded as a collective coordinate. Upon quantization, this collective coordinate leads to the existence of electrically charged states for the monopole with discrete charges. In the presence of a CP violating term in the Lagrangian, the situation is more subtle as we will discuss later. In the next subsection, we describe a limit in which the equations of motion can be solved exactly for the 'tHooft-Polyakov and the Julia-Zee ansatz. This is the limit in which the soliton mass saturates the Bogomol'nyi bound.

The Bogomol'nyi bound and the BPS states.

In this subsection, we derive the Bogomol'nyi bound [9] on the mass of a dyon in term of its electric and magnetic charges which are the sources for $F^{\mu\nu} = \vec{\phi} \cdot \vec{G}^{\mu\nu}/a$. Using the Bianchi identity (15.337) and the first equation in (15.336), we can write the charges as

$$g \equiv \int_{S^2_\infty} B_i dS^i = \frac{1}{a} \int \mathcal{B}_i^a (D^i\phi)^a d^3x \,,$$
$$q \equiv \int_{S^2_\infty} E_i dS^i = \frac{1}{a} \int \mathcal{E}_i^a (D^i\phi)^a d^3x \,. \tag{15.351}$$

Now, in the center of mass frame, the dyon mass is given by

$$\mathcal{M} \equiv \int d^3x\theta_{00} = \int d^3x \left(\frac{1}{2} \left[(\mathcal{E}_k^a)^2 + (\mathcal{B}_k^a)^2 + (D_k\phi^a)^2 + (D_0\phi^a)^2 \right] + V(\phi) \right) \,, \tag{15.352}$$

where, $\theta_{\mu\nu}$ is the energy momentum tensor. Using (15.351) and some algebra we obtain

$$\mathcal{M} = \int d^3x \left(\frac{1}{2} \left[(\mathcal{E}_k^a - D_k\phi^a \sin\theta)^2 + (\mathcal{B}_k^a - D_k\phi^a\cos\theta)^2 + (D_0\phi^a)^2 \right] + V(\phi) \right) + a(q\sin\theta + g\cos\theta) \,, \tag{15.353}$$

where θ is an arbitrary angle. Since the terms in the first line are positive, we can write $\mathcal{M} \geq (q\sin\theta + g\cos\theta)$. This bound is maximized for $\tan\theta = q/g$. Thus we get the Bogomol'nyi bound on the dyon mass as

$$\mathcal{M} \geq a\sqrt{g^2 + q^2} \,. \tag{15.354}$$

For the 't Hooft-Polyakov solution, we have $q = 0$, and thus, $\mathcal{M} \geq a|g|$. But $|g| = 4\pi/e$ and $\mathcal{M}_W = ae = aq$, so that

$$\mathcal{M} \geq a\frac{4\pi}{e} = \frac{4\pi}{e^2}\mathcal{M}_W = \frac{4\pi}{q^2}\mathcal{M}_W = \frac{\nu}{\alpha}\mathcal{M}_W \,.$$

Here, α is the fine structure constant and $\nu = 1$ or $1/4$, depending on whether the electron charge is q or $q/2$. Since α is a small ($\sim 1/137$ for electromagnetism), the above relation implies that the monopole is much heavier than the W-bosons associated with the symmetry breaking.

From (15.353) it is clear that the bound is not saturated unless $\lambda \to 0$, so that $V(\phi) = 0$. This is the Bogomol'nyi-Prasad-Sommerfield (BPS) limit of the theory [9, 10]. Note that in this limit, $\phi^2_{vac} = a^2$ is no longer determined by the theory and, therefore, has to be imposed as a boundary condition on the Higgs field. Moreover, in this limit, the Higgs scalar becomes massless. Now, to saturate the bound we set

$$\begin{aligned} &D_0\phi^a = 0\,, \\ &\mathcal{E}^a_k = (D_k\phi)^a \sin\theta\,, \\ &\mathcal{B}^a_k = (D_k\phi)^a \cos\theta\,, \end{aligned} \tag{15.355}$$

where, $\tan\theta = q/g$. In the BPS limit, one can use the 't Hooft-Polyakov (or the Julia-Zee) ansatz either in (15.336), or in (15.355) to obtain the exact monopole (or dyon) solutions [9, 10]. These solutions automatically saturate the Bogomol'nyi bound and are referred to as the BPS states. Also, note that in the BPS limit, all the perturbative excitations of the theory saturate this bound and, therefore, belong to the BPS spectrum. As we will see later, BPS states appears in a very natural way in theories with $N = 2$ supersymmetry.

The θ parameter and the Witten effect.

In this section we will show that in the presence of a θ-term in the Lagrangian, the magnetic charge of a particle always contributes to its electric charge in the way given by formula (15.313) [11].

To study the effect of CP violation, we consider the Georgi-Glashow model with an additional θ-term as the only source of CP violation:

$$\begin{aligned} \mathcal{L} = &-\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + \frac{1}{2}(D_\mu\phi^a)^2 - \lambda(\phi^2 - a^2)^2 \\ &+ \frac{\theta e^2}{32\pi^2}F^a_{\mu\nu}\tilde{F}^{a\mu\nu}\,. \end{aligned} \tag{15.356}$$

Here, $\tilde{F}^{a\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}F^a_{\rho\sigma}$. The presence of the θ-term does not affect the equations of motion but changes the physics since the theory is no longer CP invariant. We want to construct the electric charge operator in this theory. The theory has an $SO(3)$ gauge symmetry but the electric charge is associated with an unbroken $U(1)$ which keeps the Higgs vacuum invariant. Hence, we define an operator N which implements a gauge rotation around the $\hat{\phi}$ direction with gauge parameter $\Lambda^a = \phi^a/a$. These transformations correspond to the electric charge. Under N, a vector v^a and the gauge fields A^a_μ transform as

$$\delta v^a = \frac{1}{a}\,\epsilon^{abc}\phi^b v^c\,, \quad \delta A^a_\mu = \frac{1}{ea}\,D_\mu\phi^a\,.$$

Clearly, ϕ^a is kept invariant. At large distances where $|\phi| = a$, the operator $e^{2\pi i N}$ is a 2π-rotation about $\hat{\phi}$ and therefore $\exp(2\pi i N) = 1$. Elsewhere, the rotation angle is $2\pi|\phi|/a$. However, by Gauss' law, if the gauge transformation is 1 at ∞, it leaves the physical states invariant. Thus, it is only the large distance behavior of the transformation which matters and the eigenvalues of N are quantized in integer units. Now, we use Noether's formula to compute N:

$$N = \int d^3x \left(\frac{\delta\mathcal{L}}{\delta\partial_0 A^a_i}\,\delta A^a_i + \frac{\delta\mathcal{L}}{\delta\partial_0\phi^a}\,\delta\phi^a\right)\,.$$

Since $\delta\vec{\phi} = 0$, only the gauge part (which also includes the θ-term) contributes:

$$\frac{\delta}{\delta\partial_0 A_i^a}\,)F^a_{\mu\nu}F^{a\mu\nu}) = 4F^{aoi} = -4\mathcal{E}^{ai}\,,$$
$$\frac{\delta}{\delta\partial_0 A_i^a}\left(\tilde{F}^a_{\mu\nu}F^{a\mu\nu}\right) = 2\epsilon^{ijk}F^a_{jk} = -4\mathcal{B}^{ai}\,.$$

Thus,

$$\begin{aligned} N &= \frac{1}{ae}\int d^3x D_i\vec{\phi}\cdot\vec{\mathcal{E}}^i - \frac{\theta e}{8\pi^2 a}\int d^3x\, D_i\vec{\phi}\cdot\vec{\mathcal{B}}^i \\ &= \frac{1}{e}Q_e - \frac{\theta e}{8\pi^2}Q_m\,, \end{aligned}$$

where, we have used (15.351). Here, Q_e and Q_m are the electric and magnetic charge operators with eigenvalues q and g, respectively, and N is quantized in integer units. This leads to the following formula for the electric charge:

$$q = ne + \frac{\theta e^2}{8\pi^2}g\,.$$

For the 't Hooft-Polyakov monopole, $n = 1$, $g = -4\pi/e$, and therefore, $q = e(1 - \theta/2\pi)$. For a general dyonic solution we get

$$g = \frac{4\pi}{e}n_m, \qquad q = n_e e + \frac{\theta e}{2\pi}n_m\,. \tag{15.357}$$

and we recover (15.313) and (15.314) for $q_0 = e$. In the presence of a θ-term, a magnetic monopole always carries an electric charge which is not an integral multiple of some basic unit. In section III we introduced the charge lattice of periods e and $e\tau$. In this parameterization, the Bogomol'nyi bound (15.354) takes the form

$$\mathcal{M} \geq \sqrt{2}|ae(n_e + n_m\tau)|\,. \tag{15.358}$$

Notice that for a BPS state, equation (15.358) implies that its mass is proportional to the distance of its lattice point from the origin.

15.4.7 The Confining phase.

The Abelian projection.

In non-Abelian gauge theories, gauge fixing is a subject full of interesting surprises (ghosts, phantom solitons,...) which often obscure the physical content of the theory [12].

't Hooft gave a qualitative program to overcome these difficulties and provided a scenario that explains confinement in a gauge theory. The idea is to perform the gauge fixing procedure in two steps. In the first one a unitary gauge is chosen for the non-Abelian degrees of freedom. It reduces the non-Abelian gauge symmetry to the maximal Abelian subgroup of the gauge group. Here one gets particle gauge singularities *. This procedure is called the Abelian projection [12]. In this way, the dynamics of the Yang-Mils theory will be reduced to an Abelian gauge theory with certain additional degrees of freedom.

We need a field that transforms without derivatives under gauge transformations. An example is a real field, X in the adjoint representation of $SU(N)$,

$$X \to \Omega X \Omega^{-1}. \tag{15.359}$$

*We will discuss the physical meaning of them later on.

Such a field can always be found; take for instance $X^a = G^a_{12}$. We will use the field X to implement the unitary gauge condition which will carry us to the Abelian projection of the $SU(N)$ gauge group. The gauge is fixed by requiring that X be diagonal:

$$X = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{pmatrix}. \tag{15.360}$$

The eigenvalues of the matrix X are gauge invariant. Generically they are all different, and the gauge condition (15.360) leaves an Abelian $U(1)^{N-1}$ gauge symmetry. It corresponds to the subgroup generated by the gauge transformations

$$\Omega = \begin{pmatrix} e^{i\omega_1} & & 0 \\ & \ddots & \\ 0 & & e^{i\omega_N} \end{pmatrix}, \quad \sum_{i=1}^{N} \omega_i = 0\,. \tag{15.361}$$

There is also a discrete subgroup of transformations which still leave X in diagonal form. It is the Weyl group of $SU(N)$, which corresponds to permutations of the eigenvalues λ_i. We also fix the Weyl group with the convention $\lambda_1 > \lambda_2 > \cdots \lambda_N$.

At this stage, we have an Abelian $U(1)$ gauge theory with $N-1$ photons, $N(N-1)$ charged vector particles and some additional degrees of freedom that will appear presently.

The nature of the gauge singularities.

So far we assumed that the eigenvalues λ_i coincide nowhere. But there are some gauge field configurations that produce two consecutive eigenvalues to coincide at some spacetime points

$$\lambda_i = \lambda_{i+1} = \lambda, \quad \text{for certain } i. \tag{15.362}$$

These spacetime points are 'singular' points of the Abelian projection. The $SU(2)$ gauge subgroup corresponding to the 2×2 block matrix with coinciding eigenvalues leaves invariant the gauge-fixing condition (15.360).

Let us consider the vicinity of such a point. Prior to the complete gauge-fixing we may take X to be

$$X = \begin{pmatrix} D_1 & 0 & 0 & 0 \\ 0 & \lambda + \epsilon_3 & \epsilon_1 - i\epsilon_2 & 0 \\ 0 & \epsilon_1 + i\epsilon_2 & \lambda - \epsilon_3 & 0 \\ 0 & 0 & 0 & D_2 \end{pmatrix}, \tag{15.363}$$

where D_1 and D_2 may safely be considered to be diagonalized because the other eigenvalues do not coincide. With respect to that $SU(2)$ subgroup of $SU(N)$ that corresponds to rotations among the ith and $i+1$st components, the three fields $\epsilon_a(x)$ form an isovector. We may write the central block as

$$\lambda I_2 + \epsilon_a \sigma^a, \tag{15.364}$$

where σ^a are the Pauli matrices.

Consider static field configurations. The points of space where the two eigenvalues coincide correspond to the points $\mathbf{x}_0$ that satisfy

$$\epsilon^a(\mathbf{x}_0) = 0\,. \tag{15.365}$$

These three equations define a single space point, and then the singularity is particle-like. Which is its physical interpretation?.

By analyticity we have that $\epsilon^a \sim (x - x_0)^a$, and our gauge condition corresponds to rotating the isovector ϵ^a such that

$$\epsilon = \begin{pmatrix} 0 \\ 0 \\ |\epsilon_3| \end{pmatrix}. \quad (15.366)$$

From the previous sections, we know that the zero-point of ϵ^a at $\mathbf{x}_0$ behaves as a magnetic charge with respect to the remaining $U(1) \subset SU(2)$ rotations. We realize that those gauge field configurations that produce such a gauge 'singularities' correspond to magnetic monopoles.

The non-Abelian $SU(N)$ gauge theory is topologically such that it can be cast into a $U(1)^{N-1}$ Abelian gauge theory, which will feature not only electrically charged particles but also magnetic monopoles.

The phases of the Yang-Mills vacuum.

We can now give a qualitative description of the possible phases of the Yang-Mills vacuum. It is only the dynamics which, as a function of the microscopic bare parameters, determines in which phase the Yang-Mills vacuum is actually realized.

Classically, the Yang-Mills Lagrangian is scale invariant. One can write down field configurations with magnetic charge and arbitrarily low energy. But quantum corrections are likely to violate their masslessness. If dynamics simply chooses to give a positive mass to the monopoles, we are in a Higgs or Coulomb phase. We must look for the magnetic vortex tubes to figure out if we are in a Higgs phase. It will be a signal that the ordinary Higgs mechanism has taken place in the Abelian gauge formulation of the Yang-Mills theory. The role of the dynamically generated Higgs field could be done by some scalar composite operator charged respect the $U(1)^{N-1}$ gauge symmetries. There is also the possibility that no Higgs phenomenon occurs at all in the Abelian sector, or that some $U(1)$ gauge symmetries are not spontaneously broken. In this case we are in the Coulomb phase, with some massless photons, or in a mixed Coulomb-Higgs phase.

There is a third possibility however. Maybe the quantum corrections give a formally negative mass squared for the monopole: a magnetically charged object condenses. We apply an 'electric-magnetic dual transformation' to write an effective Lagrangian which encodes the relevant magnetic degrees of freedom in the infrared limit. In such effective Lagrangian, the Higgs mechanism takes place in terms of dual variables. We are in a dual Higgs phase. We have electric flux tubes with finite energy per unit of length. There is a confining potential between electrically charged objects, like quarks.

In 1994, Seiberg and Witten gave a quantitative proof that such dynamical mechanism of color confinement takes place in $N = 2$ super-QCD (SQCD) broken to $N = 1$ [13], giving a non-trivial realization of 't Hooft scenario. When $N = 2$ SQCD is softly broken to $N = 0$ the same mechanism of confinement persists [14, 15].

Oblique confinement.

For simplicity let us consider an $SU(2)$ gauge group. We have seen that for a non-zero CP violating parameter θ, the physical electric charge of a particle with electric (resp. magnetic) number n_e (resp. n_m) is:

$$q = (n_e + \frac{\theta}{2\pi} n_m)e. \quad (15.367)$$

Dyons with large electric charges may have larger self-energies contributing positively to their mass squared. If the state (n_e, n_m) condenses at $\theta \simeq 0$, it is likely that the state $(n_e - 1, n_m)$ condenses at $\theta \simeq 2\pi$. It suggests that there is a phase transition around $\theta \simeq \pi$. Such first order phase transitions has been observed in softly broken $N = 2$ SQCD to $N = 0$ [16].

't Hooft proposed a new condensation mode at $\theta \simeq \pi$ [12]. He imagined the possibility that a bound state of the dyons (n_e, n_m) and $(n_e - 1, n_m)$, with zero electric charge at $\theta = \pi$, could be formed. Its smaller electric charge could favor its condensation, leading to what he called an oblique confinement mode. These oblique modes have also been observed in softly broken $N = 2$ SQCD with matter [14, 15].

15.4.8 The Higgs/confining phase.

In the previous section we have characterized the confining phase as the dual of the Higgs phase, *i.e.*, the physical states are gauge singlets made by the electric degrees of freedom bound by stable electric flux tubes. A good gauge invariant order parameter measuring such behavior is the Wilson loop [17]:

$$W(C) = \mathrm{Tr}\,\exp\left(ig\oint_C dx^\mu A_\mu\right). \tag{15.368}$$

For $SU(N)$ Yang-Mills in the confining phase, for contours C, the Wilson loop obeys the area law,

$$\langle W(C)\rangle \sim \exp(-\sigma\cdot(\text{area})), \tag{15.369}$$

with σ the string tension of the electric flux tube.

But dynamical matter fields in the fundamental representation immediately create a problem in identifying the confining phase of the theory through the Wilson loop. The criterion used for confinement in the pure gauge theory, the energy between static sources, no longer works. Even if the energy starts increasing as the sources separate, it eventually becomes favorable to produce a particle-antiparticle pair out of the vacuum. This pair shields the gauge charge of the sources, and the energy stops growing. So even in a theory that 'looks' very confining our signal fails, and the perimeter law replaces (15.369),

$$\langle W(C)\rangle =\sim \exp(-\Lambda\cdot(\text{perimeter})) \tag{15.370}$$

If some scalar field is in the fundamental representation of the gauge group, there is no distinction at all between the confinement phases and the Higgs phase. Using the scalar field in the fundamental representation one can build gauge invariant interpolating operators for all possible physical states. As the vacuum expectation value of the Higgs field in the fundamental representation continuously changes from large values to smaller ones, the spectrum of all physical states, and all other measurable quantities, changes smoothly [18]. There is no gauge invariant operator which can distinguish between the Higgs or confining phases. We are in a Higgs/confining phase.

In supersymmetric gauge theories, it is common to have scalar fields in the fundamental representation of the gauge group, the scalar quarks. In such situation, when the theory is not in the Coulomb phase, we will see that the theory is presented in a Higgs/confining phase. We could take the phase description which is more appropriate for the theory. For instance, if the theory is in the weak coupling region, it is better to realize it in the Higgs phase; if the theory in the strong coupling region, it is better to think it in a confining phase.

15.4.9 Supersymmetry

The supersymmetry algebra and its massless representations.

The $N = 1$ supersymmetry algebra is written as [19]

$$\begin{aligned}&\{Q_\alpha, \overline{Q}_{\dot\alpha}\} = 2\sigma^\mu_{\alpha\dot\alpha}P_\mu\\ &\{Q_\alpha, Q_\beta\} = 0\,,\ \{\overline{Q}_{\dot\alpha}\,,\,\overline{Q}_{\dot\beta}\} = 0\,.\end{aligned} \tag{15.371}$$

Here, Q and $\overline{Q}$ are the supersymmetry generators and transform as spin 1/2 operators, $\alpha, \dot{\alpha} = 1, 2$. Moreover, the supersymmetry generators commute with the momentum operator P_μ and hence, with P^2. Therefore, all states in a given representation of the algebra have the same mass. For a theory to be supersymmetric, it is necessary that its particle content form a representation of the above algebra. The irreducible representations of (15.371) can be obtained using Wigner's method.

For massless states, we can always go to a frame where $P^\mu = E(1, 0, 0, 1)$. Then the supersymmetry algebra becomes

$$\{Q_\alpha, \overline{Q}_{\dot{\alpha}}\} = \left) \begin{matrix} 0 & 0 \\ 0 & 4E \end{matrix} \right) .$$

In a unitary theory the norm of a state is always positive. Since Q_α and $\overline{Q}_{\dot{\alpha}}$ are conjugate to each other, and $\{Q_1, \overline{Q}_{\dot{1}}\} = 0$, it follows that $Q_1|phys >= \overline{Q}_{\dot{1}}|phys >= 0$. As for the other generators, it is convenient to re-scale them as

$$a = \frac{1}{2\sqrt{E}} Q_2 , \qquad a^\dagger = \frac{1}{2\sqrt{E}} \overline{Q}_{\dot{2}} .$$

Then, the supersymmetry algebra takes the form

$$\{a, a^\dagger\} = 1 , \quad \{a, a\} = 0 , \quad \{a^\dagger, a^\dagger\} = 0 .$$

This is a Clifford algebra with 2 fermionic generators and has a 2-dimensional representation. From the point of view of the angular momentum algebra, a is a rising operator and $a^\dagger$ is a lowering operator for the helicity of massless states. We choose the vacuum such that $J_3|\Omega_\lambda >= \lambda|\Omega_\lambda >$ and $a|\Omega_\lambda >= 0$. Then

$$J_3(a^\dagger|\Omega_\lambda >) = (\lambda - \frac{1}{2})(a^\dagger|\Omega_\lambda >). \tag{15.372}$$

The irreducible representations are not necessarily CPT invariant. Therefore, if we want to assign physical states to these representations, we have to supplement them with their CPT conjugates $| - \lambda >_{CPT}$. If a representation is CPT self-conjugate, it is left unchanged. Thus, from a Clifford vacuum with helicity $\lambda = 1/2$ we obtain the $N = 1$ supermultiplet:

$$\begin{pmatrix} \{\ |1/2 >, \ | - 1/2 >_{CPT}\ \} \\ \{\ |0 >, \qquad |0 >_{CPT}\ \} \end{pmatrix} \tag{15.373}$$

which contains a Weyl spinor ψ and a complex scalar ϕ. It is called the scalar multiplet.

The other relevant representation of a renormalizable quantum field theory is the vector multiplet. It is constructed from a Clifford vacuum with helicity $\lambda = 1$:

$$\begin{pmatrix} \{\ |1 >, \quad | - 1 >_{CPT}\ \} \\ \{\ |1/2 >, \ | - 1/2 >_{CPT}\ \} \end{pmatrix} . \tag{15.374}$$

It contains a vector A_μ and a Weyl spinor λ.

Superspace and superfields.

To make supersymmetry linearly realized it is convenient to use the superspace formalism and superfields [20]. Superspace is obtained by adding four spinor degrees of freedom $\theta^\alpha, \overline{\theta}_{\dot{\alpha}}$ to the spacetime coordinates x^μ. Under the supersymmetry transformations implemented by the operator $\xi^\alpha Q_\alpha + \overline{\xi}_{\dot\alpha}\overline{Q}^{\dot\alpha}$ with transformation parameters ξ and $\overline{\xi}$, the superspace coordinates transform as

$$\begin{aligned} x^\mu &\to x'^\mu = x^\mu + i\theta\sigma^\mu\overline{\xi} - i\xi\sigma^\mu\overline{\theta}\,, \\ \theta &\to \theta' = \theta + \xi\,, \\ \overline{\theta} &\to \overline{\theta}' = \overline{\theta} + \overline{\xi}\,. \end{aligned} \tag{15.375}$$

These transformations can easily be obtained by the following representation of the supercharges acting on (x, θ):

$$\begin{aligned} Q_\alpha &= \frac{\partial}{\partial\theta^\alpha} - i\sigma^\mu_{\alpha\dot\alpha}\overline{\theta}^{\dot\alpha}\,\partial_\mu\,, \\ \overline{Q}_{\dot\alpha} &= -\frac{\partial}{\partial\overline{\theta}^{\dot\alpha}} + i\theta^\alpha\sigma^\mu_{\alpha\dot\alpha}\,\partial_\mu\,. \end{aligned} \tag{15.376}$$

These satisfy $\{Q_\alpha, \overline{Q}_{\dot\alpha}\} = 2i\sigma^\mu_{\alpha\dot\alpha}\,\partial_\mu$. Moreover, using the chain rule, it is easy to see that $\partial/\partial x^\mu$ is invariant under (15.375) but not $\partial/\partial\theta$ and $\partial/\partial\overline{\theta}$. Therefore, we introduce the super-covariant derivatives

$$\begin{aligned} D_\alpha &= \frac{\partial}{\partial\theta^\alpha} + i\sigma^\mu_{\alpha\dot\alpha}\,\partial_\mu\,, \\ \overline{D}_{\dot\alpha} &= -\frac{\partial}{\partial\overline{\theta}^{\dot\alpha}} - i\sigma^\mu_{\alpha\dot\alpha}\theta^\alpha\,\partial_\mu\,. \end{aligned} \tag{15.377}$$

They satisfy $\{D_\alpha, \overline{D}_{\dot\alpha}\} = -2i\sigma^\mu_{\alpha\dot\alpha}\,\partial_\mu$ and anti-commute with Q and $\overline{Q}$.

The quantum fields transform as components of a superfield defined on superspace, $F(x, \theta, \overline{\theta})$. Since the θ-variables are anti-commuting, the Taylor expansion of $F(x, \theta, \overline{\theta})$ in $(\theta, \overline{\theta})$ is finite, indicating that the supersymmetry representations are finite dimensional. The coefficients of the expansion are the component fields.

To have irreducible representations we must impose supersymmetric invariant constraints on the superfields. The scalar multiplet (15.373) is represented by a chiral scalar superfield, Φ, satisfying the chiral constraint

$$\overline{D}_{\dot\alpha}\Phi = 0\,. \tag{15.378}$$

Note that for $y^\mu = x^\mu + i\theta\sigma^\mu\overline{\theta}$, we have $\overline{D}_{\dot\alpha}y^\mu = 0, \quad \overline{D}_{\dot\alpha}\theta^\beta = 0$. Therefore, any function of (y, θ) is a chiral superfield. It can be shown that this also is a necessary condition. Hence, any chiral superfield can be expanded as

$$\Phi(y, \theta) = \phi(y) + \sqrt{2}\theta\psi(y) + \theta\theta F(y)\,. \tag{15.379}$$

Here, ψ and ϕ are the fermionic and scalar components respectively and F is an auxiliary field linear and homogeneous. Similarly, an anti-chiral superfield is defined by $D_\alpha\Phi^\dagger = 0$ and can be expanded as

$$\Phi^\dagger(y^\dagger, \overline{\theta}) = \phi^\dagger(y^\dagger) + \sqrt{2}\overline{\theta\psi}(y^\dagger) + \overline{\theta\theta}F^\dagger(y^\dagger)\,, \tag{15.380}$$

where, $y^{\mu\dagger} = x^\mu - i\theta\sigma^\mu\overline{\theta}$.

The vector multiplet (15.374) is represented off-shell by a real scalar superfield

$$V = V^\dagger. \tag{15.381}$$

In local quantum field theories, spin one massless particles carry gauge symmetries [21]. These symmetries commute with the supersymmetry transformations. For a vector superfield, many of its component fields can be gauged away using the Abelian gauge transformation

$V \to V + \Lambda + \Lambda^\dagger$, where Λ ($\Lambda^\dagger$) are chiral (anti-chiral) superfields. In the Wess-Zumino gauge [19], it becomes

$$V = -\theta\sigma^\mu\overline{\theta}A_\mu + i\theta^2\overline{\theta\lambda} - i\overline{\theta}^2\theta\lambda + \frac{1}{2}\theta^2\overline{\theta}^2 D\,.$$

In this gauge, $V^2 = \frac{1}{2}A_\mu A^\mu\theta^2\overline{\theta}^2$ and $V^3 = 0$. The Wess-Zumino gauge breaks supersymmetry, but not the gauge symmetry of the Abelian gauge field A_μ. The Abelian superfield gauge field strength is defined by

$$W_\alpha = -\frac{1}{4}\overline{D}^2 D_\alpha V\,, \qquad \overline{W}_{\dot\alpha} = -\frac{1}{4}D^2\overline{D}_{\dot\alpha}V\,.$$

It can be verified that W_α is a chiral superfield. Since it is gauge invariant, it can be computed in the Wess-Zumino gauge,

$$\begin{aligned} W_\alpha &= -i\lambda_\alpha(y) + \theta_\alpha D - \frac{i}{2}(\sigma^\mu\overline{\sigma}^\nu\theta)_\alpha\,F_{\mu\nu} \\ &\quad + \theta^2(\sigma^\mu\partial_\mu\overline{\lambda})_\alpha\,, \end{aligned} \tag{15.382}$$

where, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

In the non-Abelian case, V belongs to the adjoint representation of the gauge group: $V = V_A T^A$, where, $T^{A\dagger} = T^A$. The gauge transformations are now implemented by

$$e^{-2V} \to e^{-i\Lambda^\dagger}e^{-2V}e^{i\Lambda}\,,$$

where $\Lambda = \Lambda_A T^A$ is a chiral superfield. The non-Abelian gauge field strength is defined by

$$W_\alpha = \frac{1}{8}\overline{D}^2 e^{2V}D_\alpha e^{-2V}$$

and transforms as

$$W_\alpha \to W'_\alpha = e^{-i\Lambda}W_\alpha e^{i\Lambda}\,.$$

In components, in the WZ gauge it takes the form

$$\begin{aligned} W^a_\alpha &= -i\lambda^a_\alpha + \theta_\alpha D^a - \frac{i}{2}(\sigma^\mu\overline{\sigma}^\nu\theta)_\alpha F^a_{\mu\nu} \\ &\quad + \theta^2\sigma^\mu D_\mu\overline{\lambda}^a\,, \end{aligned} \tag{15.383}$$

where,

$$\begin{aligned} F^a_{\mu\nu} &= \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + f^{abc}A^b_\mu A^c_\nu\,, \\ D_\mu\overline{\lambda}^a &= \partial_\mu\overline{\lambda}^a + f^{abc}A^b_\mu\overline{\lambda}^c\,. \end{aligned}$$

Now we are ready to construct supersymmetric Lagrangians in terms of superfields.

Supersymmetric Lagrangians.

Clearly, any function of superfields is, by itself, a superfield. Under supersymmetry, the superfield transforms as $\delta F = (\xi Q + \overline{\xi Q})F$, from which the transformation of the component fields can be obtained. Note that the coefficient of the $\theta^2\overline{\theta}^2$ component is the field component of highest dimension in the multiplet. Then, its variation under supersymmetry is always a total derivative of other components. Thus, ignoring surface terms, the spacetime integral of this component is invariant under supersymmetry. This tells us that a supersymmetric Lagrangian density may be constructed as the highest dimension component of an appropriate superfield.

Let us first consider the product of a chiral and an anti-chiral superfield $\Phi^\dagger\Phi$. This is a general superfield and its highest component can be computed using (15.379) as

$$\begin{aligned}\Phi^\dagger\Phi\,|_{\theta^2\overline{\theta}^2} &= -\frac{1}{4}\,\phi^\dagger\Box\phi - \frac{1}{4}\Box\phi^\dagger\phi + \frac{1}{2}\,\partial_\mu\phi^\dagger\partial^\mu\phi \\ &\quad - \frac{i}{2}\,\psi\sigma^\mu\partial_\mu\overline{\psi} + \frac{i}{2}\partial_\mu\psi\sigma^\mu\overline{\psi} + F^\dagger F\,. \end{aligned} \tag{15.384}$$

Dropping some total derivatives we get the free field Lagrangian for a massless scalar and a massless fermion with an auxiliary field.

The product of chiral superfields is a chiral superfield. In general, any arbitrary function of chiral superfields is a chiral superfield:

$$\begin{aligned}\mathcal{W}(\Phi_i) &= \mathcal{W}(\phi_i + \sqrt{2}\theta\psi_i + \theta\theta F_i) \\ &= \mathcal{W}(\phi_i) + \frac{\partial\mathcal{W}}{\partial\phi_i}\sqrt{2}\theta\psi_i \\ &\quad + \theta\theta\left(\frac{\partial\mathcal{W}}{\partial\phi_i}F_i - \frac{1}{2}\,\frac{\partial^2\mathcal{W}}{\partial\phi_i\phi_j}\,\psi_i\psi_j\right)\,. \end{aligned} \tag{15.385}$$

$\mathcal{W}$ is referred to as the superpotential. Moreover, the space of the chiral fields Φ may have a non-trivial metric g^{ij} in which case the scalar kinetic term, for example, takes the form $g^{ij}\partial_\mu\phi_i^\dagger\partial^\mu\phi_j$, with appropriate modifications for other terms. In such cases, the free field Lagrangian above has to be replaced by a non-linear σ-model [22]. Thus, the most general $N = 1$ supersymmetric Lagrangian for the scalar multiplet is given by

$$\mathcal{L} = \int d^4\theta\, K(\Phi,\Phi^\dagger) + \int d^2\theta\mathcal{W}(\Phi) + \int d^2\overline{\theta}\overline{\mathcal{W}}(\Phi^\dagger)\,.$$

Note that the θ-integrals pick up the highest component of the superfield and in our conventions, $\int d^2\theta\;\theta^2 = 1$ and $\int d^2\overline{\theta}\;\overline{\theta}^2 = 1$. In terms of the non-holomorphic function $K(\phi,\phi^\dagger)$, the metric in field space is given by $g^{ij} = \partial^2 K/\partial\phi_i\partial\phi_j^\dagger$, *i.e.*, the target space for chiral superfields is always a Kähler space. For this reason, the function $K(\Phi,\Phi^\dagger)$ is referred to as the Kähler potential.

Remember that the super-field strength W_α is a chiral superfield spinor. Using the normalization $\mathrm{Tr}(T^aT^b) = \frac{1}{2}\delta^{ab}$, we have that

$$\begin{aligned}\mathrm{Tr}(W^\alpha W_\alpha)\,|_{\theta\theta} &= -i\lambda^a\sigma^\mu D_\mu\overline{\lambda}^a + \frac{1}{2}D^aD^a \\ &\quad - \frac{1}{4}F^{a\mu\nu}F^a_{\mu\nu} + \frac{i}{8}\epsilon^{\mu\nu\rho\sigma}F^a_{\mu\nu}F^a_{\rho\sigma}\,. \end{aligned} \tag{15.386}$$

The first three terms are real and the last one is pure imaginary. It means that we can include the gauge coupling constant and the θ parameter in the Lagrangian in a compact form

$$\begin{aligned}\mathcal{L} &= \frac{1}{4\pi}\mathrm{Im}\left(\tau\,\mathrm{Tr}\int d^2\theta\, W^\alpha W_\alpha\right) \\ &= -\frac{1}{4g^2}F^a_{\mu\nu}F^{a\mu\nu} + \frac{\theta}{32\pi^2}F^a_{\mu\nu}\tilde{F}^{a\mu\nu} \\ &\quad + \frac{1}{g^2}(\frac{1}{2}D^aD^a - i\lambda^a\sigma^\mu D_\mu\overline{\lambda}^a)\,, \end{aligned} \tag{15.387}$$

where, $\tau = \theta/2\pi + 4\pi i/g^2$.

We now include matter fields by the introduction of the chiral superfield Φ in a given representation of the gauge group in which the generators are the matrices T^a_{ij}. The kinetic energy term $\Phi^\dagger\Phi$ is invariant under global gauge transformations $\Phi' = e^{-i\Lambda}\Phi$. In the local case, to insure that Φ' remains a chiral superfield, Λ has to be a chiral superfield. The supersymmetric gauge invariant kinetic energy term is then given by $\Phi^\dagger e^{-2V}\Phi$. We are now in a position to write down the full N=1 supersymmetric gauge invariant Lagrangian as

$$\begin{aligned}\mathcal{L} = \frac{1}{8\pi}\,\mathrm{Im}\left(\tau \mathrm{Tr}\int d^2\theta\, W^\alpha W_\alpha\right) \\ + \int d^2\theta d^2\overline{\theta}\,(\Phi^\dagger e^{-2V}\Phi) + \int d^2\theta\,\mathcal{W} + \int d^2\overline{\theta}\,\overline{\mathcal{W}}\,.\end{aligned} \tag{15.388}$$

Note that since each term is separately invariant, the relative normalization between the scalar part and the Yang-Mills part is not fixed by $N = 1$ supersymmetry. In fact, under loop effects, by virtue of the perturbative non-renormalization theorem [23], only the term with the complete superspace integral $\int d^2\theta d^2\overline{\theta}$ gets an overall renormalization factor $Z(\mu, g(\mu))$, with μ the renormalization scale and $g(\mu)$ the renormalized gauge coupling constant. Observe the unique dependence on $Re(\tau)$ in Z, breaking the holomorphic τ-dependence of the Lagrangian $\mathcal{L}$. But quantities as the superpotential $\mathcal{W}$ are renormalization group invariant under perturbation theory [23] (we will see dynamically generated superpotentials by nonperturbative effects).

In terms of component fields, the Lagrangian (15.388) becomes

$$\begin{aligned}\mathcal{L} &= -\frac{1}{4g^2}F^a_{\mu\nu}F^{a\mu\nu} + \frac{\theta}{32\pi^2}F^a_{\mu\nu}\widetilde{F}^{a\mu\nu} \\ &- \frac{i}{g^2}\,\lambda^a\sigma^\mu D_\mu\overline{\lambda}^a + \frac{1}{2g^2}D^aD^a \\ &+ (\partial_\mu\phi - iA^a_\mu T^a\phi)^\dagger(\partial^\mu\phi - iA^{a\mu}T^a\phi) - D^a\phi^\dagger T^a\phi \\ &- i\,\overline{\psi}\overline{\sigma}^\mu(\partial_\mu\psi - iA^a_\mu T^a\psi) + F^\dagger F \\ &+ \left(-i\sqrt{2}\,\phi^\dagger T^a\lambda^a\psi + \frac{\partial\mathcal{W}}{\partial\phi}\,F - \frac{1}{2}\frac{\partial^2\mathcal{W}}{\partial\phi\partial\phi}\,\psi\psi + h.c.\right)\,.\end{aligned} \tag{15.389}$$

Here, $\mathcal{W}$ denotes the scalar component of the superpotential. The auxiliary fields F and D^a can be eliminated by using their equations of motion:

$$F = \frac{\partial\mathcal{W}}{\partial\phi}\,, \tag{15.390}$$

$$D^a = g^2(\phi^\dagger T^a\phi)\,. \tag{15.391}$$

The terms involving these fields, thus, give rise to the scalar potential

$$V = |F|^2 + \frac{1}{2g^2}D^aD^a\,. \tag{15.392}$$

Using the supersymmetry algebra (15.371) it is not difficult to see that the hamiltonian $P^0 = H$ is a positive semi-definite operator, $\langle H\rangle \geq 0$, and that the ground state has zero energy if and only if it is supersymmetric invariant. At the level of local fields, the equation (15.392) means that the supersymmetric ground state configuration is such that

$$F = D^a = 0\,. \tag{15.393}$$

***R*-symmetry.**

The supercharges Q_α and $\overline{Q}_{\dot\alpha}$ are complex spinors. In the supersymmetry algebra (15.371) there is a $U(1)$ symmetry associated to the phase of the supercharges:

$$\begin{aligned}Q \to Q' = e^{i\beta}Q \\ \overline{Q} \to \overline{Q}' = e^{-i\beta}\overline{Q}.\end{aligned} \tag{15.394}$$

This symmetry is called the R-symmetry. It plays an important role in the study of supersymmetric gauge theories.

In terms of superspace, the R-symmetry is introduced through the superfield generator $(\theta Q + \overline{\theta Q})$. Then, it rotates the phase of the superspace components θ and $\overline{\theta}$ in the opposite way as Q and $\overline{Q}$. It gives different R-charges for the component fields of a superfield. Consider that the chiral superfield Φ has R-charge n,

$$\Phi(x,\theta) \to \Phi'(x,\theta) = e^{in\beta}\Phi(x, e^{-i\beta}\theta)\,. \tag{15.395}$$

In terms of its component fields we have that:

$$\begin{aligned} \phi &\to \phi' = e^{in\beta}\phi\,,\\ \psi &\to \psi' = e^{i(n-1)\beta}\psi\,,\\ F &\to F' = e^{i(n-2)\beta}F\,. \end{aligned}$$

Since $d^2(e^{-i\beta}\theta) = e^{2i\beta}d^2\theta$, we derive that the superpotential has R-charge two,

$$\mathcal{W}(\Phi) \to \mathcal{W}(\Phi',\theta) = e^{2i\beta}\mathcal{W}(\Phi, e^{-i\beta}\theta)\,, \tag{15.396}$$

and that the Kähler potential is R-neutral.

15.4.10 The uses of supersymmetry.

Flat directions and super-Higgs mechanism

We have seen that the fields configuration of the supersymmetric ground state are those corresponding to zero energy. To find them we solve (15.393). Consider a supersymmetric gauge theory with gauge group G, and matter superfields Φ_i in the representation $R(f)$ of G. The classical equations of motion of the D^a ($a = 1, ..., \dim G$) auxiliary fields give

$$D^a = \sum_f \phi_f^\dagger T_f^a \phi_f. \tag{15.397}$$

The solutions of $D^a = 0$ usually lead to the concept of flat directions. They play an important role in the analysis of SUSY theories. These flat directions may be lifted by F-terms in the Lagrangian, as for instance mass terms.

As an illustrative example of flat directions and some of its consequences, consider the $SU(2)$ gauge group, one chiral superfield Q in the fundamental representation of $SU(2)$ and another chiral superfield $\tilde{Q}$ in the anti-fundamental representation of $SU(2)$. This is supersymmetric QCD (SQCD) with one massless flavor. In this particular case, the equation (15.397) becomes

$$D^a = q^\dagger \sigma^a q - \tilde{q}\sigma^a \tilde{q}^\dagger. \tag{15.398}$$

The equations $D^a = 0$ have the general solution (up to gauge and global symmetry transformations)

$$q = \tilde{q}^\dagger = \begin{pmatrix} a \\ 0 \end{pmatrix}, \quad a \text{ arbitrary}\,. \tag{15.399}$$

The scalar superpartners of the fermionic quarks, (q, q), called squarks, play the role of Higgs fields. As these are in the fundamental representation of the gauge group, $SU(2)$ is completely broken by the super-Higgs mechanism (for $a \neq 0$). It is just the supersymmetric generalization of the familiar Higgs mechanism: three real scalars are eaten by the gluon, in the adjoint representation, and three Weyl spinor combinations of the quark spinors are eaten by the gluino to form a massive Dirac spinor in the adjoint of $SU(2)$. Gluons and gluinos acquire the classical square mass

$$\mathcal{M}_g^2 = 2g_0^2|a|^2, \tag{15.400}$$

where g_0 is the bare gauge coupling. We see that the theory is in the Higgs/confining phase. But there is not mass gap; it remains a massless superfield. Its corresponding massless scalar must move along some flat direction of the classical potential. This flat direction is given by the arbitrary value of the real number $|a|$. This degeneracy is not unphysical, as in the spontaneous breaking of a symmetry. When we move along the supersymmetric flat direction the physical observables change, as for instance the gluon mass (15.400). Different values of $|a|$ correspond to physically inequivalent vacua. The space they expand is called the moduli space. It would be nice to have a gauge invariant parameterization of such an additional parameter of the gauge invariant vacuum. It can only come from the vacuum expectation value of some gauge invariant operator, since it is an independent new classical parameter which does not appear in the bare Lagrangian. The simplest choice is to take the following gauge invariant chiral superfield:

$$M = Q\tilde{Q}\,. \tag{15.401}$$

Classically, its vacuum expectation value is

$$\langle M\rangle = |a|^2, \tag{15.402}$$

a gauge invariant statement and a good parameterization of the flat direction.

There is one consequence of the flat directions in supersymmetric gauge theories that, when combined with the property of holomorphy, will be important to obtain exact results in supersymmetric theories. SQCD depends of the complex coupling $\tau(\mu) = \theta(\mu)/2\pi + 4\pi i/g^2(\mu)$ at scale μ. The angle $\theta(\mu)$ measures the strength of CP violation at scale μ. By asymptotic freedom, the theory is weakly coupled at scales higher than the dynamically generated scale $|\Lambda|$, which is defined by

$$\Lambda \equiv \mu_0 e^{\frac{2\pi i\tau(\mu_0)}{b_0}}\,, \tag{15.403}$$

where μ_0 is the ultraviolet cut-off where the bare parameter $\tau_0 = \tau(\mu_0)$ is defined, and b_0 is the one-loop coefficient of the beta function,

$$\mu\frac{\partial g}{\partial\mu}(\mu) = g\left(-b_0(g^2/16\pi^2) + \mathcal{O}(g^4)\right)\,. \tag{15.404}$$

The complex parameter Λ is renormalization group invariant in the scheme of the Wilsonian effective actions, where holomorphy is not lost (see below). Observe also that the bare instanton angle θ_0 plays the role of the complex phase of Λ^{b_0}.

At scales $\mu \le \mathcal{M}_g$ all the gluons decouple and the relevant degrees of freedom are those of the 'meson' M. Its self-interactions are completely determined by the 'microscopic' degrees of freedom of the super-gluons and super-quarks. We must perform a matching condition for the physics at some scale of order $\mathcal{M}_g$; the renormalization group will secure the physical equivalence at the other energies. If $\mathcal{M}_g\hat{g}\Lambda$, this matching takes place at weak coupling, where perturbation theory in the gauge coupling g is reliable, and we can trust the semiclassical arguments, like those leading to formulae (15.400) and (15.402).

So far we have shown the existence of a flat direction at the classical level. When quantum corrections are included, the flat direction may disappear and a definite value of $\langle M\rangle$ is selected. For the Wilsonian effective description in terms of the relevant degrees of freedom M, this is only possible if a superpotential $\mathcal{W}(M)$ is dynamically generated for M. By the perturbative non-renormalization theorem, this superpotential can only be generated by nonperturbative effects, since classically there was no superpotential for the massless gauge singlet M because of the masslessness of the quark multiplet.

If we turn on a bare mass for the quarks, m, the flat direction is lifted at classical level and a determined value of mass dependent function $\langle M \rangle$ is selected. But the advantage of the flat direction to carry $\langle M \rangle \to \infty$ to be at weak coupling is not completely lost. This limit can now be performed by sending the free parameter m to the appropriate limit, as far as we are able to know the mass dependence of the vacuum expectation value of the meson superfield M. Here holomorphy is very relevant.

Wilsonian effective actions and holomorphy.

The concept of Wilsonian effective action is simple. Any physical process has a typical scale. The idea of the Wilsonian effective action is to give the Lagrangian of some physical processes at its corresponding characteristic scale μ:

$$\mathcal{L}^{(\mu)}(x) = \sum_i g^i(\mu) \boldsymbol{O}_i(x, \mu) \,. \tag{15.405}$$

$\boldsymbol{O}_i(x, \mu)$ are some relevant local composite operators of the effective fields $\varphi_a(p, \mu)$. These are the effective degrees of freedom at scale μ, with momentum modes p running from zero to μ. There could be some symmetries in the operators $\boldsymbol{O}_i$ that our physical system could realize in some way, broken or unbroken. The constants $g^i(\mu)$ measure the strength of the interaction $\boldsymbol{O}_i$ of φ_a at scale μ.

Behind some macroscopic physical processes, there is usually a microscopic theory, with a bare Lagrangian $\mathcal{L}^{(\mu_0)}(x)$ defined at scale μ_0. The microscopic theory has also its characteristic scale μ_0, much higher than the low energy scale μ. Also its corresponding microscopic degrees of freedom, $\phi_j(p, \mu_0)$, may be completely different than the macroscopic ones $\varphi_a(p, \mu)$. The bare Lagrangian encodes the dynamics at scales below the ultraviolet cut-off μ_0. The effective Lagrangian (15.405) is completely determined by the microscopical Lagrangian $\mathcal{L}^{(\mu_0)}(x)$. It is obtained by integrating out the momentum modes p between μ and μ_0. It gives the values of the effective couplings in terms of the bare couplings $g_0^i(\mu_0)$,

$$g^i(\mu) = g^i(\mu; \mu_0, g_0^i(\mu_0)) \,. \tag{15.406}$$

In the macroscopic theory there is no reference to the scale μ_0. Physics is independent of the ultraviolet cut-off μ_0:

$$\frac{\partial g^i}{\partial \mu_0} = 0 \,. \tag{15.407}$$

The μ_0-dependence on the bare couplings $g_0^i(\mu_0)$ cancel the explicit μ_0-dependence in (15.406). This is the action of the renormalization group. It allows to perform the continuum limit $\mu_0 \to \infty$ without changing the low energy physics.

In supersymmetric theories, there are some operators $\boldsymbol{O}_i(z)$, depending only on $z = (x, \theta)$, the chiral superspace coordinate, not on $\overline{\theta}$. Clearly, their field content can only be made of chiral superfields. Those of most relevant physical importance are the superpotential $\mathcal{W}(\Phi_i, \tau_0, m_f)$, and the gauge kinetic operator $\tau(\mu/\mu_0, \tau_0) W^\alpha W_\alpha$. We say that the superpotential $\mathcal{W}$ and the effective gauge coupling τ are holomorphic functions, with the chiral superfields Φ_i, the dimensionless quotient μ/μ_0 and the bare parameters τ_0 and m_f playing the role of the complex variables. The Kähler potential $K(\Phi^\dagger, \Phi)$ is a real function of the variables Φ_i, but as far as supersymmetry is not broken and the theory is not on some Coulomb phase, the vacuum structure is determined by the superpotential in the limit $\mu \to 0$.

We know that complex analysis is substantially more powerful than real analysis. For instance, there are a lot of real functions $f(x)$ that at $x \to 0$ and $x \to \infty$ go like $f(x) \to x$. But there is only one holomorphic function $f(z)$ ($\partial_{\bar{z}} f(z) = 0$) with those properties: $f(z) = z$. The holomorphic constraint is so strong that sometimes the symmetries of the theory, together with some consistency conditions, are enough to determine the unique possible form of the functions $\mathcal{W}$ and τ [24].

An illustrative example is the saturation at one-loop of the holomorphic gauge coupling $\tau(\mu/\mu_0, \tau_0)$ at any order of perturbation theory. Since $\tau_0 = \theta_0/2\pi + i4\pi/g_0^2$, physical periodicity in θ_0 implies

$$\tau(\frac{\mu}{\mu_0}, \tau_0) = \tau_0 + \sum_{n=0}^{\infty} c_n \left(\frac{\mu}{\mu_0}\right) e^{2\pi n i \tau_0}, \tag{15.408}$$

where the sum is restricted to $n \geq 0$ to ensure a well defined weak coupling limit $g_0 \to 0$. The unique term compatible with perturbation theory is the $n = 0$ term. Terms with $n > 0$ corresponds to instanton contributions. The function $c_0(t)$ must satisfy $c_0(t_1 t_2) = c_0(t_1) + c_0(t_2)$ and hence it must be a logarithm. Hence

$$\tau_{\text{pert}}\left(\frac{\mu}{\mu_0}, \tau_0\right) = \tau_0 + \frac{ib_0}{2\pi} \ln \frac{\mu}{\mu_0}\,, \tag{15.409}$$

with b_0 the one-loop coefficient of the beta function. We can use the definition (15.403) of the dynamically generated scale Λ to absorb the bare coupling constant inside the logarithm

$$\tau_{\text{pert}}\left(\frac{\mu}{\Lambda}\right) = \frac{ib_0}{2\pi} \ln \frac{\mu}{\Lambda}\,, \tag{15.410}$$

showing explicitly the independence of the effective gauge coupling in the ultraviolet cut-off μ_0.

We would like to comment that the one-loop saturation of the perturbative beta function and the renormalization group invariance of the scale Λ can be lost by the effect of the Konishi anomaly [25, 26]. In general, after the integration of the modes $\mu < p < \mu_0$ the kinetic terms of the matter fields Φ_i are not canonically normalized,

$$\mathcal{L}^{(\mu)} = \sum_i Z_i(\frac{\mu}{\mu_0}, g_0) \int d^4\theta \Phi_i^\dagger e^{-2V} \Phi_i + \cdots \tag{15.411}$$

These terms have an integral on the whole superspace $(\theta, \overline{\theta})$ and hence are not protected by any non-renormalization theorem. For $N = 1$ gauge theories, holomorphy is absent there, and the functions $Z_i(\frac{\mu}{\mu_0}, g_0)$ are just real functions with perturbative multi-loops contributions. A canonical normalization of the matter fields in the effective action, defining the canonical fields $\Phi_i' = Z_i^{1/2} \Phi_i$ do not leaves invariant the path integral measure $\Pi_i \mathcal{D}\Phi_i$. The anomaly is proportional to $(\sum_i \ln Z_i)\ W^\alpha W_\alpha$, giving a non-holomorphic contribution to the effective coupling τ. For $N = 2$ theories, $Z_i = 1$ and holomorphy is not lost for τ [26, 27].

15.4.11 $N = 1$ SQCD.

Classical Lagrangian and symmetries.

We now analyze $N = 1$ SQCD with gauge group $SU(N_c)$ and N_f flavors * . The field content is the following: There is a spinor chiral superfield W_α in the adjoint of $SU(N_c)$, which contains the gluons A_μ and the gluinos λ. The matter content is given by $2N_f$ scalar chiral superfields Q_f and $\tilde{Q}_f$, $f, \tau f = 1, ..., N_f$, in the $\mathbf{N_c}$ and $\overline{\mathbf{N_c}}$ representations of $SU(N_c)$ respectively. The

*Some reviews on exacts results in $N = 1$ supersymmetric gauge theories are [28].

renormalizable bare Lagrangian is the following:

$$\begin{aligned} \mathcal{L}_{SQCD} &= \frac{1}{8\pi}\text{Im}\left(\tau_0 \int d^2\theta\; W^\alpha W_\alpha\right) \\ &+ \int d^4\theta \left(Q_f^\dagger e^{-2V} Q_f + \tilde{Q}_f e^{2V} \tilde{Q}_f^\dagger\right) \\ &+ \left(\int d^2\theta\; m_f \tilde{Q}_f Q_f \;\; + \text{h.c.}\right), \end{aligned} \tag{15.412}$$

with $\tau_0 = \theta_0/2\pi + i4\pi/g_0^2$ and m_f the bare couplings. In the massless limit the global symmetry of the classical Lagrangian is $SU(N_f)_L \times SU(N_f)_R \times U(1)_B \times U(1)_A \times U(1)_R$. For $N_c = 2$ the representations $\mathbf{2}$ and $\overline{\mathbf{2}}$ are equivalent, and the global symmetry group is enlarged. In general we consider $N_c > 2$. The $U(1)_A$ and $U(1)_R$ symmetries are anomalous and are broken by instanton effects. But we can perform a linear combination of $U(1)_A$ and $U(1)_R$, call it $U(1)_{AF}$, that is anomaly free. We have the following table of representations for the global symmetries of SQCD:

	$SU(N_f)_L$	$SU(N_f)_R$	$U(1)_B$	$U(1)_{AF}$
W_α	$\mathbf{1}$	$\mathbf{1}$	0	1
Q_f	$\mathbf{N_c}$	$\mathbf{1}$	1	$\frac{(N_f - N_c)}{N_f}$
$\tilde{Q}_f$	$\mathbf{1}$	$\overline{\mathbf{N_c}}$	-1	$\frac{(N_f - N_c)}{N_f}$

The anomaly free R-charges, R_{AF}, are derived by the following. The superfield W_α is neutral under $U(1)_A$ and its R-transformation is fixed to be

$$W_\alpha(x,\theta) \to e^{i\beta} W_\alpha(x, e^{-i\beta}\theta). \tag{15.413}$$

Consider now that the fermionic quarks ψ have charge R_ψ under an $U(1)_{AF}$ transformation. In the one-instanton sector, λ has $2N_c$ zero modes, and one for each Q_f and $\tilde{Q}_f$. In total we have $2N_c + 2N_f R_\psi = 0$ to avoid the anomalies. We derive that $R_\psi = -N_c/N_f$. Since this is the charge of the fermions, the superfields $(Q_f, \tilde{Q}_f)$ have R_{AF} charge $1 - N_c/N_f = (N_f - N_c)/N_f$.

The classical moduli space.

The classical equations of motion of the auxiliary fields are

$$\begin{aligned} \overline{F}_{q_f} &= -m_f \tilde{q}_f = 0\,, \\ \overline{F}_{\tilde{q}_f} &= -m_f q_f = 0\,, \\ D^a &= \sum_f \left(q_f^\dagger T^a q_f - \tilde{q}_f T^a \tilde{q}_f^\dagger\right) = 0\,. \end{aligned} \tag{15.414}$$

If there is a massive flavor $m_f \neq 0$, then we must have $q_f = \tilde{q}_f = 0$. As we want to go to the infrared limit to analyze the vacuum structure, the interesting case is the situation of N_f massless flavors. If some quark has a non-zero mass m, its physical effects can be decoupled at very low energy, by taking into account the appropriate physical matching conditions at the decoupling scale m (see below). If all quarks are massive, in the infrared limit we only have a pure $SU(N_c)$ supersymmetric gauge theory. The Witten index of pure $SU(N_c)$ super Yang-Mills is $\text{tr}(-1)^F = N_c$ [29]. We know that supersymmetry is not broken dynamically in this theory, and that there are N_c equivalent vacua. The $2N_c$ gaugino zero modes break the $U(1)_R$ symmetry to Z_{2N_c} by the instantons. Those N_c vacua corresponds to the spontaneously broken discrete symmetry Z_{2N_c} to Z_2 by the gaugino condensate $\langle\lambda\lambda\rangle \neq 0$.

If there are some massless super-quarks, they can have non-trivial physical effects on the vacuum structure. Consider that we have N_f massless flavors. We can look at the q_f and $\tilde{q}_f$ scalar quarks as $N_c \times N_f$ matrices. Using $SU(N_c) \times SU(N_f)$ transformations, the q_f matrix can be rotated into a simple form. There are two cases to be distinguished:

a) $N_f < N_c$:

In this case we have that the general solution of the classical vacuum equations (15.414) is:

$$q_f = \tilde{q}_f^\dagger = \begin{pmatrix} v_1 & 0 & \cdots & 0 \\ 0 & v_2 & & \\ & & \ddots & \\ 0 & \cdots & & v_{N_f} \\ \vdots & & & \vdots \\ 0 & \cdots & & 0 \end{pmatrix}, \tag{15.415}$$

with v_f arbitrary. These scalar quark's vacuum expectation values break spontaneously the gauge group to $SU(N_c - N_f)$. By the super-Higgs mechanism, $N_c^2 - (N_c - N_f)^2 = 2N_c N_f - N_f^2$ chiral superfields are eaten by the vector superfields. This leaves $2N_f N_c - (2N_f N_c - N_f^2) = N_f^2$ chiral superfields. They can be described by the meson operators

$$M_{fg} \equiv \tilde{Q}_f Q_g. \tag{15.416}$$

which provide a gauge invariant description of the classical moduli space.

b) $N_f \geq N_c$:

In this case the general solution of (15.414) is:

$$q_f = \begin{pmatrix} v_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & v_2 & & \vdots & & \vdots \\ & & \ddots & & & \vdots \\ & & & v_{N_c} & \cdots & 0 \end{pmatrix}, \tag{15.417}$$

$$\tilde{q}_f^\dagger = \begin{pmatrix} \tilde{v}_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \tilde{v}_2 & & \vdots & & \vdots \\ & & \ddots & & & \vdots \\ & & & \tilde{v}_{N_c} & \cdots & 0 \end{pmatrix}, \tag{15.418}$$

with the parameters v_i, $\tilde{v}_i$ $(i = 1, ..., N_c)$ subject to the constraint

$$|v_i|^2 - |\tilde{v}_i|^2 = \text{constant independent of } i. \tag{15.419}$$

Now the gauge group is completely higgsed. The gauge invariant parameterization of the classical moduli space must be done by $2N_f N_c - (N_c^2 - 1)$ chiral superfields. For instance, if $N_f = N_c$, we need $N_c^2 + 1$ superfields. The meson operators M_{fg} provide N_c^2. The remaining degree of freedom comes from the baryon-like operators

$$\begin{aligned} B &= \epsilon^{f_1 \cdots f_{N_f}} Q_{f_1} \cdots Q_{f_{N_f}}, \\ \tilde{B} &= \epsilon^{f_1 \cdots f_{N_f}} \tilde{Q}_{f_1} \cdots \tilde{Q}_{f_{N_f}}, \end{aligned} \tag{15.420}$$

with the color indices also contracted by the ϵ-tensor. These are two superfields, but there is a holomorphic constraint

$$\det M - \tilde{B}B = 0\,. \tag{15.421}$$

For $N_f = N_c + 1$, we need $2N_c(N_c+1) - (N_c^2 - 1) = N_c^2 + 2N_c + 1$ independent chiral superfields. We can construct the baryon operators:

$$B^f = \epsilon^{f f_1 \cdots f_{N_c}} Q_{f_1} \cdots Q_{f_{N_c}},$$
$$\tilde{B}^f = \epsilon^{f f_1 \cdots f_{N_c}} \tilde{Q}_{f_1} \cdots \tilde{Q}_{f_{N_c}}. \tag{15.422}$$

M_{fg}, B^f and $\tilde{B}^f$ have $(N_c+1)^2 + 2(N_c+1)$ components. The matrix M_{fg} has rank N_c, which can be expressed by the $2(N_c+1)$ constraints:

$$M_{fg} B^g = M_{fg} \tilde{B}^g = 0\,. \tag{15.423}$$

And in total we get the needed $N_c^2 + 2N_c + 1$ independent chiral superfields.

As N_f increases, we get more and more constraints. Each case with $N_f \geq N_c$ is interesting by itself and we will have to look at them in different ways.

15.4.12 The vacuum structure of SQCD with $N_f < N_c$.

The Affleck-Dine-Seiberg's superpotential.

First we consider the case of massless flavors. At the classical level there are flat directions parameterized by the free vacuum expectation values of the meson fields M_{fg}. They belong to the representation $(\mathbf{N_f}, \overline{\mathbf{N}}_\mathbf{f}, 0, 2(N_f - N_c)/N_f)$ of the global symmetry group $SU(N_f)_L \times SU(N_f)_R \times U(1)_B \times U(1)_{AF}$. If nonperturbative effects generate a Wilsonian effective superpotential $\mathcal{W}$, it must depend in a holomorphic way of the light chiral superfields M_{fg} and the bare coupling constant τ_0. The renormalization group invariance of the Wilsonian effective action demands that the dependence on the bare coupling constant τ_0 of $\mathcal{W}$ enters thought the dynamically generated scale Λ_{N_f,N_c}. The invariance of $\mathcal{W}$ under $SU(N_f)_L \times SU(N_f)_R$ rotations reduces the dependence in the mesons fields to the combination $\det M$. There is only one holomorphic function $\mathcal{W} = \mathcal{W}(\det M, \Lambda_{N_f,N_c})$, with R_{AF} charge two that can be built from the variables $\det M$ and Λ_{N_f,N_c}, which have R_{AF} charge $2(N_f - N_c)$ and zero, respectively. It is the Affleck-Dine-Seiberg's superpotential [30, 31]

$$\mathcal{W} = c_{N_f,N_g} \left(\frac{\Lambda_{N_f,N_c}}{\det M} \right)^{\frac{1}{(N_c - N_f)}}, \tag{15.424}$$

where c_{N_f,N_c} are some undetermined dimensionless constants. If $c_{N_f,N_c} \neq 0$, (15.424) corresponds to an exact nonperturbative dynamically generated Wilsonian superpotential. It has catastrophic consequences, the theory has no vacuum. If we try to minimize the energy derived from the superpotential (15.424) we find that $|\langle \det M \rangle| \to \infty$.

Massive flavors.

When we add mass terms for all the flavors we expect to find some physical vacua. In fact, by Witten index, we should find N_c of them. To verify this, let us try to compute $\langle M_{fg} \rangle$ taking advantage of its holomorphy and symmetries.

A bare mass matrix $m_{fg} \neq 0$ breaks explicitly the $SU(N_f) \times SU(N_f)_R \times U(1)_{AF}$ global symmetry of the bare Lagrangian (15.412). In terms of the meson operator the mass term is

$$\mathcal{W}_{\text{tree}} = \operatorname{tr}(mM). \tag{15.425}$$

We see that, under an L and R rotation of $SU(N_f)_L$ and $SU(N_f)_R$ respectively, we can recover the $SU(N_f)_L \times SU(N_f)_R$ invariance if we require m to transform as $m \to L^{-1} m R$. In the same way, as the superpotential has R-charge two, the $U(1)_{AF}$ invariance is recovered if we assign the charge $2-2(N_f-N_c)/N_f = 2N_c/N_f$ to the mass matrix m. The vacuum expectation value of the matrix chiral superfield M is a holomorphic function of Λ_{N_f,N_c} and m. To implement the same action under $SU(N_f)_L \times SU(N_f)_R$ rotations, we must have

$$\langle M \rangle = f(\det m, \Lambda_{N_f,N_c}) m^{-1}. \tag{15.426}$$

The dependence in $\det m$ of the function f is determined by the R_{AF} charge. Then, the Λ_{N_f,N_c} dependence is worked out by dimensional analysis. The result is

$$\langle M \rangle = (\text{const}) \left(\Lambda_{N_f,N_c}^{3N_c-N_f} \det m \right)^{\frac{1}{N_c}} m^{-1} . \tag{15.427}$$

The N_c roots give N_c vacua. Observe that this is an exact result, and valid also for $N_f \geq N_c$. There is only an dimensionless constant (in general N_f and N_c dependent) to be determined. It would be nice to be able to carry its computation in the weak coupling limit, since holomorphy would allow to extend (15.427) also to the strong coupling region.

The result (15.427) suggest the existence of an effective superpotential out of which (15.427) can be obtained. Holomorphy and symmetries tell us that the possible superpotential would have to be

$$\begin{aligned} \mathcal{W}(M, \Lambda_{N_f,N_c}, m) = & \left(\frac{\Lambda_{N_f,N_c}}{\det M} \right)^{\frac{1}{(N_c-N_f)}} . \\ & f\left(t = \operatorname{tr}(mM) \left(\frac{\Lambda_{N_f,N_c}}{\det M} \right)^{\frac{-1}{(N_c-N_f)}} \right). \end{aligned} \tag{15.428}$$

In the limit of weak coupling, $\Lambda_{N_f,N_c} \to 0$, we know that $f(t) = c_{N_f,N_c} + t$. But we can play at the same time with the free values of m to reach any desired value of t. This fixes the function $f(t)$ and the superpotential $\mathcal{W}(M, \Lambda_{N_f,N_c}, m)$ to be

$$\begin{aligned} \mathcal{W}(M, \Lambda_{N_f,N_c}, m) = & \; c_{N_f,N_c} \left(\frac{\Lambda_{N_f,N_c}}{\det M} \right)^{\frac{1}{(N_c-N_f)}} \\ & + tr\,(mM). \end{aligned} \tag{15.429}$$

As a consistency check, when we solve the equations $\partial \mathcal{W}/\partial M = 0$, we obtain the previously determined vacuum expectation values (15.427).

Finally, we have to check the non-vanishing of c_{N_f,N_c}. We take advantage of the decoupling theorem to obtain further information about the constants c_{N_f,N_c}. Let us add a mass term m only for the N_f flavor,

$$\begin{aligned} \mathcal{W}(M, \Lambda_{N_f,N_c}, m) = & \left(\frac{\Lambda_{N_f,N_c}}{\det M} \right)^{\frac{1}{(N_c-N_f)}} \\ & + m M_{N_f N_f}. \end{aligned} \tag{15.430}$$

Solving for the equations:

$$\begin{aligned} \frac{\partial \mathcal{W}}{\partial M_{f N_f}}(M, \Lambda_{N_f,N_c}, m) = 0, \\ \frac{\partial \mathcal{W}}{\partial M_{N_f f}}(M, \Lambda_{N_f,N_c}, m) = 0, \end{aligned} \tag{15.431}$$

for $f < N_f$ gives that $M_{f N_f} = M_{N_f f} = 0$. Hence $\det M = M_{N_f N_f} \cdot \det \hat{M}$, with $\hat{M}$ the $(N_f-1) \times (N_f-1)$ matrix meson operator of the $N_f - 1$ massless flavors. At scales below m, the N_f-th flavor decouples and its corresponding $M_{N_fN_f}$ meson operator is frozen to the value that satisfies:

$$\frac{\partial \mathcal{W}}{\partial M_{N_fN_f}}(M, \Lambda_{N_f,N_c}, m) = -\frac{c_{N_f,N_c}}{(N_f - N_c)} \cdot \Lambda_{N_f,N_c}^{(3N_f-N_c)/(N_f-N_c)} (\det M)^{\frac{1}{(N_c-N_f)}-1} \det \hat{M} + m = 0. \tag{15.432}$$

If we substitute the solution $\langle M_{N_fN_f} \rangle$ of the previous equation into the superpotential $\mathcal{W}(M, \Lambda_{N_f,N_c}, m)$, we should obtain the superpotential $\mathcal{W}(\hat{M}, \Lambda_{N_f-1,N_c}, 0)$ of $N_f - 1$ massless flavors with the dynamically generated scale Λ_{N_f-1,N_c}. The matching conditions at scale m between the theory with N_f flavors and the theory with $N_f - 1$ flavors gives the relation

$$m\Lambda_{N_f,N_c}^{3N_c-N_f} = \Lambda_{N_f-1,N_c}^{3N_c-N_f+1}, \tag{15.433}$$

thus,

$$\mathcal{W}(M, \Lambda_{N_f,N_c}, m)|_{\langle M_{N_fN_f} \rangle} = (N_c - N_f + 1) \cdot \left(\frac{c_{N_f,N_c}}{N_c - N_f}\right)^{\frac{N_c-N_f}{N_c-N_f+1}} \left(\frac{\Lambda_{N_f-1,N_c}}{\det \hat{M}}\right)^{\frac{1}{(N_c-N_f+1)}}, \tag{15.434}$$

and we obtain the relation

$$\left(\frac{c_{N_f,N_c}}{N_c - N_f}\right)^{N_c-N_f} = \left(\frac{c_{N_f-1,N_c}}{N_c - N_f + 1}\right)^{N_c-N_f+1}. \tag{15.435}$$

Similarly, we can try to obtain another relation between the constants c_{N_f,N_c} for different numbers of colors. To this end we give a large expectation value to $M_{N_fN_f}$ with respect the expectation values of $\hat{M}$. Then below the scale $\langle M_{N_fN_f} \rangle$ we have SQCD with $N_c - 1$ colors and $N_f - 1$ flavors. Following the same strategy as before we find that $c_{N_f-1,N_c-1} = c_{N_c,N_f}$. It means that $c_{N_c,N_f} = c_{N_f-N_c}$, which together with the relation (15.435) gives

$$c_{N_f,N_c} = (N_c - N_f) c_{1,2}\,. \tag{15.436}$$

We just have to compute the dimensionless constant $c_{1,2}$ of the gauge group $SU(2)$ with one flavor. In this case, or for the general case of $N_f = N_c - 1$, the gauge group is completely higgsed and there are not infrared divergences in the instanton computation. In the weak coupling limit the unique surviving nonperturbative contributions come from the one-instanton sector. A direct instanton calculation reveals that the constant $c_{2,1} \neq 0$ [31] *.

For $N_f < N_c - 1$ there is an unbroken gauge group $SU(N_c - N_f)$. At scales below the smallest eigenvalue of the matrix $\langle M_{fg} \rangle$ we have a pure super Yang-Mills theory with $N_c - N_f$ colors. This theory is believed to confine with a mass gap given by the gaugino condensate $\langle \lambda\lambda \rangle \neq 0$. Consider the simplest case of $\langle M_{fg} \rangle = \mu^2 \mathbf{1}_{N_f}$. Matching the gauge couplings at scale μ gives $\Lambda_{N_f,N_c}^{3N_c-N_f} = (\det M)\, \Lambda_{0,N_c-N_f}^{3(N_c-N_f)}$, which implies for the effective superpotential

$$\mathcal{W} = (N_c - N_f) \Lambda_{0,N_c-N_f}^3. \tag{15.437}$$

On the other hand, the gaugino bilinear $\lambda\lambda$ is the lowest component of the chiral superfield $S = W^\alpha W_\alpha$, which represents the super-glueball operator. The bare gauge coupling τ_0 acts as the source of the operator S. If we differentiate (15.437) with respect to $\ln \Lambda^{3(N_c-N_f)}$ we obtain the gaugino condensate

$$\langle \lambda\lambda \rangle = \Lambda_{0,N_c-N_f}^3. \tag{15.438}$$

*In the $\overline{\text{DR}}$ scheme $c_{2,1} = 1$ [32]. If we do not say the contrary, we will work on such a scheme.

In fact, following the 'integrating in' procedure [33, 34], we would obtain the Veneziano-Yankielowicz effective Lagrangian [35].

It is not possible to extend the Affleck-Dine-Seiberg's superpotential to the case of $N_f \geq N_c$. For these values the quantum corrections do not lift the flat directions, and we still have a moduli space which may be different from the classical one. This is the case of $N_f = N_c$.

15.4.13 The vacuum structure of SQCD with $N_f = N_c$.

A quantum modified moduli space.

For $N_f = N_c$, the classical moduli space is spanned by the gauge singlet operators M_{fg}, B and $\tilde{B}$ subject to the constraint $\det M - \tilde{B}B = 0$. At quantum level, instanton effects could change the classical constraint to

$$\det M - \tilde{B}B = \Lambda^{2N_c}, \tag{15.439}$$

since $\Lambda^{2N_c} \sim e^{-8\pi/g^2 + i\theta}$ corresponds to the one-instanton factor, it has the right dimensions, and the operators $(Q_f, \tilde{Q}_f)$ have R_{AF} charge zero.

To check if the quantum correction (15.439) really takes place, add a mass term for the quarks. The unique possible holomorphic term with R_{AF} charge two that can be generated with the variables $(M_{fg}, B, \tilde{B}, \Lambda, m)$ is

$$\mathcal{W} = \operatorname{tr} mM\,. \tag{15.440}$$

Imagine now that the N_c-flavor is much heavier, with bare mass m, than the $N_c - 1$ other ones, with bare mass matrix $\hat{m}$. The degree of freedom $M_{N_c N_c}$ is given by the constraint. Locate at $B = \tilde{B} = M_{fN_c} = 0$. By equation (15.427) we know that the $(N_c - 1) \times (N_c - 1)$ matrix $\hat{M}$ is determined to be

$$\hat{M} = \left(\Lambda^{2N_c+1}_{N_c-1,N_c} \det\hat{m}\right)^{\frac{1}{N_c}} \hat{m}^{-1}, \tag{15.441}$$

which has a non-zero determinant. It indicates that the constraint (15.439) is really generated at quantum level [36]. As a final check, consider the simplest situation of $N_c - 1$ massless flavors. When we use the constraint (15.439) to express $M_{N_c N_c}$ as function of $\det\hat{M}$ we obtain

$$\mathcal{W} = \frac{m\Lambda^{2N_c}}{\det\hat{M}}\,, \tag{15.442}$$

the Affleck-Dine-Seiberg's superpotential for $N_f = N_c - 1$ massless flavors.

Far from the origin of the moduli field space we are at weak coupling and the quantum moduli space given by the constraint (15.439) looks like the classical moduli space (15.421). But far from the origin of order Λ, the one-instanton sector is sufficiently strong to change significatively the vacuum structure. Observe that the classically allowed point $M = B = \tilde{B} = 0$ is not a point of the quantum moduli space and the gluons never become massless.

Patterns of spontaneous symmetry breaking and 't Hooft's anomaly matching conditions.

Our global symmetries are $SU(N_f)_L \times SU(N_f)_R \times U(1)_B \times U(1)_{AF}$. Since for $N_f = N_c$ the super-quarks are neutral with respect to the non-anomalous symmetry $U(1)_{AF}$, it is never spontaneously broken. The other symmetries present different patterns of symmetry breaking depending on which point of the moduli space the vacuum is located *.

For instance, the point

$$M = \Lambda^2 \mathbf{1}_{N_f}, \quad B = \tilde{B} = 0, \tag{15.443}$$

*Different patterns of symmetry breaking have also been observed in softly broken $N = 2$ SQCD [15].

suggests the spontaneous symmetry breaking

$$SU(N_f)_L \times SU(N_f)_R \times U(1)_B \times U(1)_{AF} \\ \longrightarrow SU(N_f)_V \times U(1)_B \times U(1)_{AF}, \tag{15.444}$$

with $SU(N_f)_V$ the diagonal part of $SU(N_f) \times SU(N_f)_R$. To check it, the unbroken symmetries must satisfy the 't Hooft's anomaly matching conditions [37].

With respect to the unbroken symmetries the quantum numbers of the elementary and composite massless fermions, at high and low energy respectively, are

	$SU(N_f)_V$	$U(1)_B$	$U(1)_{AF}$
λ	$\mathbf{1}$	0	1
ψ_q	$\mathbf{N_f}$	1	-1
$\psi_{\tilde{q}}$	$\overline{\mathbf{N}}_\mathbf{f}$	-1	-1
ψ_M	$\mathbf{N_f^2 - 1}$	0	-1
ψ_B	$\mathbf{1}$	N_f	-1
$\psi_{\tilde{B}}$	$\mathbf{1}$	$-N_f$	-1

Observe there are only $N_f^2 - 1$ independent meson fields, arranged in the adjoint of $SU(N_f)_V$, since the constraint (15.439) eliminates one of them. There are $N_f^2 - 1$ gluinos and N_f extra components for each quark ψ_q and anti-quark $\psi_{\tilde{q}}$ because of the gauge group $SU(N_c)$. The anomaly coefficients are:

triangles	high energy	low energy
$SU(N_f)^2 \times U(1)_{AF}$	$-2N_f T(\mathbf{N_f})$	$-T(\mathbf{N_f^2 - 1})$
$U(1)_{AF}^3$	$-2N_f^2 + (N_f^2 - 1)$	$-(N_f^2 - 1) - 2$
$U(1)_B^2 \times U(1)_{AF}$	$-N_f^2 - N_f^2$	$-2N_f^2$
$\operatorname{tr} U(1)_{AF}$	$-2N_f^2 + N_f^2 - 1$	$-(N_f^2 - 1) - 2$

The constants $T(R)$ are defined by $\operatorname{tr}(T^a T^a) = T(R)\delta^{ab}$, with T^a in the representation R of the group $SU(N)$. For the fundamental representation, $T(\mathbf{N}) = 1/2$. For the adjoint representation, $T(\mathbf{N^2 - 1}) = N$. The coefficient of $\operatorname{tr} U(1)_{AF}$ corresponds to the gravitational anomaly. One can check that all the anomalies match perfectly, supporting the spontaneous symmetry breaking pattern of (15.444).

The quantum moduli space of $N_f = N_c$ allows another particular point with a quite different breaking pattern. It is:

$$M = 0, \quad B = -\tilde{B} = \Lambda^{N_c}. \tag{15.445}$$

At this point, only the vectorial baryon symmetry is broken, all the chiral symmetries $SU(N_f)_L \times SU(N_f)_R \times U(1)_{AF}$ remain unbroken. We check this pattern with the help of the 't Hooft's anomaly matching conditions again. In this case we have the quantum numbers:

	$SU(N_f)_L$	$SU(N_f)_R$	$U(1)_{AF}$
λ	$\mathbf{1}$	$\mathbf{1}$	1
ψ_q	$\mathbf{N_f}$	$\mathbf{1}$	-1
$\psi_{\tilde{q}}$	$\mathbf{1}$	$\overline{\mathbf{N}}_\mathbf{f}$	-1
ψ_M	$\mathbf{N_f}$	$\overline{\mathbf{N}}_\mathbf{f}$	-1
ψ_B	$\mathbf{1}$	$\mathbf{1}$	-1
$\psi_{\tilde{B}}$	$\mathbf{1}$	$\mathbf{1}$	-1

and the anomaly coefficients are:

triangles	high energy	low energy
$SU(N_f)^3_L$	$N_f C_3$	$N_f C_3$
$SU(N_f)^3_R$	$N_f C_3$	$N_f C_3$
$SU(N_f)^2 \times U(1)_{AF}$	$-N_f T(\mathbf{N_f})$	$-N_f T(\mathbf{N_f})$
$U(1)^3_{AF}$	$-2N_f^2 + N_f^2 - 1$	$-N_f^2 - 1$

where C_3 is defined by $\mathrm{tr}(T^a\{T^b, T^c\}) = C_3 d^{abc}$, with T^a in the fundamental representation of $SU(N_f)$. Because of the constraint (15.439) there is only one independent baryonic degree of freedom. The anomaly coefficients match perfectly.

15.4.14 The vacuum structure of SQCD with $N_f = N_c + 1$.

The quantum moduli space.

First we consider if the classical constraints:

$$M_{fg}B^g = M_{fg}\tilde{B}^f = 0, \tag{15.446}$$

$$\det M (M^{-1})^{fg} - B^f \tau B^g = 0, \tag{15.447}$$

are modified quantum mechanically. For $N_f = N_c + 1$ the quark multiplets $(Q_f, \tau Q_f)$ have R_{AF} charge equal to $1/N_f$. The mass matrix breaks the $U(1)_{AF}$ symmetry with a charge of $2 - 2/N_f = 2N_c/N_f$. It is exactly the charge $U(1)_{AF}$ of equation (15.447). On the other hand, the instanton factor Λ^{2N_c-1} supplies the right dimensionality. Then, there is the possibility that the classical constraint (15.447) is modified by nonperturbative contributions to

$$\det M (M^{-1})^{fg} - B^f \tau B^g = \Lambda^{2N_c - 1} m^{fg}. \tag{15.448}$$

On the other hand, one can see that the classical constraints (15.446) do not admit modification. Then if $M \neq 0$ we have $B^f = \tau B^g = 0$. Using (15.427), we obtain

$$\det M (M^{-1})^{fg} = \Lambda^{2N_c - 1} m^{fg}\,, \tag{15.449}$$

and the quantum modification (15.448) really takes place [36].

S-confinement.

In the massless limit $m^{fg} \to 0$, (15.446) and (15.447) are satisfied at the quantum level. It means that the origin of field space, $M = B = \tau B = 0$, is an allowed point of the quantum moduli space. On such a point, there is no spontaneous symmetry breaking at all. We use the 't Hooft's anomaly matching conditions to check it. The quantum numbers of the massless fermions at high and low energy are:

	$SU(N_f)_L$	$SU(N_f)_R$	$U(1)_B$	$U(1)_{AF}$
λ	$\mathbf{1}$	$\mathbf{1}$	0	1
ψ_q	$\mathbf{N_f}$	$\mathbf{1}$	1	$\frac{1}{N_f} - 1$
$\psi_{\tilde{q}}$	$\mathbf{1}$	$\overline{\mathbf{N_f}}$	-1	$\frac{1}{N_f} - 1$
ψ_M	$\mathbf{N_f}$	$\overline{\mathbf{N_f}}$	0	$\frac{2}{N_f} - 1$
ψ_B	$\overline{\mathbf{N_f}}$	$\mathbf{1}$	$N_f - 1$	$-\frac{1}{N_f}$
$\psi_{\tilde{B}}$	$\mathbf{1}$	$\mathbf{N_f}$	$1 - N_f$	$-\frac{1}{N_f}$

and the anomaly coefficients are:

triangles	high energy	low energy
$SU(N_f)^3$	$N_c C_3$	$N_f C_3 + \overline{C}_3$
$SU(N_f)^2 \times U(1)_{AF}$	$N_c T(\mathbf{N_f})(-\frac{N_c}{N_f})$	$N_f T(\mathbf{N_f})(\frac{2}{N_f} - 1)$ $+T(\mathbf{N_f})(-\frac{1}{N_f})$
$U(1)_B^2 \times U(1)_{AF}$	$2N_c N_f(-\frac{N_c}{N_f})$	$2N_f N_c^2(-\frac{1}{N_f})$
$U(1)_{AF}^3$	$(N_c^2 - 1)$ $+2N_f N_c(-\frac{N_c}{N_f})^3$	$N_f^2(\frac{2}{N_f} - 1)^3$ $+2N_f(-\frac{1}{N_f})^3$
$\mathrm{tr}\, U(1)_{AF}$	$(N_c^2 - 1)$ $+2N_f N_c(-\frac{N_c}{N_f})$	$N_f^2(\frac{2}{N_f} - 1)$ $+2N_f(-\frac{1}{N_f})$

with complete agreement. Hence, at the origin of field space we have massless mesons and baryons, and the full global symmetry is manifest. It is a singular point, with the number of massless degrees of freedom larger than the dimensionality of the space of vacua. As we move along the moduli space away from the origin, the 'extra' fields become massive and the massless fluctuations match with the dimensionality of the moduli space. As we are in a Higgs/confining phase, there should be a smooth connection of the dynamics at the origin of field space with the one away from it. This dynamics must be given by some nonperturbative superpotential of mesons and baryons. A theory with the previous characteristics is called s-confining.

There is a unique effective superpotential yielding all the constraints [36],

$$\mathcal{W} = \frac{1}{\Lambda^{2N_f - 3}} (\tau B^g M_{gf} B^f - \det M) \,, \qquad (15.450)$$

it satisfies:

i) Invariance under all the symmetries.

ii) The equations of motion $\partial \mathcal{W}/\partial M = \partial \mathcal{W}/\partial B = \partial \mathcal{W}/\partial \tau B = 0$ give the constraints (15.446, 15.447).

iii) At the origin all the fields are massless.

iv) Adding the bare term $\mathrm{tr}\,(mM) + b_f B^f + \tau b_f \tau B^f$ we recover the $N_f < N_c + 1$ results.

15.4.15 Seiberg's duality.

The dual SQCD.

If we try to extend the same view of $SU(N_c)$ SQCD for the case of $N_f > N_c + 1$, *i.e.*, as being in a Higgs/confining phase with the vacuum structure determined by meson and baryons operators satisfying the corresponding classical constraints, to the case of $N_f > N_c+1$ (it is not possible to modify the classical constraints for $N_f > N_c + 1$), we obtain inconsistencies. It is not possible to generate a superpotential yielding to the constraints, and the 't Hooft's anomaly matching conditions are not satisfied. It indicates that for $N_f > N_c + 1$ the Higgs/confining description of SQCD at large distances in terms of just M, B and τB is no longer valid.

For $N_f > N_c + 1$, Seiberg conjectured [38] that the infrared limit of SQCD with N_f flavors admits a dual description in terms of an $N = 1$ super Yang-Mills gauge theory with $\tau N_c = N_f - N_c$ number of colors, N_f flavors D^f and τD^f in the fundamental and anti-fundamental representations of $SU(N_f - N_c)$ respectively, and N_f^2 gauge singlet chiral superfields $M^{(m)}_{gf}$. The fields $M^{(m)}_{gf}$ couple to D_f and τD_f through the relevant bare superpotential

$$\mathcal{W} = M^{(m)}_{gf} \tau D^g D^f. \tag{15.451}$$

If both theories are going to describe the same physics at large distances, we must be able to give a prescription of the gauge invariant operators M_{gf}, $B^{f_1 \cdots f_{\tau N_c}}$ and $\tau B^{f_1 \cdots f_{\tau N_c}}$ in terms of the dual microscopic operators $(D^f, \tau D^f)$ and $M^{(m)}_{gf}$. The simplest identification is:

$$\begin{aligned} M_{gf} &= \mu M^{(m)}_{gf}, \\ B^{f_1 \cdots f_{\tau N_c}} &= D^{f_1} \cdots D^{f_{\tau N_c}}, \\ \tau B^{f_1 \cdots f_{\tau N_c}} &= \tau D^{f_1} \cdots \tau D^{f_{\tau N_c}}. \end{aligned} \tag{15.452}$$

In the baryon operators the $SU(\tau N_c)$ color indices of $(D^f, \tau D^f)$ are contracted with the τN_c antisymmetric tensor. The scale μ is introduced because the dimension of the bare operator $M^{(m)}_{gf}$, derived from (15.451), is one. This mass scale relates the intrinsic scales Λ and $\tau\Lambda$ of the $SU(N_c)$ and $SU(\tau N_c)$ gauge theories through the equation

$$\Lambda^{3N_c - N_f} \tau\Lambda^{3\tau N_c - N_f} = (-1)^{N_f - N_c} \mu^{N_f}. \tag{15.453}$$

We see that an strongly coupled $SU(N_c)$ gauge theory corresponds to a weakly coupled $SU(\tau N_c)$ gauge theory, in analogy with the electric-magnetic duality. From this analogy, we call the $SU(N_c)$ gauge theory the electric one, and the $SU(\tau N_c)$ gauge theory the magnetic one.

Both theories must have the same global symmetries. The mapping (15.452) gives the quantum numbers of the magnetic degrees of freedom. Once more, 't Hooft's anomaly matching conditions for the electric and magnetic theories give a non-trivial check of (15.452). In the following table we write the quantum numbers for the fermions of the magnetic theory:

	$SU(N_f)_L$	$SU(N_f)_R$	$U(1)_B$	$U(1)_{AF}$
$\tau\lambda$	$\mathbf{1}$	$\mathbf{1}$	0	1
ψ_d	$\overline{\mathbf{N_f}}$	$\mathbf{1}$	$\frac{N_c}{\tau N_c}$	$\frac{\tau N_c}{N_f}$
$\psi_{\tilde{d}}$	$\mathbf{1}$	$\mathbf{N_f}$	$-\frac{N_c}{\tau N_c}$	$\frac{\tau N_c}{N_f}$
ψ_m	$\mathbf{N_f}$	$\overline{\mathbf{N_f}}$	0	$1 - 2\frac{N_c}{N_f}$

with $\tau\lambda$ the magnetic gluinos. One can check that both theories give the same anomalies.

It can be verified that applying duality again we obtain the original theory.

$N_c + 1 < N_f \leq 3N_c/2$. An infrared free non-Abelian Coulomb phase.

In this range of N_f the magnetic theory is not asymptotically free and has a trivial infrared fixed point. At large distances the physical effective degrees of freedom are the fields D^f, D^f, M_{gf} and the massless super-gluons of the gauge group $SU(N_f - N_c)$. At the origin of field space we are in an infrared free non-Abelian Coulomb phase, with a complete screening of its charges in the infrared limit. Observe that the strongly coupled electric theory is weakly coupled in terms of the magnetic degrees of freedom, according to the philosophy of the electric-magnetic duality.

$3N_c/2 < N_f < 3N_c$. An interacting non-Abelian Coulomb phase.

As in QCD, the $N = 1$ SQCD has a Banks-Zaks fixed point [39] for $N_c, N_f \to \infty$, when $N_f/N_c = 3 - \epsilon$ with $\epsilon\mathcal{L}1$. We still have asymptotic freedom and under the renormalization group transformations the theory flows from the ultraviolet free fixed point to an infrared fixed point with a non-zero finite value of the gauge coupling constant. If there is an interacting superconformal gauge theory the scaling dimensions of some gauge invariant operators should be non-trivial.

The superconformal invariance includes an R-symmetry, from which the scaling dimensions of the operators satisfy the lower bound

$$D \geq \frac{3}{2}|R| \tag{15.454}$$

with equality for chiral and anti-chiral operators. The R-current is in the same supermultiplet as the energy-momentum tensor, whose trace anomaly is zero on the fixed point. It implies that there the R-symmetry must be the anomaly-free $U(1)_{AF}$ symmetry. It gives the scaling dimensions of the following chiral operators:

$$D(M) = \frac{3}{2}R_{AF}(M) = 3\frac{N_f - N_c}{N_f}, \tag{15.455}$$

$$D(B) = D(\tau B) = \frac{3}{2}\frac{N_c(N_f - N_c)}{N_f}. \tag{15.456}$$

Unitarity restricts the scaling dimensions of the gauge invariant operators to be $D \geq 1$. If $D = 1$, the corresponding operator $\mathcal{O}$ satisfies the free equation of motion $\partial^2 \mathcal{O} = 0$. If $D > 1$, there are non-trivial interactions between the operators.

For the range $3N_c/2 < N_f < 3N_c$, the gauge invariant chiral operators M, B and τB satisfy the unitarity constraint with $D > 1$. Seiberg conjectured the existence of such a non-trivial fixed point for any value of $3N_c/2 < N_f < 3N_c$, at least for large N_c.

As $\frac{3}{2}(N_f - N_c) < N_f < 3(N_f - N_c)$, there is also a non-trivial fixed point in the magnetic theory. Seiberg's claim is that both theories flow to the same infrared fixed point [38].

15.4.16 $N = 2$ supersymmetry.

The supersymmetry algebra and its massless representations.

The $N = 2$ supersymmetry algebra, without central charge, is

$$\begin{aligned} \{Q_\alpha^{(I)}, \overline{Q}_{\dot\beta(J)}\} &= 2(\sigma^\mu)_{\alpha\dot\beta}P_\mu\delta^I_J\,, \\ \{Q_\alpha^{(I)}, Q_\beta^{(J)}\} &= 0 \end{aligned} \tag{15.457}$$

with $I, J = 1, 2$. The algebra (15.457) has a new symmetry. We can perform unitary rotations of the two supercharges $Q_\alpha^{(I)}$ that do leave the anti-commutator relations (15.457) invariant.

We have an $U(2)_R = U(1)_R \times SU(2)_R$ symmetry. The Abelian factor $U(1)_R$ corresponds to the familiar R-symmetry of supersymmetric theories that rotate the global phase of the supercharges $Q_\alpha^{(I)}$. With respect the $SU(2)_R$ group, the supercharges $Q_\alpha^{(I)}$ are in the doublet representation **2**.

As in massless $N = 1$ supersymmetric representations, half of the supercharges are realized as vanishing operators: $Q_2^{(I)} = 0$. We normalize the other two supercharges,

$$a_1^{(I)} = \frac{1}{2\sqrt{E}} Q_1^{(I)} \,, \tag{15.458}$$

which are an $SU(2)_R$ doublet. The massless $N = 2$ vector multiplet is a representation constructed from the Clifford vacuum $|1>$, which has helicity $\lambda = 1$ and is an $SU(2)_R$ singlet. From it we obtain two fermionic states, $|1/2>^{(I)} = (a^{(I)})^\dagger |1>$, and a scalar boson $|0> = (a^{(1)})^\dagger (a^{(2)})^\dagger |1>$. After CPT doubling we obtain the $N = 2$ vector multiplet:

$$\begin{pmatrix} \{ \ |1>, \ |-1>_{CPT} \ \} \\ \{ \ |\tfrac{1}{2}>^{(1)}, \ |-\tfrac{1}{2}>^{(1)}_{CPT} \ \} \quad \{ \ |\tfrac{1}{2}>^{(2)}, \ |-\tfrac{1}{2}>^{(2)}_{CPT} \ \} \\ \{ \ |0>, \ |0>_{CPT} \ \} \end{pmatrix} \tag{15.459}$$

In terms of local fields we have: a vector A_μ (the gauge bosons of some gauge group G, since we consider massless representations), which is $SU(2)_R$ singlet; two Weyl spinors $\lambda^{(I)}$, the gauginos, arranged in an $SU(2)_R$ doublet; and a complex scalar ϕ, playing the role of the Higgs, a singlet of $SU(2)_R$ but in the adjoint of the gauge group G. These fields arrange as

$$\begin{pmatrix} & A_\mu & \\ & \swarrow & \\ \lambda^{(1)} & & \lambda^{(2)} \\ & \swarrow & \\ & \phi & \end{pmatrix} \tag{15.460}$$

where the arrows indicate the action of the supercharge $\overline{Q}^{(1)}_{\cdot\alpha}$. We can use a manifest $N = 1$ supersymmetry representation taking into account that the $N = 2$ vector multiplet is composed of an $N = 1$ vector multiplet $W_\alpha = (A_\mu, \lambda^{(1)})$ and an $N = 1$ chiral multiplet $\Phi = (\phi, \lambda^{(2)})$.

The massless $N = 2$ hypermultiplet is a representation constructed from a Clifford vacuum $|1/2>$, which is an $SU(2)_R$ singlet. The action of the two grassmanian operators a_α^I seems to produce the same particle content that the $N = 1$ chiral multiplet, but $|1/2> = |1/2, \mathbf{R}>$ is usually in some non-trivial representation $\mathbf{R}$ of a gauge group G. As $\mathbf{R} \to \overline{\mathbf{R}}$ under a CPT transformation, it forces to make the CPT doubling, and the $N = 2$ hypermultiplet is built from two $N = 1$ chiral multiplets in complex conjugate gauge group representations:

$$\begin{pmatrix} \{ \ |\tfrac{1}{2}, \mathbf{R}> , \ |-\tfrac{1}{2}, \overline{\mathbf{R}}>_{CPT} \ \} \\ \{ \ |0, \mathbf{R}>^{(1)}, \ |0, \overline{\mathbf{R}}>^{(1)}_{CPT} \} \quad \{ \ |0, \mathbf{R}>^{(2)}, \ |0, \overline{\mathbf{R}}>^{(2)}_{CPT} \} \\ \{ \ |-\tfrac{1}{2}, \mathbf{R}> , \ |\tfrac{1}{2}, \overline{\mathbf{R}}>_{CPT} \ \} \end{pmatrix} \tag{15.461}$$

Which represents the local fields

$$\begin{pmatrix} & \psi_q & \\ & \swarrow & \\ q & & \tau q^\dagger \\ & \swarrow & \\ & \overline{\psi}_{\tau q} & \end{pmatrix} \tag{15.462}$$

with the complex scalar fields $(q, \tau q^\dagger)$ in a doublet representation of $SU(2)_R$. In terms of $N = 1$ superfields we have one chiral superfield $Q = (q, \psi_q)$ in gauge representation $\mathbf{R}$ and another chiral superfield $\tau Q = (\tau q, \tau\psi_{\tau q})$ in gauge representation $\overline{\mathbf{R}}$. All the field in the hypermultiplet have spin $\leq 1/2$. Because of the CPT doubling, the matter content of extended supersymmetry $(N > 1)$ is always in vectorial representations of the gauge group.

The central charge and massive short representations.

As shown by Haag, Lapuszanski and Sohnius [40], the $N = 2$ supersymmetry algebra admits a central extension:

$$\begin{aligned} \{Q^a_\alpha, Q^b_\beta\} &= 2\sqrt{2}\epsilon_{\alpha\beta}\epsilon^{ab} Z\,, \\ \{\overline{Q}_{\dot{\alpha}a}, \overline{Q}_{\dot{\beta}b}\} &= 2\sqrt{2}\epsilon_{\dot{\alpha}\dot{\beta}}\epsilon_{ab}\overline{Z}\,. \end{aligned} \tag{15.463}$$

Since Z commutes with all the generators, we can fix it to be the eigenvalue for the given representation. Now, let us define:

$$a_\alpha = \frac{1}{2}\{Q^1_\alpha + \epsilon_{\alpha\beta}(Q^2_\beta)^\dagger\}\,, \tag{15.464}$$

$$b_\alpha = \frac{1}{2}\{Q^1_\alpha - \epsilon_{\alpha\beta}(Q^2_\beta)^\dagger\}\,. \tag{15.465}$$

Then, in the rest frame, the $N = 2$ supersymmetry algebra reduces to

$$\{a_\alpha, a^\dagger_\beta\} = \delta_{\alpha\beta}(\mathcal{M} + \sqrt{2}Z)\,, \tag{15.466}$$

$$\{b_\alpha, b^\dagger_\beta\} = \delta_{\alpha\beta}(\mathcal{M} - \sqrt{2}Z)\,, \tag{15.467}$$

with all other anti-commutators vanishing. Since all physical states have positive definite norm, it follows that for massless states, the central charge is trivially realized (*i.e.*„ $Z = 0$), as we used before. For massive states, this leads to a bound on the mass $\mathcal{M} \geq \sqrt{2}|Z|$. When $\mathcal{M} = \sqrt{2}|Z|$, the operators in (15.467) are trivially realized and the algebra resembles the massless case. The dimension of the representation is greatly reduced. For example, a reduced massive $N = 2$ multiplet has the same number of states as a massless $N = 2$ multiplet. Thus the representations of the $N = 2$ algebra with a central charge can be classified as either long multiplets (when $\mathcal{M} > \sqrt{2}|Z|$) or short multiplets (when $\mathcal{M} = \sqrt{2}|Z|$).

From (15.467) it is clear that the BPS states [9, 10] (which saturate the bound) are annihilated by half of the supersymmetry generators and thus belong to reduced representations of the supersymmetry algebra. An important consequence of this is that, for BPS states, the relationship between their charges and masses is dictated by supersymmetry and does not receive perturbative or nonperturbative corrections in the quantum theory. This is so because a modification of this relation implies that the states no longer belong to a short multiplet. On the other hand, quantum corrections are not expected to generate the extra degrees of freedom needed to convert a short multiplet into a long multiplet. Since there is no other possibility, we conclude that for short multiplets the relation $\mathcal{M} = \sqrt{2}|Z|$ is not modified either perturbatively or nonperturbatively.

15.4.17 $N = 2$ $SU(2)$ super Yang-Mills theory in perturbation theory.

The $N = 2$ Lagrangian.

The $N = 2$ superspace has two independent chiral spinors $\theta^{(I)}$, $I = 1, 2$. The $N = 2$ vector multiplet can be written in terms of $N = 2$ superspace by the $N = 2$ superfield $\Psi(x, \theta^{(I)})$

subject to the superspace constraints [41]:

$$\begin{aligned} \overline{\nabla}^{(I)}_{\cdot\alpha}\Psi &= 0\,, \\ \nabla_{(I)}\nabla_{(J)}\Psi &= \epsilon_{IK}\epsilon_{JL}\overline{\nabla}^{(K)}\overline{\nabla}^{(L)}\overline{\Psi}\,. \end{aligned} \tag{15.468}$$

where $\nabla_{(I)\alpha} = D_{(I)\alpha} + \Gamma_{(I)\alpha}$ is the generalized supercovariant derivative of the variable $\theta^{(I)}$, with $\Gamma_{(I)\alpha}$ the superconnection. The $N = 1$ superfields are connected to the $N = 2$ vector superfield through the equations:

$$\begin{aligned} \Psi|_{\theta^{(2)}=\overline{\theta}^{(2)}=0} &= \Phi(x, \theta^{(1)}, \overline{\theta}^{(1)})\,, \\ \nabla_{(2)\alpha}\Psi|_{\theta^{(2)}=\overline{\theta}^{(2)}=0} &= i\sqrt{2}W_\alpha(x, \theta^{(1)}, \overline{\theta}^{1})\,. \end{aligned} \tag{15.469}$$

It results that the renormalizable $N = 2$ super Yang-Mills Lagrangian is

$$\mathcal{L} = \frac{1}{8\pi}\mathrm{Im}\left(\tau \int d^2\theta^{(1)} d^2\theta^{(2)}\ \Psi^a\Psi^a\right) \tag{15.470}$$

with our old friend $\tau = \theta/2\pi + i4\pi/g^2$. In terms of $N = 1$ superspace, using (15.468) and (15.469), with $\theta \equiv \theta^{(1)}$, the Lagrangian is

$$\mathcal{L} = \frac{1}{8\pi}\mathrm{Im}\left(\tau \int d^2\theta\ W^\alpha W_\alpha\right) + \frac{1}{g^2}\int d^2\theta d^2\overline{\theta}\ \Phi^\dagger e^{-2V}\Phi\,. \tag{15.471}$$

It looks like $N = 1$ $SU(2)$ gauge theory with an adjoint chiral superfield Φ. The point is that the $1/g^2$ normalization in front of the kinetic term of Φ gives $N = 2$ supersymmetry. In fact, when we perform the remaining superspace integral in (15.471), we obtain a Lagrangian that looks like a Georgi-Glashow model with a complex Higgs triplet and the addition of a Dirac spinor $(\lambda^{(1)}, \overline{\lambda}^{(2)})$ in the adjoint also. This Lagrangian does not have all the gauge invariant renormalizable terms. $N = 2$ supersymmetry restricts the possible terms and gives relations between their couplings, such that at the end there are only the parameters g^2 and θ.

If we apply perturbation theory to the Lagrangian (15.470) we only have to perform a one loop renormalization. This is an indication that in $N = 2$ supersymmetry, holomorphy is not lost by radiative corrections. The reason is the following: We expained that the multi-loop renormalization of the coupling τ came from the generation of non-holomorphic factors $Z(\mu/\mu_0, g)$ in front of the complete $N = 1$ superspace integrals. At the level of the Lagrangian (15.471), consider the bare coupling τ_0 at scale μ_0 and integrate out the modes between μ_0 and μ. If we consider only the renormalizable terms, $N = 1$ supersymmetry gives us

$$\begin{aligned} \mathcal{L}_{ren} &= \frac{1}{8\pi}\mathrm{Im}\left(\tau(\mu/\Lambda)\int d^2\theta W^\alpha W_\alpha\right) \\ &+ Z\left(\frac{\mu}{\mu_0}, g_0\right)\frac{1}{g^2(\frac{\mu}{\Lambda})}\int d^2\theta d^2\overline{\theta}\quad \Phi^\dagger e^{-2V}\Phi \end{aligned} \tag{15.472}$$

where

$$\tau(\frac{\mu}{\Lambda}) = \frac{2i}{\pi}\ln\frac{\mu}{\Lambda} + \sum_{n=0}^{\infty} c_n \left(\frac{\Lambda}{\mu}\right)^{4n} \tag{15.473}$$

is the renormalized coupling constant at scale μ. We used the one-loop beta function of $N = 2$ $SU(2)$ gauge theory $b_0 = 4$ and the renormalization group invariant scale $\Lambda \equiv \mu_0 \exp(i\pi\tau_0/2)$. The dimensionless constants c_n are the coefficients of the n-instanton contribution $(\Lambda/\mu)^{4n} = \exp(-8\pi n/g^2(\mu) + i\theta(\mu)n)$.

If we compare with the $N = 2$ renormalizable Lagrangian (15.471) we derive that $Z(\mu/\mu_0, g_0) = 1$. Then, there is no Konishi anomaly and the one-loop renormalization of τ is all there is in perturbation theory.

The flat direction.

Unlike $N = 1$ super Yang-Mills, $N = 2$ super Yang-Mills theory includes a complex scalar ϕ in the adjoint of the gauge group. This scalar plays the role of a Higgs field through the potential derived from the Lagrangian (15.471),

$$V(\phi, \phi^\dagger) = \frac{1}{2g^2}[\phi^\dagger, \phi]^2 \,. \tag{15.474}$$

The supersymmetric minimum is obtained by the solution of

$$[\phi^\dagger, \phi] = 0 \,, \tag{15.475}$$

whose solution, up to gauge transformations, is $\phi = a\sigma^3$, with a an arbitrary complex number. This is our flat direction. Along it, the $SU(2)$ gauge group is spontaneously broken to the $U(1)$ subgroup. The $\Psi^\pm = \frac{1}{\sqrt{2}}(\Psi^1 \pm i\Psi^2)$ superfield components have $U(1)$ electric charge $Q_e = \pm g$, respectively, and they have the classical squared mass

$$\mathcal{M}_W^2 = 2|a|^2 \,. \tag{15.476}$$

The $\mathcal{A} \equiv \Psi^3$ superfield component remains massless. We know that the Lagrangian (15.470) admits semi-classical dyons with electric charge $Q_e = n_e g + \theta/2\pi$ and magnetic charge $Q_m = (4\pi/g)$, *i.e.*, the points $(1, n_e)$ in the charge lattice. They have the classical squared mass

$$\mathcal{M}^2(1, n_e) = 2|a|^2|n_e + \tau|^2 \,. \tag{15.477}$$

Physical masses are gauge invariant. We can use the gauge invariant parametrization of the moduli space in terms of the chiral superfield

$$U = \mathrm{tr}\Phi^2 \,, \tag{15.478}$$

and traslate the a-dependence in previous formulae by an u-dependence through the relation $u = \mathrm{tr}\langle\phi^2\rangle$. The classical relation is just $u = a^2/2$.

Then, semi-classical analysis gives $\mathcal{A}$ as the unique light degree of freedom. Only at $u = 0$ the full $SU(2)$ gauge symmetry is restored. How is this picture modified by the nonperturbative corrections?. The Seiberg-Witten solution answers this question [13] *.

15.4.18 The low energy effective Lagrangian.

The $N = 2$ vector superfield $\mathcal{A}$ is invariant under the unbroken $U(1)$ gauge transformations. At a scale of the order of the $\mathcal{M}_W$ mass, *i.e.*, of the order or $|u|^{1/2}$, the most general $N = 2$ Wilsonian Lagrangian, with two derivatives and four fermions terms, that can be constructed from the light degrees of freedom in $\mathcal{A}$ is

$$\mathcal{L}_{eff} = \frac{1}{4\pi}\mathrm{Im}\left(\int d^2\theta^{(1)} d^2\theta^{(2)}\ \mathcal{F}(\mathcal{A})\right) \tag{15.479}$$

with $\mathcal{F}$ a holomorphic function of $\mathcal{A}$, called the prepotential. We stress that the unique inputs to equation (15.479) are $N = 2$ supersymmetry and that $\mathcal{A}$ is a vector multiplet. We derive an immediate consequence of the general form of the effective Lagrangian (15.479): $N = 2$ supersymmetry prevents the generation of a superpotential for the $N = 1$ chiral superfield of $\mathcal{A}$. It means that the previously derived flat direction, parametrized by the arbitrary value $u = \mathrm{tr}\langle\phi^2\rangle$, is not lifted by nonperturbative corrections.

*Some additional reviews on the Seiberg-Witten solution are [42].

In terms of $N = 1$ superspace we have

$$\mathcal{L}_{eff} = \frac{1}{4\pi}\mathrm{Im}\left(\int d^2\theta \frac{1}{2}\tau(A)W^\alpha W_\alpha\right) + \int d^2\theta d^2\overline{\theta}\ K(A,\overline{A}), \tag{15.480}$$

where

$$\tau(A) = \frac{\partial^2\mathcal{F}}{\partial A^2}(A), \tag{15.481}$$

$$K(A,\overline{A}) = \mathrm{Im}\left(\frac{\partial\mathcal{F}}{\partial A}\overline{A}\right), \tag{15.482}$$

and A is the $N = 1$ chiral multiplet of $\mathcal{A}$.

The Wilsonian Lagrangian (15.480) is an Abelian gauge theory defined at some scale of order $\mathcal{M}_W \sim |u|^{1/2}$. Interaction terms come out after the expansion $A = a + \hat{A}$, with a the vacuum expectation value of the Higgs field, and $\hat{A}$ the quantum fluctuations of the chiral superfield. The matching at scale $|u|^{1/2}$ with the high energy $SU(2)$ theory is performed by the renormalization group:

$$\tau(u) = \frac{i}{\pi}\ln\frac{u}{\Lambda^2} + \sum_{n=0}^{\infty} c_n\left(\frac{\Lambda^2}{u}\right)^{2n}. \tag{15.483}$$

Observe that the phase of the dimensionless quotient u/Λ^2 plays the role of the bare θ_0 angle. If we are able to know the relation between the u and a variables, *i.e.*, the function $u(a)$, we can replace it into (15.483) to obtain $\tau(a)$. Integrating twice in the variable a we obtain the prepotential

$$\mathcal{F}(a) = \frac{i}{2\pi}a^2\ln\frac{a^2}{\Lambda^2} + a^2\sum_{n=1}^{\infty}\mathcal{F}_k\left(\frac{\Lambda}{a}\right)^{4k}. \tag{15.484}$$

If we look at the terms of the Lagrangian (15.480) proportional to the dimensionless constant $\mathcal{F}_n$, they correspond to the effective interaction terms created by the n-instanton contribution, as expected. For $a \to \infty$, the instanton contributions go to zero. This is an expected result, since at $a \to \infty$ the matching takes place at weak coupling due to asymptotic freedom. In this region perturbation theory is applicable and we can believe the semi-classical relation, $u \sim a^2/2$.

15.4.19 BPS bound and duality.

The $N = 2$ supersymmetry algebra gives the mass bound

$$\mathcal{M} \geq \sqrt{2}|Z|, \tag{15.485}$$

with Z the central charge. The origin of the central charge is easy to understand: the supersymmetry charges Q and $\overline{Q}$ are space integrals of local expressions in the fields (the time component of the super-currents). In calculating their anti-commutators, one encounters surface terms which are normally neglected. However, in the presence of electric and magnetic charges, these surface terms are non-zero and give rise to a central charge. When one calculates the central charge that arises from the classical Lagrangian (15.470) one obtains [43]

$$Z = ae(n + m\tau), \tag{15.486}$$

so that $\mathcal{M} \geq \sqrt{2}|Z|$ coincides with the Bogomol'nyi bound (15.354).

But the equation (15.486) is a classical result. The effective Lagrangian (15.479) includes all the nonperturbative quantum corrections of the higher modes. To get their contribution to the BPS bound, we just have to compute the central charge that is derived from the effective Lagrangian (15.479). The result is

$$Z(n_m, n_e) = n_e a + n_m a_D \,, \tag{15.487}$$

for a supermultiplet located in the charge lattice at (n_m, n_e). We have defined the a_D function

$$a_D \equiv \frac{\partial \mathcal{F}}{\partial a}(a) \,. \tag{15.488}$$

This function plays a crucial role in duality. Observe that under the $SL(2, \mathbf{Z})$ transformation $M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ of the charge lattice,

$$(n_m, n_e) \to (n_m, n_e) M^{-1} \,, \tag{15.489}$$

the invariance of the central charge demands

$$\begin{pmatrix} a_D \\ a \end{pmatrix} \to M \begin{pmatrix} a_D \\ a \end{pmatrix} \,. \tag{15.490}$$

Its action on the effective gauge coupling $\tau = \partial a_D / \partial a$ is

$$\tau \to \frac{\alpha \tau + \beta}{\gamma \tau + \delta} \,. \tag{15.491}$$

The S-transformation, that interchanges electric with magnetic charges, makes

$$\begin{aligned} a_D &\to a \,, \\ a &\to -a_D \,. \end{aligned} \tag{15.492}$$

Then, a_D is the dual scalar photon, that couples locally with the monopole $(1, 0)$ through the dual gauge coupling $\tau_D = -1/\tau$.

From (15.481) and (15.482), we see that $\mathrm{Im}\tau(a)$ is the Kähler metric of the Kähler potential $K(a, \overline{a})$,

$$d^2 s = [\mathrm{Im}\tau(a)] da d\overline{a} \,. \tag{15.493}$$

Physical constraints demands the metric be positive definite, $\mathrm{Im}\tau > 0$. However, if $\tau(a)$ is globally defined the metric cannot be positive definite as the harmonic function $\mathrm{Im}\tau(a)$ cannot have a minimum. This indicates that the above description of the metric in terms of the variable a must be valid only locally. In the weak coupling region, $|u|\hat{g}|\Lambda|$, where $\tau(a) \sim (2i/\pi)\ln(a/\Lambda)$, we have that $\mathrm{Im}\tau(a) > 0$, but for $a \sim \Lambda$, when the theory is at strong coupling and the nonperturbative effects become important, the perturbative result does not give the correct physical answer. Two things should happen: the instanton corrections must secure the positivity of the metric and physics must be described in terms of a new local variable a'. Which is this new local variable? If we do not want to change the physics, the change of variables must be an isometry of the Kähler metric (15.493). In terms of the variables (a_D, a) the Kähler metric is

$$d^2 s = \mathrm{Im}(da_D d\overline{a}) = -\frac{i}{2}(da_D d\overline{a} - da d\overline{a}_D) \,, \tag{15.494}$$

The complete isometry group of (15.494) is $\begin{pmatrix} a_D \\ a \end{pmatrix} \to M \begin{pmatrix} a_D \\ a \end{pmatrix} + \begin{pmatrix} p \\ q \end{pmatrix}$ with $M \in SL(2, \mathbf{R})$ and $p, q \in \mathbf{R}$. But the invariance of the central charge puts $p = q = 0$ * and the Dirac quantization condition restricts $M \in SL(2, \mathbf{Z})$. We arrive to an important result: in some region of the moduli space we have to perform an electric-magnetic duality transformation.

*In $N = 2$ SQCD with massive matter, the central charge allows to have $p, q \neq 0$ [44].

15.4.20 Singularities in the moduli space.

As Imτ cannot be globally defined on the u plane, there must be some singularities u_i indicating the multivaluedness of $\tau(u)$. If we perform a loop arround a singularity u_i, there is a non-trivial monodromy action M_i on $\tau(u)$. This action should be an isometry of the Kähler metric, if we do not want to change the physics. It implies that the monodromies M_i are elements of the $SL(2, \mathbf{Z})$ group.

In fact, we have found already one non-trivial monodromy because of the perturbative contributions. The multivalued logarithmic dependence of τ gives the monodromy. For $u \sim \infty$, $\tau \sim (i/\pi)\ln(u/\Lambda^2)$. In that region, the loop $u \to e^{2\pi i}u$ applied on $\tau(u)$ gives

$$\tau \to \tau - 2\,. \tag{15.495}$$

Its associated monodromy is

$$M_\infty = \begin{pmatrix} -1 & 2 \\ 0 & -1 \end{pmatrix} = PT^{-2}\,. \tag{15.496}$$

which acts on the variables (a_D, a) as

$$a_D \to -a_D + 2a\,, \tag{15.497}$$

$$a \to -a\,. \tag{15.498}$$

As it should be, the monodromy is a symmetry of the theory. T^{-2} just shifts the θ parameter by -4π, and P is the action of the Weyl subgroup of the $SU(2)$ gauge group. Then, the monodromy at infinity M_∞ leaves the a variable invariant (up to a gauge transformation).

The monodromy at infinity means there must be some singularity in the u plane. How many singularities?. We know that the anomalous $U(1)_R$ symmetry is broken by instantons, and that there is an unbroken $\mathbf{Z}_8$ subgroup because the one-instanton sector has eight fermionic zero modes. The $U = \mathrm{tr}\,\Phi^2$ operator has R-charge four. It means that the $u \to -u$ symmetry is spontaneously broken, leading to equivalent physical vacua. Then, if u_0 is a singular point, $-u_0$ must be also another singular point.

Let us assume that there is only one singularity. If this were the situation, the monodromy group would be Abelian, generated only by the monodromy at infinity. From the monodromy invariance of the variable a under M_∞, we would have that a is a good variable to describe the physics of the whole moduli space. This is in contradiction with the holomorphy of $\tau(a)$.

Seiberg and Witten made the assumption that there are only two singularities, which they normalized to be $u_1 = \Lambda^2$ and $u_2 = -\Lambda^2$. This assumption leads to a unique and elegant solution that passes many tests.

15.4.21 The physical interpretation of the singularities.

The most natural physical interpretation of singularities in the u plane is that some additional massless particles appear at the singular point $u = u_0$.

The particles will arrange in some $N = 2$ supermultiplet and will be labeled by some quantum numbers (n_m, n_e). If the massless particle is purely electric, the Bogomol'nyi bound implies $a(u_0) = 0$. It would mean that the W-bosons become massless at u_0 and the whole $SU(2)$ gauge symmetry is restored there. It would imply the existence of a non-Abelian infrared fixed point with $\langle \mathrm{tr}\phi^2 \rangle \neq 0$. By conformal invariance, the scaling dimension of the operator $\mathrm{tr}\phi^2$ at this infrared fixed point would have to be zero, *i.e.*, it would have to be the identity operator. It is not possible since $\mathrm{tr}\phi^2$ is odd under a global symmetry.

Then, the particles that become massless at the singular point u_0 are arranged in an $N = 2$ supermultiplet of spin $\leq 1/2$. The possibilities are severely restricted by the structure of $N = 2$ supersymmetry: the multiplet must be an hypermultiplet that saturates the BPS bound. As we have derived that we should have $a \neq 0$ for all the points of the moduli space, the singular BPS state must have a non-zero magnetic charge.

Near its associated singularity, the light $N = 2$ hypermultiplet is a relevant degree of freedom to be considered in the low energy Lagrangian. The coupling to the massless photon of the unbroken $U(1)$ gauge symmetry has to be local. Therefore, we apply a duality transformation to describe the relevant degree of freedom (n_m, n_e) as a purely electric state $(0, 1)$,

$$(0,1) =) n_m, n_e) N^{-1} \,, \tag{15.499}$$

with N the appropiate $SL(2, \mathbf{Z})$ transformation. The dual variables are the good local variables near the u_0 singularity. It implies that the monodromy matrix must leave invariant the singular state (n_m, n_e). This constraint plus the $U(1)$ β-function give the monodromy matrix

$$M(n_m, n_e) = \begin{pmatrix} 1 + 2n_m n_e & 2n_e^2 \\ -2n_m^2 & 1 - 2n_e n_m \end{pmatrix} . \tag{15.500}$$

In fact, in terms of the local variables,

$$\begin{pmatrix} a'_D \\ a' \end{pmatrix} = N \begin{pmatrix} a_D \\ a \end{pmatrix} , \tag{15.501}$$

the monodromy matrix is just T^2. This result can be understood as follows: The renormalizable part of the low energy Lagrangian is just $N = 2$ QED with one light hypermultiplet with mass $\sqrt{2}|a'| = \sqrt{2}|n_m a_D + n_e a|$. It has a trivial infrared fixed point, and the theory is weakly coupled at large distances. Perturbation theory gives

$$\tau' \simeq -\frac{i}{\pi} \ln a' . \tag{15.502}$$

On the other hand, by the monodormy invariance of a', we have $a'(u) \simeq c_0(u - u_0)$, this gives the monodromy matrix T^2: $\tau' \to \tau' + 2$.

With all the monodromies taken in the counter clockwise direction, and the monodromy base point chosen in the negative imaginary part of the complex u plane, we have the topological constraint

$$M_{-\Lambda^2} M_{\Lambda^2} = M_\infty . \tag{15.503}$$

If we use the expression (15.500) for the monodromies $M_{\pm\Lambda^2}$ and that $M_\infty = PT^{-2}$, (15.503) implies that the magnetic charge of the singular states must be ± 1. Then, they exist semi-classically and are continuousy connected with the weak coupling region. Moreover, if the state $(1, n_e)$ becomes massless at $u = \Lambda^2$, then (15.503) gives the massless state $(1, n_e - 1)$ at $u = -\Lambda^2$. It is consistent with the action of the spontaneously broken symmetry $u \to -u$, since by the expression of $\tau(u)$ in (15.483) we have that $\theta_{eff}(-\Lambda^2) = 2\pi \mathrm{Re}(\tau(-\Lambda^2)) = 2\pi$, and by the Witten effect gives the same physical electric charge to the massless states at $u = \pm\Lambda^2$.

Seiberg and Witten took the simplest solution: a purely magnetic monopole $(1, 0)$ *

becomes massless at $u = \Lambda^2$. With our chosen monodromy base point, the state with quantum numbers $(1, -1)$ has vanishing mass at $u = -\Lambda^2$.

*Observe that by Witten effect, the shift $\theta \to \theta + 2\pi n$ transforms $(1, 0) \to (1, n)$. There is a complete democracy between the semi-classical stable dyons.

15.4.22 The Seiberg-Witten solution.

The inputs.

After this long preparation, we can present the solution of the model. The moduli space is the compactified u-plane punctured at $u = \Lambda^2, -\Lambda^2, \infty$. These singular points generate the monodromies:

$$M_{\Lambda^2} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} ,$$
$$M_{-\Lambda^2} = \begin{pmatrix} -1 & 2 \\ -2 & 3 \end{pmatrix} ,$$
$$M_{\infty} = \begin{pmatrix} -1 & 2 \\ 0 & -1 \end{pmatrix} , \tag{15.504}$$

which act on the holomorphic function $\tau(u)$ by the corresponding modular transformations. Physically, the function $\tau(u)$ is the effective coupling at the vacuum u and its asymptotic behavior near the punctured points $u = \Lambda^2, -\Lambda^2, \infty$, is known.

The geometrical picture.

A torus is a two dimensional compact Riemann surface of genus one. Topologically it can be described by a two dimensional lattice with complex periods ω and ω_D. The construction is the following: a point z in the complex plane is identifyed with the points $z+\omega$ and $z+\omega_D$ (with the convention $\mathrm{Im}(\omega_D/\omega) > 0$), to get the topology of a torus. Then, the $SL(2,\mathbf{Z})$ transformations

$$\begin{pmatrix} \omega_D \\ \omega \end{pmatrix} \to M \begin{pmatrix} \omega_D \\ \omega \end{pmatrix} \tag{15.505}$$

leave invariant the torus. If we rescale the lattice with $1/\omega$, the torus is characterized just by the modulus

$$\tau \equiv \frac{\omega_D}{\omega} ,$$

up to $SL(2,\mathbf{Z})$ transformations,

$$\tau \sim \frac{\alpha\tau + \beta}{\gamma\tau + \delta} .$$

Algebraically the torus can be described by a complex elliptic curve

$$y^2 = 4(x - e_1)(x - e_2)(x - e_3) . \tag{15.506}$$

The toric structure arises because of the two Riemman sheets in the x plane joined through the two branch cuts going from e_1 to e_2 and e_3 to infinity (see fig. 2).

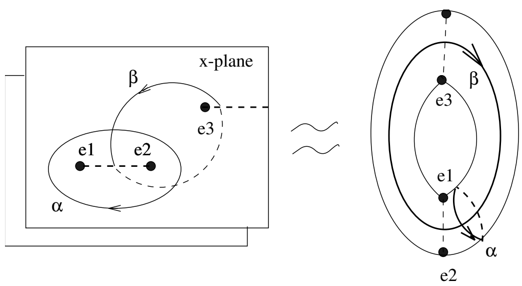


Figure 2: The elliptic curve (15.506) gives the topology of a torus.

The lattice periods are obtained by integrating the Abelian differential of first kind dx/y along the two homologically non-trivial one-cycles α and β, with intersection number $\beta \cdot \alpha = 1$,

$$\omega_D = \oint_\beta \frac{dx}{y}\,,$$
$$\omega = \oint_\alpha \frac{dx}{y}\,. \tag{15.507}$$

They have the property that $\mathrm{Im}\tau > 0$.

The Physical connection with $N = 2$ super Yang-Mills.

The breakthough of Seiberg and Witten for the solution of the model was the identification of the complex effective coupling $\tau(u)$ at a given vacuum u with the modulus of a u-dependent torus. At any point u of the moduli space, they associated an elliptic curve

$$y^2 = 4\prod_{i=1}^{3}(x - e_i(u))\,, \tag{15.508}$$

with its lattice periods given by (15.507).

The identification of the physical coupling $\tau(u) = \partial a_D/\partial a$ with the modulus $\tau_u = \omega_D(u)/\omega(u)$ of the elliptic curve (15.508),

$$\tau(u) = \frac{\partial a_D/\partial u}{\partial a/\partial u} = \frac{\oint_\beta dx/y}{\oint_\alpha dx/y} = \tau_u\,, \tag{15.509}$$

leads to the formulae:

$$a_D = \oint_\beta \lambda(u)\,, \tag{15.510}$$

$$a = \oint_\alpha \lambda(u)\,, \tag{15.511}$$

where $\lambda(u)$ is an Abelian differential with the property that

$$\frac{\partial\lambda}{\partial u} = f(u)\frac{dx}{y} + dg\,. \tag{15.512}$$

Then, the solution of the problem is reduced to finding the family of elliptic curves (15.508) and the holomorphic function $f(u)$. The conditions at the begining of this section fix a unique solution. The family of elliptic curves is determined by the monodromy group generated by the monodromy matrices. The matrices (15.504) generate the group $\Gamma(2)$, the subgroup of $SL(2, \mathbf{Z})$ consisting of matrices congruent to the identity modulo 2. It gives the elliptic curves

$$y^2 = (x^2 - \Lambda^4)(x - u)\,. \tag{15.513}$$

Finally, the function $f(u)$ is determined by the asymptotic behavior of (a_D, a) at the singular points. The answer is $f = -\sqrt{2}/4\pi$.

15.4.23 Breaking $N = 2$ to $N = 1$. Monopole condensation and confinement.

In this section we will exhibit an explicit realization of the confinement mechanism envisaged by Mandelstam [45] and 't Hooft's through the condensation of light monopoles.

In the $N = 2$ model, we have found points in the moduli space where the relevant light degrees of freedom are magnetic particles. Since we have the exact solution of the low energy $N = 2$ model, it would be nice to answer in which phase the dynamics of the model, or controlable deformations of it, locates the vacuum.

For the $N = 2$ model we already know from section XVIII that $N = 2$ supersymmetry does not allow the generation of a superpotential just for the $N = 1$ chiral multiplet of the $N = 2$ vector multiplet. It means that the theory is always in an Abelian Coulomb phase. The exact solution of the model allowed us to know which are all the instanton corrections to the low energy Lagrangian. Remarkably enough, the instanton series admits a resumation in terms of magnetic variables.

To go out of the Coulomb branch, we need a superpotential for the chiral superfield Φ. In [13] an explicit mass term for the chiral superfield was added in the bare Lagrangian,

$$\mathcal{W}_{tree} = m \,\mathrm{tr}\, \Phi^2 \,. \tag{15.514}$$

It breaks $N = 2$ to $N = 1$ supersymmetry. At low energy, we will have an effective superpotential $\mathcal{W}(m, M, \tau M, A_D)$. Once again, holomorphy of the superpotential and selection rules from the symmeries will fix the exact form of $\mathcal{W}$. In terms of $N = 1$ superspace, only the subgroup $U(1)_J \subset SU(2)_R$ is manifestly a symmetry. It is a non-anomalous R-symmetry (rotates the complex phases of $\theta^{(I)}$, $I = 1, 2$, in opposite directions.). The corresponding charge of Φ is zero. As superpotentials should have charge two, from (15.514) we derive that the parameter $m \neq 0$ breaks the $U(1)_J$ symmetry by two units. On the other hand, the $N = 1$ chiral superfields M and τM are in an $N = 2$ hypermultiplet and therefore, both have charge one. Imposing that $\mathcal{W}$ is a regular function at $m = \tau M M = 0$, we find that it is of the form $\mathcal{W} = m f_1(A_D) + \tau M M f_2(A_D)$. For $m \to 0$, the effective superpotential flows to the tree level superpotential (15.514) plus the term $\sqrt{2} A_D \tau M M$. As the functions f_1 and f_2 are independent of m, we obtain the exact result

$$\mathcal{W} = \sqrt{2} A_D \tau M M + m U(A_D) \,. \tag{15.515}$$

We found what we were looking for: an exact effective superpotential with a term which depends only of the $N = 1$ chiral composite operator U. It presumely will remove the flat direction. The $N = 2$ to $N = 1$ breaking makes no loger valid the hiden $N = 2$ holomorphy in the Kähler potential $K(A, \overline{A})$. But as long as there is an unbroken supersymmetry, the vacuum configuration corresponds to the solution of the equations

$$d\mathcal{W} = 0 \,, \tag{15.516}$$

$$D = |M|^2 - |\tau M|^2 = 0 \,. \tag{15.517}$$

From the exact solution we know that $du/da_D \neq 0$ at $a_D = 0$. Thus (up to gauge transformations)

$$\begin{aligned} M = \tau M &= \left(-m u'(0)/\sqrt{2}\right)^{1/2} \,, \\ a_D &= 0 \,. \end{aligned} \tag{15.518}$$

Expanding around this vacuum we find:

i) There is a mass gap of the order $(m\Lambda)^{1/2}$.

ii) The objects that condense are magnetic monopoles. There are electric flux tubes with a non-zero string tension of the order of the mass gap, that confines the electric charges of the $U(1)$ gauge group.

The spontaneously broken symmetry $u \to -u$ carries the theory to the ‘dyon region’, with the local variable $a_D - a$. The perturbing superpotential there, $mU(a_D - a)$, also produces the condensation of the 'dyon' with physical electric charge zero at the point $a_D - a = 0$. Then, we have two physically equivalent vacua, related by an spontaneously broken symmetry, in agreement with the Witten index of $N = 1$ $SU(2)$ gauge theory.

15.4.24 Breaking $N = 2$ to $N = 0$.

When the $N = 2$ theory is broken to the $N = 1$ theory through the decoupling of the chiral superfield Φ in the adjoint, we have seen that the mechanism of confinement takes place because of the condensation of a magnetic monopole. The natural question is if this results can be extended to non supersymmetric gauge theories.

The $N = 1, 2$ results were based on the use of holomorphy; the question is whether the properties connected with holomorphy can be extended to the $N = 0$ case. The answer is positive provided supersymmetry is broken via soft breaking terms.

The method is to promote some couplings in the supersymmetric Lagrangian to the quality of frozen superfields, called spurion superfields. We could think they correspond to some heavy degrees of freedom which at low energies have been decoupled. Their trace is only through their vacuum expectation values appearing in the Lagrangian and are parametrized by the spurion superfields [46].

In the $N = 2$ theory we will promote some couplings to the status of spurion superfields. The property of holomorphy in the prepotential will be secured if the introduced spurions are $N = 2$ vector superfields [14, 15]
*.

In the bare Lagrangian of the $N = 2$ $SU(2)$ gauge theory (15.470), there is only one parameter: τ_0. The $N = 2$ softly broken theory is obtained by the bare prepotential

$$\mathcal{F}_0 = \frac{1}{\pi} \mathcal{S} \mathcal{A}^a \mathcal{A}^a \, , \tag{15.519}$$

where S is an dimensionless $N = 2$ vector multiplet whose scalar component gives the bare coupling constant, $s = \frac{\pi}{2}\tau_0$. The factor of proporcionality is related with the one loop coefficient of the beta function, such that $\Lambda = \mu_0 \exp(is)$. Inspired by String Theory, we call S the dilaton spurion. The source of soft breaking comes from the non vanishing auxiliary fields, F_0 and D_0, in the dilaton spurion S.

The tree level mass terms arising from the softly broken bare Lagrangian (15.519) are the following: the W-bosons get a mass term by the usual Higgs mechanism, with the mass square equal to $2|a|^2$; the photon of the unbroken $U(1)$ remains massless; the gauginos get a mass square $\mathcal{M}_{1/2}^2 = (|F_0|^2 + D_0^2/2)(4\text{Im}s)^{-1}$; all the scalar components, except the real part of ϕ^3 which do not have a bare mass term, get a square mass $\mathcal{M}_0^2 = 4\mathcal{M}_{1/2}^2$.

At low energy, *i.e.*, at scales of the order $|u|^{1/2} \sim \Lambda$, the Wilsonian effective Lagrangian up to two derivatives and four fermions terms is given by the effective prepotential $\mathcal{F}(a, \Lambda)$ found in the $N = 2$ model, but with the difference that the bare coupling constant is replaced by the dilaton spurion, *i.e.*, $\Lambda \to \mu_0 \exp(i\mathcal{S})$. Then, the prepotential depends on two vector multiplets and the effective Lagrangian becomes

$$\begin{aligned} \mathcal{L} = & \frac{1}{4\pi} \text{Im} \left(\int d^4\theta \frac{\partial \mathcal{F}}{\partial A^i} \overline{A}^i + \int d^2\theta \frac{1}{2} \frac{\partial^2 \mathcal{F}}{\partial A^i \partial A^j} W^i W^j \right) \\ & + \mathcal{L}_{HM} \, . \end{aligned} \tag{15.520}$$

with $A^i = (S, A)$ and $\mathcal{L}_{HM}$ the $N = 2$ Lagrangian that includes the monopole hypermultiplet. Observe that the dilaton spurion do not enter in the Lagrangian of the hypermultiplets, in

*Soft breaking of $N = 1$ SQCD has been studied in [47].

agreement with the $N = 2$ non-renormalization theorem of [27]. The low energy couplings are determined by the 2×2 matrix

$$\tau_{ij}(a,s) = \frac{\partial^2 \mathcal{F}}{\partial a^i \partial a^j}\,. \tag{15.521}$$

The supersymmetry breaking generates a non-trivial effective potential for the scalar fields,

$$\begin{aligned} V_{eff} &= \left(b_{00} - \frac{b_{01}^2}{b_{11}}\right)\left(|F_0|^2 + \frac{1}{2}D_0^2\right) \\ &+ \frac{b_{01}}{b_{11}}\left[\sqrt{2}(F_0 m\tau m + \overline{F}_0\overline{m\tau m}) + D_0)|m|^2 - |\tau m|^2)\right] \\ &+ \frac{1}{2b_{11}}(|m|^2 + |\tau m|^2)^2 + 2|a|^2(|m|^2 + |\tau m|^2)\,, \end{aligned} \tag{15.522}$$

where we have defined $b_{ij} = (4\pi)^{-1}\mathrm{Im}\tau_{ij}$. m and τm are the scalar components of the chiral superfields M and τM of the monopole hypermultiplet, respectively. Observe that the first line of (15.522) is independent of the monopole degrees of freedom. To be sure that such quantity gives the right amount of energy at any point of the moduli space, where different local descriptions of the physics are necessary, it must be duality invariant. This is the case for any $SL(2,\mathbf{Z})$ transformation.

The auxiliary fields of the dilaton spurion are in the adjoint representation of the group $SU(2)_R$ and have $U(1)_R$ charge two. We can consider the situation of $D_0 = 0, F_0 = f_0 > 0$ without any loss of generality, since it is related with the case of $D_0 \neq 0$ and complex F_0 just by the appropiate $SU(2)_R$ rotation.

We have to be careful with the validity of our approximations. Because of supersymmetry, the expansion in derivatives is linked with the expansion in fermions and the expansion in auxiliary fields. The exact solution of Seiberg and Witten is only for the first terms in the derivative expansion of the effective Lagrangian, in particular up to two derivatives. At the level of the softly broken effective Lagrangian, the exact solution of Seiberg and Witten only gives us the terms at most quadratic in the supersymmetry breaking parameter f_0. The expansion is performed in the dimensionless parameter f_0/Λ. Our ignorance on the higher derivative terms of the effective Lagrangian is traslated into our ignorance the terms of $\boldsymbol{O})(f_0/\Lambda)^4)$. Hence our results are reliable for small values of f_0/Λ, and this is far from the supersymmetry decoupling limit $f_0/\Lambda \to \infty$.

But for moderate values of the supersymmetry breaking parameter, the effective Lagrangian (15.520) gives the large distance physics of a non-supersymmetric gauge theory at strong coupling. If we minimize the effective potential (15.522) with respect to the monopoles, we obtain the energy of the vacuum u

$$\begin{aligned} V_{eff}(u) &= \left(b_{00}(u) - \frac{b_{01}^2(u)}{b_{11}(u)}\right)|F_0|^2 \\ &- \frac{2}{b_{11}(u)}\rho^4(u)\,, \end{aligned} \tag{15.523}$$

where $\rho(u)$ is a positive function that gives the monopole condensate at u

$$|m|^2 = |\tau m|^2 = \rho^2(u) = \frac{|b_{01}|f_0}{\sqrt{2}} - b_{11}|a|^2 > 0 \tag{15.524}$$

or $m = \tau m = \rho(u) = 0$ if $|b_{01}|f_0 < \sqrt{2}b_{11}|a|^2$.

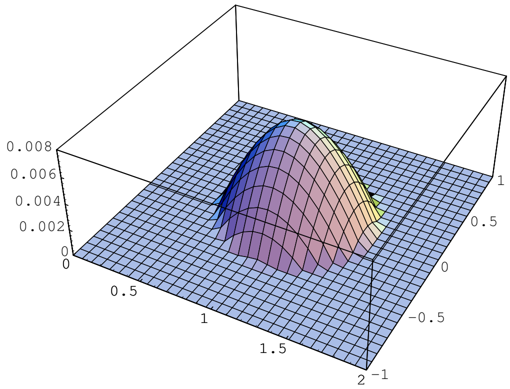


Figure 3: The monopole condensate ρ^2, at the monopole region $u \sim \Lambda^2$, for $f_0 = \Lambda/10$.

Notice that b_{11} diverges logarithmically at the singularities $u = \pm\Lambda^2$, but the corresponding local variable a vanishes linearly at $u = \pm\Lambda^2$. It implies that $b_{11}|a|^2 \to 0$ for $u \to \pm\Lambda^2$. It can be shown that the Seiberg-Witten solution gives $b_{01} \sim \Lambda/8\pi$ for $u \sim \Lambda$. It means that the monopole condenses at the monopole region (see fig. 3), since from the expression of the effective potential (15.523), such condensation is energetically favoured. If we look at the dyon region, we find that $b_{01} \to 0$ for $u \to -\Lambda^2$. Numerically, there is a very small dyon condensate without any associated minimum in the effective potential in that region. On the other hand, there is a clear absolute minimum in the monopole region (see fig. 4). The different behaviors of the broken theory under the transformation $u \to -u$ is an expected result if we take into account that $f_0 \neq 0$ breaks explicitly the $U(1)_R$ symmetry.

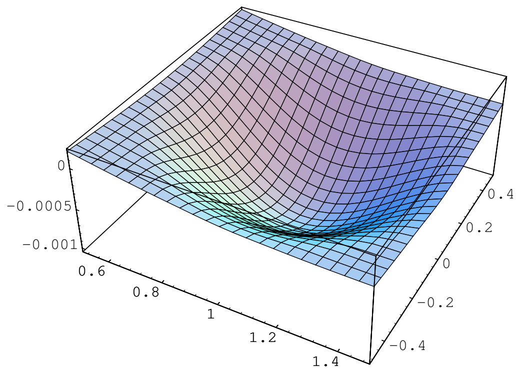


Figure 4: The effective potential $V_{eff}(u)$ (15.523), at the monopole region $u \sim \Lambda^2$, for $f_0 = \Lambda/10$.

The softly broken theory selects a unique minimum at the monopole region, with a non vanishing expectation value for the monopole. The theory confines and has a mass gap or order $(f_0\Lambda)^{1/2}$.

15.4.25 String Theory in perturbation theory.

String Theory is a multifaceted subject. In the sixties strings were first introduced to model the dynamics of hadron dynamics. In section VII we described the confining phase as the dual Higgs phase, where magnetic degrees of freedom condense. The topology of the gauge group allows the existence of electric vortex tubes, ending on quark-antiquark bound states. The transverse size of the electric tubes is of the order of the compton wave length of the 'massive' W-bosons. At large distances, these electric tubes can be considered as open strings with a quark and an anti-quark at their end points. This is the QCD string, with an string tension of the order of the characteristic length square of the hadrons, $\alpha' \sim (1\text{GeV})^{-2}$.

But the major interest in String Theory comes from being a good candidate for quantum gravity [48]. The macroscopic gravitational force includes an intrinsic constant, G_N, with dimensions of length square

$$G_N = l_p^2 = (1.6 \times 10^{-33}\text{cm})^2 \,. \tag{15.525}$$

In a physical process with an energy scale E for the fundamental constituents of matter, the strength of the gravitational interaction is given by the dimensionless coupling $G_N E^2$ to the graviton. This interaction can be neglected when the graviton probes length scales much larger than the Planck's size, $G_N E^2 \mathcal{L} 1$. The interaction is also non-renormalizable. From the point of view of Quantum Field Theory, it corresponds to an effective low energy interaction, with l_p the natural length scale at which the effects of quantum gravity become important. The natural suspicion is that there is new physics at such short distances, which smears out the interaction. The idea of String Theory is to replace the point particle description of the interactions by one-dimensional objects, strings with size of the order of the Planck's length $l_p \sim 10^{-33}$cm (see fig. 5). Such simple change has profound consequences on the physical behavior of the theory, as we will briefly review below. It is still not clear whether the stringy solution to quantum gravity should work. Because Planck's length scale is so small, up to now String Theory is only constructed from internal consistency. But it is at the moment the best candidate we have. Let us quickly review some of the major implications of String Theory, derived already at perturbative level.

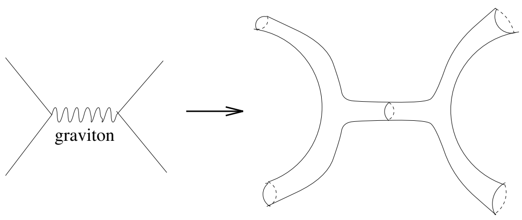


Figure 5: The point particle graviton interchange is replaced by the smeared string interaction.

The first important consequence of String Theory is the existence of vibrating modes of the string. They correspond to the physical particle spectrum. For phenomenology the relevant part comes from the massless modes, since the massive modes are excited at energies of the order of the Planck's mass l_p^{-1}. At low energies all the massive modes decouple and we end with an effective Quantum Field Theory for the massless modes. In the massless spectrum of the closed string, there is a particle of spin two. It is the graviton. Then String Theory includes gravity. If we know how to make a consistent and phenomenologically satisfactory quantum theory of strings, we have quantized gravity.

Up to now, String Theory is only well understood at the perturbative level. The field theory diagrams are replaced by two dimensional Riemann surfaces, with the loop expansion being performed by an expansion in the genus of the surfaces. It is a formulation of first quantization, where the path integral is weighed by the area of the Riemann surface and the external states are included by the insertion of the appropiate vertex operators (see fig. 6). The perturbative string coupling constant is determined by the vacuum expectation value of a massless real scalar field, called the dilaton, through the relation $g_s = \exp\langle s \rangle$. The thickening of Feynman diagrams into 'surface' diagrams improves considerably the ultraviolet behavior of the theory. String Theory is ultraviolet finite.

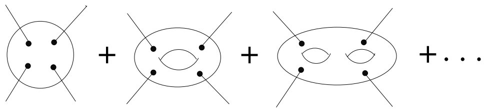


Figure 6: The preturbative loop expansion in String Theory is equivalent to expand in the number of genus of the Riemann surfaces.

The third important consequence is the introduction of supersymmetry. For the bosonic string, the lowest vibrating mode correponds to a tachyon. It indicates that we are performing perturbation theory arround an unestable minimum. Supersymmetry gives a very economical solution to this problem. In a supersymmetric theory the hamiltonian operator is positive semi-definite and the ground state has always zero energy. It is also very appealing from the point of view of the cosmological constant problem. Furthermore, supersymmetry also introduces fermionic degrees of freedom in the physical spectrum. If nature really chooses to be supersymmetric at sort distances, the big question is: How is supersymmetry dynamically broken? The satisfactory answer must include the observed low energy phenomena of the standard model and the vanishing of the cosmological constant. As a last comment on supersymmetry we will say that the Green-Schwarz formulation of the superstring action demands invariance under a world-sheet local fermionic symmetry, called κ-symmetry. It is only possible to construct κ-symmetric world-sheet actions if the number of spacetime symmetries is $N \leq 2$ (in ten spacetime dimensions).

The fourth important consequence is the prediction on the number of dimensions of the target space where the perturbative string propagates. Lorentz invariance on the target space or conformal invariance on the world-sheet fixes the number of spacetime dimensions (twenty-six for bosonic strings and ten for superstrings). As our low energy world is four dimensional, String Theory incorporates the Kaluza-Klein idea in a natural way. But again the one-dimensional nature of the string gives a quite different behavior of String Theory with respect to field theory. The dimensional reduction of a field theory in D spacetime dimensions is another field theory in $D-1$ dimensions. The effect of a non-zero finite radius R for the compactified dimension is just a tower of Kaluza-Klein states with masses n/R. But in String Theory, the string can wind m times around the compact dimension. This process gives a contribution to the momentum of the string proportional to the compact radius, mR/α'. These quantum states become light for $R \to 0$. The dimensional reduction of a String Theory in D dimensions is another String Theory in D dimensions. This is T duality [49].

The fifth important consequence comes from the cancellation of spacetime anomalies (gauge, gravitational and mixed anomalies). It gives only the following five anomaly-free superstring theories in ten spacetime dimensions.

The type IIA and type IIB string theories.

A type II string theory is constructed from closed superstrings with $N=2$ spacetime supersymmetries. The spectrum is obtained as a tensor product of a left- and right-moving world-sheet sectors of the closed string. Working in the light-cone gauge, the massless states of each sector are in the representation $\mathbf{8}_v \oplus \mathbf{8}_\pm$ of the little group $SO(8)$. The representations $\mathbf{8}_v$ and $\mathbf{8}_\pm$ are the vector representation and the irreducible chiral spinor representations of $SO(8)$, respectively.

The type IIA string theory corresponds to the choice of opposite chiralities for the spinorial representations in the left- and right-moving sectors,

$$\text{Type IIA}: \quad (\mathbf{8}_v \oplus \mathbf{8}_+) \otimes (\mathbf{8}_v \oplus \mathbf{8}_-)\,. \tag{15.526}$$

The bosonic massless spectrum is divided between the NS-NS fields:

$$\mathbf{8}_v \otimes \mathbf{8}_v = \mathbf{1} \oplus \mathbf{28} \oplus \mathbf{35}\,, \tag{15.527}$$

which corresponds to the dilaton s, the antisymmetric tensor $B_{\mu\nu}$ and the gravitation field $g_{\mu\nu}$, respectively, and the R-R fields:

$$\mathbf{8}_+ \otimes \mathbf{8}_- = \mathbf{8}_v \oplus \mathbf{56}, \tag{15.528}$$

which correspond to the light-cone degrees of freedom of the antisymmetric tensors A_μ and $A_{\mu\nu\rho}$, respectively. As the chiral spinors have opposite chiralities, in the vertex operators of the R-R fields only even forms appear, F_2 and F_4. The physical state conditions on the massless states give the following equations on these even forms:

$$dF = 0 \quad d \star F = 0\,, \tag{15.529}$$

with $\star F$ the Poincare dual $(10-n)$-form of the n-form F_n. These are the Bianchi identity and the equation of motion for a field strength. Their relation with the R-R fields is then $F_n = dA_{n-1}$. The Abelian field strengths F_n are gauge invariant, and since these are the fields that appear in the vertex operators, the fundamental strings do not carry RR charges.

The fermionic massless spectrum is given by the $NS-R$ and $R-NS$ fields:

$$\begin{aligned} \mathbf{8}_v \otimes \mathbf{8}_- &= \mathbf{8}_+ \oplus \mathbf{56}_-\,, \\ \mathbf{8}_+ \otimes \mathbf{8}_v &= \mathbf{8}_- \oplus \mathbf{56}_+\,. \end{aligned} \tag{15.530}$$

The $\mathbf{8}_\pm$ states are the two dilatini. The $\mathbf{56}_\pm$ states are the two gravitini, with a spinor and a vector index. Observe that the fermions have opposite chiralities, which prevent the type IIA theory from gravitational anomalies.

The Type IIB String Theory corresponds to the choice of the same chirality for the spinor representations of the left- and right-moving sector,

$$\text{Type IIB}: \quad (\mathbf{8}_v \oplus \mathbf{8}_+) \otimes (\mathbf{8}_v \oplus \mathbf{8}_+)\,. \tag{15.531}$$

The NS-NS fields are the same as for the type IIA string. The difference comes from the R-R fields:

$$\mathbf{8}_+ \otimes \mathbf{8}_+ = \mathbf{1} + \oplus \mathbf{28} \oplus \mathbf{35}_+\,. \tag{15.532}$$

They correspond, respectively, to the forms A_0, A_2 and A_4 (self-dual).

For the massless fermions there are two dilatini and two gravitini, but now all of them have the same chirality. In spite of it, the theory does not have gravitational anomalies [50].

Under spacetime compactifications, the type IIA and the type IIB string theories are unified by the T-duality symmetry. It is an exact symmetry of the theory already at the perturbative level and maps a type IIA string with a compact dimension of radius R to a type IIB string with radius α'/R.

The Type I string theory.

It is constructed from unoriented open and closed superstrings, leading only $N=1$ spacetime supersymmetry. The massless states are:

$$\text{Open}: \quad \mathbf{8}_v \otimes \mathbf{8}_+ \tag{15.533}$$

$$\begin{aligned} \text{Closed sym.}: \quad & [(\mathbf{8}_v \oplus \mathbf{8}_+) \otimes (\mathbf{8}_v \oplus \mathbf{8}_+)]_{\text{sym}} = \\ &= [\mathbf{1} \oplus \mathbf{28} \oplus \mathbf{35}]_{\text{bosonic}} \oplus [\mathbf{8}_- \oplus \mathbf{56}_-]_{\text{fermionic}}\,. \end{aligned} \tag{15.534}$$

The massless sector of the spectrum that comes from the unoriented open superstring (15.533) gives $N=1$ super Yang-Mills theory, with a gauge group $SO(N_c)$ or $USp(N_c)$ introduced by Chan-Paton factors at the ends of the open superstring. The sector coming from the unoriented closed string (15.534) gives $N=1$ supergravity. Cancellation of spacetime anomalies restricts the gauge group to $SO(32)$.

The $SO(32)$ and $E_8 \times E_8$ heterotic strings.

The heterotic string is constructed from a right-moving closed superstring and a left-moving closed bosonic string. Conformal anomaly cancellation demands twenty-six bosonic target space coordinates in the left-moving sector. The additional sixteen left-moving coordinates X_L^I, $I = 1, ..., 16$, are compactified on a T^{16} torus, defined by a sixteen-dimensional lattice, Λ_{16}, with some basis vectors $\{e_i^I\}$, $i = 1, ..., 16$. The left-moving momenta p_L^I live on the dual lattice $\tau\Lambda_{16}$. The mass operator gives an even lattice ($\sum_{I=1}^{16} e_i^I e_i^I = 2$ for any i). The modular invariance of the one-loop diagrams restricts the lattice to be self-dual ($\tau\Lambda_{16} = \Lambda_{16}$). There are only two even self-dual sixteen-dimensional lattices. They correspond to the root lattices of the Lie groups $SO(32)/Z_2$ and $E_8 \times E_8$.

For the physical massless states, the supersymmetric right-moving sector gives the factor $\mathbf{8}_v \otimes \mathbf{8}_+$, which together with the lattice points of length squared two of the left-moving sector, give an $N = 1$ vector multiplet in the adjoint representation of the gauge group $SO(32)$ or $E_8 \times E_8$.

There is also a T-duality symmetry relating the two heterotic strings.

15.4.26 D-branes.

Perturbation theory is not the whole history. In the field theory sections we have learned how much the nonperturbative effects could change the perturbative picture of a theory. In particular, there are nonperturbative stable field configurations (solitons) that can become the relevant degrees of freedom in some regime. In that situation it is convenient to perform a duality transformation to have an effective description of the theory in terms of these solitonic degrees of freedom as the fundamental objects.

What about the nonperturbative effects in String Theory?. Does String Theory incorporate nonperturbative excitations (string solitons)?. Are there also strong-weak coupling duality transformations in String Theory?. Before the role of D-branes in String Theory were appreciated, the answers to these three questions were not clear.

For instance, it was known, by the study of large orders of string perturbation theory, that the nonperturbative effects in string theory had to be stronger than in field theory, in the sense of being of the order of $\exp(-1/g_s)$ instead of order $\exp(-1/g_s^2)$ [51], but it was not known which were the nature of such nonperturbative effects.

With respect the existence of nonperturbative objects, the unique evidence came form solitonic solutions of the supergravity equations of motion which are the low energy limits of string theories. These objects were in general extended membranes in $p+1$ dimensions, called p-branes [52].

In relation to the utility of the duality transformation in String Theory, there is strong evidence of some string dualities [53]. There is for instance the $SL(2, \mathbf{Z})$ self-duality conjecture of the type IIB theory [54]. Under an S-transformation,the string coupling value g_s is mapped to the value $1/g_s$, and the NS-NS field $B_{\mu\nu}$ is mapped to the R-R field $A_{\mu\nu}$. Then, self-duality of type IIB demands the existence of an string with a tension scaling as g_s^{-1} and non-zero RR charge.

Dirichlet boundary conditions.

In open string theory, it is possible to impose two different boundary conditions at the ends of the open string:

$$\text{Neuman}: \quad \partial_\perp X^\mu = 0\,. \tag{15.535}$$

$$\text{Dirichlet}: \quad \partial_t X^\mu = 0\,. \tag{15.536}$$

An extended topological defect with $p+1$ dimensions is described by the following boundary conditions on the open strings:

$$\partial_{\perp} X^{0,1,\cdots p} = \partial_t X^{p+1,\cdots 9} = 0\,. \qquad (15.537)$$

We call it a D p-brane (for Dirichlet [55]), an extended (p+1)-dimensional object (located at $X^{p+1,\cdots 9} = \text{const}$) with the end points of open strings attached to it.

The Dirichlet boundary conditions are not Lorentz invariant. There is a momentum flux going from the ends of open strings to the D-branes to which they are attached. In fact, the quantum fluctuations of the open string endpoints in the longitudinal directions of the D-brane live on the world-volume of the D-brane. The quantum fluctuations of the open string endpoints in the transverse directions of the D-brane, makes the D-brane fluctuate locally. It is a dynamical object, characterized by a tension T_p and a RR charge μ_p. If $\mu_p \neq 0$, the world-volume of a p-brane will couple to the R-R $(p+1)$-form A_{p+1}.

Far from the D-brane, we have closed superstrings, but the world-sheet boundaries (15.537) relates the right-moving supercharges to the left-moving ones, and only a linear combination of both is a good symmetry of the given configuration. In presence of the D-brane, half of the supersymmetries are broken. The D-brane is a BPS state. In fact, in [56] it was shown that the D-brane tension arises from the disk and therefore that it scales as g_s^{-1}. This is the same coupling constant dependence as for BPS solitonic branes carrying RR charges [52].

The Dirichlet boundary condition becomes the Neuman boundary condition in terms of the T-dual coordinates, and vice versa. It implies that if we T-dualize a direction longitudinal to the world- volume of the D p-brane, it becomes a $(p-1)$-brane. Equally, if the T-dualized direction is transverse to the D p-brane, we obtain a D $(p+1)$-brane. Consider a 9-brane in a type IIB background. The 9-brane fills the spacetime and the endpoints of the open strings attached to it are free to move in all the directions. It is a type I theory, with only $N = 1$ supersymmetry. Now T-dualize one direction of the target space. We obtain an 8-brane in a type IIA background. If we proceed further, we obtain that a type IIB background can hold $p = 9, 7, 5, 3, 1, -1$ p-branes. A D (-1)-brane is a D-instanton, a localized spacetime point. For a type IIA background we obtain $p = 8, 6, 4, 2, 0$ p-branes.

BPS states with RR charges.

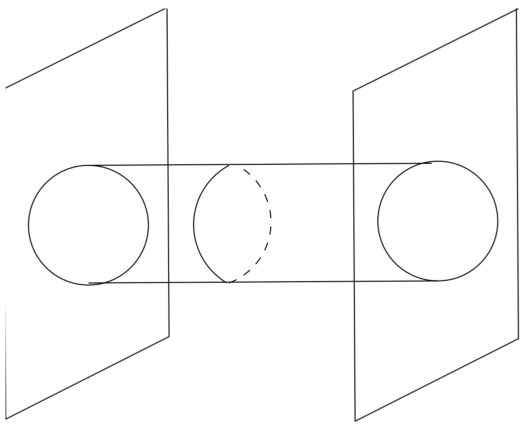


Figure 7: Two parallel D-branes with the one-loop vacuum fluctuation of an open string attached between them. By modular invariance, it also corresponds to a tree level interchange of a closed string.

To check if really the D-branes are the nonperturbative string solitons required by string duality, Polchinski computed explicitly the tension and RR charge of a D p-brane [57]. He first computed the one-loop amplitude of an open string attached to two parallel D p-branes. The resulting Casimir force between the D-branes was zero, supporting its BPS nature. By modular invariance, it can also be interpreted as the amplitude for the interchange of a closed string between the D-branes (see fig. 7). In the large separation limit, only the massless closed modes contribute. These are the NS-NS fields (graviton and dilaton) and the R-R $(p+1)$ form. On the space between the D-branes these fields follow the low energy type II action (type IIA for p even and type IIB for p odd). On the D p-branes, the coupling to the NS-NS and R-R fields is

$$S_p = T_p \int d^{p+1}\xi \; e^{-s} \, |\det G_{ab}|^{1/2} + \mu_p \int_{p-\text{brane}} A_{p+1} \,. \tag{15.538}$$

From (15.538) we see that the actual D-brane action includes a dilaton factor $\tau_p = T_p/g_s$, with g_s the coupling constant of the closed string theory. Comparing the field theory calculation with the contribution of the massless closed modes in the string theory computation, one can obtain the values of T_p and μ_p. The result is [57]

$$\mu_p^2 = 2T_p^2 = (4\pi^2\alpha')^{3-p} \,. \tag{15.539}$$

Observe that the R-R charge is really non-zero. In fact, if one checks (the generalization of) the Dirac's quantization condition for the charge μ_p and its dual charge $\mu_{(6-p)}$, one obtains that $\mu_p\mu_{(6-p)} = 2\pi$. They satisfy the minimal quantization condition. It means that the D-branes carry the minimal allowed RR charges.

15.4.27 Some final comments on nonperturbative String Theory.

D-instantons and S-duality.

The answers to the three questions at the beginning of the previous section can now be more concrete, since some nonperturbative objects in String Theory has been identified: the D-branes.

Consider a D p-brane wrapped around a non-trivial $(p+1)$-cycle. This configuration is topologically stable. Its action is $T_p V_{p+1}/g_s$, with V_{p+1} the volume of the non-trivial $(p+1)$ cycle. It contributes in amplitudes with factors $e^{-T_p V_{p+1}/g_s}$, a generalized instanton effect. Now we understand why the nonperturbative effects in String Theory are stronger than in field theory, it is related to the peculiar nature of the string solitons.

The D-branes also give the necessary ingredient for the $SL(2,\mathbf{Z})$ self-duality of the type IIB string theory. This theory allows D 1-branes, with a mass $\tau_1 \sim (2\pi\alpha' g_s)^{-1}$ in the string metric and non-zero RR charge. Also, one can see that on the D 1-brane there are the same fluctuations of a fundamental IIB string [58]. Then, it is the required object for the S-duality transformation of the type IIB string. In fact, at strong coupling the D 1-string becomes light and it is natural to formulate the type IIB theory in terms of weakly coupled D 1-branes.

There is another S-duality relation in String Theory. Observe that the type I theory and the $SO(32)$ heterotic theory have the same low energy limit. It could be that they correspond to the same theory but for different values of the string coupling constant. Again D-branes help to make this picture clearer. Consider a D 1-brane in a type I background with open strings attached to it, but also with open strings with one end point attached to a 9-brane. We call them $1-9$ strings. The 9-brane fills the spacetime, and the $1-9$ strings, having one Chan-Paton index, are vectors of $SO(32)$. One can see that the world-sheet theory of the D 1-brane is precisely that of the $SO(32)$ heterotic string [59]. Having a tension that scales as g_s^{-1}, one can argue that this D heterotic string sets the lightest scale in the theory when $g_s\hat{g}1$. The strong coupling behavior of the type I string can be modeled by the weak coupling behavior of the heterotic string.

An eleventh dimension.

Type IIA allows the existence of 0-branes that couple to the R-R one-form A_1. The 0-brane mass is $\tau_0 \sim (\alpha')^{-1/2}/g_s$ in the string metric. At strong coupling in the type IIA theory, $g_s \hat{g} 1$, this mass is the lightest scale of the theory. In fact, n 0-branes can form a BPS bound state with mass $n\tau_0$. This tower of states becoming a continuum of light states at strong coupling is characteristic of the appearance of an additional dimension. Type IIA theory at strong coupling feels an eleventh dimension of some size $2\pi R$, with the 0-branes playing the role of the Kaluza-Klein states [60].

If we compactify 11D supergravity [61] on a circle of radius R and compare its action with the 10D type IIA supergravity action, we obtain the relation

$$R \sim g_s^{2/3} \,. \tag{15.540}$$

This eleventh dimension is invisible in perturbation theory, where we perform an expansion near $g_s = 0$.

This has been a lightning review of some aspects of duality in String Theory. We hope it will serve to whet the appetite of the reader and encourage her/him to learn more about the subject and to eventually contribute to some of the outstanding open problems. More information can be found from the references [62].

15.5 Other dualities

15.5.1 Duality in Poincare group

(подробно обсужу это, пока не до этого.)

15.5.2 Duality and coset manifolds (???)

(Проен сказал, что это может быть связано, но пока это для меня разные вещи.)

Part V
Appendix

.1 Additional general information

.1.1 Additional motivation

В дуальности простая идея, но чтобы проделать преобразования, нужно на самом деле уметь работать с формулами, так что можно очень круто прокачать навык организации работы с информацией

(раскрою эту мысль, поэтому очень прикольно когда-то думать про то, как доделать какую-то формулу в теории этой. На самом деле это часто как вызов в сложных вычислениях мотивирует даже больше, чем другие аспекты.)

.1.2 Literature

Литература об основах

[3] Freedman, Daniel Z., Van Proeyen, Antoine Supergravity

Книга, где очень много теоретических основ, в том числе общий формализм дуальности обсуждается.

[2] JM Figueroa-O'Farrill Electromagnetic Duality for Children

Примерно 200 страниц очень крутой, вводной теории, Зи тоже ее рекомендует.

[5] Mary K. Gaillard, Bruno Zumino Duality rotations for interacting fields

Статья, где разбирается дуальность. Написана ужасно, математика соврешенно непонятная, физику вообще крайне сложно увидеть. Но хотя бы результаты получены известные многим, так что если не известна лучшая статья, можно и ее разбирать, но это будет неприятно и многое время не так уж полезно.

[4] Fré, Pietro Lectures on Special Kähler geometry and electric—magnetic duality rotations

Хорошие лекции для прояснения дуальности и не только. Тема 2х недель изучений. Больше там про Кэлерову геометрию, а про дуальность не знаю, насколько она полезная, но много хорошего обзора точно.

References from Gillard, Zumino

[1] C.W. Misner and J.A. Wheeler, Ann. of Phys. 2 (1957) 525

[2] S. Deser and C. Teitelboim, Phys. Rev. D13 (1976) 1592

[3] S. Ferrara, J. Scherk and B. Zumino, Nucl. Phys. B121 (1977) 393

[4] E. Cremmer, J. Scherk and S. Ferrara, Phys. Lett. 68B (1977) 234

[5] E. Cremmer and J. Scherk, Nucl. Phys. B127 (1977) 259

[6] E. Cremmer, J. Scherk and S. Ferrara, Phys. Lett. 74B (1978) 61

[7] E. Cremmer and B. Julia, Nucl. Phys. B159 (1979) 141

[8] S. Weinberg and E. Witten, Phys. Lett. 96B (1980) 59

[9] S. Coleman and E. Witten, Phys. Rev. Lett. 45 (1980) 100

[10] S. Coleman, J. Wess and B. Zumino, Phys. Rev. 177 (1969) 2239

[11] C. Callan, S. Coleman, J. Wess and B. Zumino, Phys. Rev. 177 (1969) 2247

[12] K. Cahill, Phys. Rev. D18 (1978) 2930

[13] E. Cremmer and J. Scherk, Phys. Lett. 74B (1978) 341

[14] B. Julia and J.F. Luciani, Phys. Lett. 90B (1980) 270

[15] J. Ellis, M.K. Gaillard, L. Maiani and B. Zumino, Unification of the fundamental particle interactions, ed. S. Ferrara, J. Ellis and P. Van Nieuwenhuizen (Plenum Press, New York, 1980) p. 69

[16] J. Ellis, M.K. Gaillard and B. Zumino, Phys. Lett. 94B (1980) 143

Литература о приложениях

[1] Andrianopoli, L. et. al. N = 2 supergravity and N = 2 super Yang-Mills theory on general scalar manifolds: Symplectic covariance gaugings and the momentum map

Применение в супергравитации дуальности, с нее начинается и что-то дальше с ней делается, не особо въехал, но если нужно будет - только ее и буду читать про это. Она же первая гуглится.

Другая литература об основах

Киселев Лекции по квантовой механике 2022

Там есть небольшой раздел про дуальность и про вывод уравнений электродинамики, так что тоже чуть почитать полезно, но совсем мало про нее.

.1.3 History of research in duality

.2 Мышление профессионала в дуальности

.2.1 Насколько же на самом деле дуальность полезна и работает?

(словами критический обзор)

.2.2 Об изучении дуальности

Что на самом деле стоит изучать в дуальности, а что нет?

(тут же рассуждения, что некоторые вещи не стоит изучать, например, Гиллард Зумино статью, потому что трата времени. не знаю, продумаю это лучше потом, все-таки стоит, просто тормозить много на этом не нужно, многие вещи просто такие, что можно более простыми способами доказать, а не через крышесносную математику..)

Не нужно недооценивать теории дуальности!

Это большие теории, в которых можно пару месяцев только сидеть. Не нужно думать, что за неделю работа сделается и понятно хотя бы многое будет.

(раскрою мысль, итак это очевидно из всей этой записи)

.3 Mathematical tricks for duality

.3.1 On variational methods in duality (!?)

Main variational properties

Let's discuss some properties for variations that we will use later.

1.

$$F\tilde{G} = \tilde{F}G.$$

This is because $F\tilde{G} \equiv \frac{1}{2}F^{\mu\nu}\varepsilon_{\mu\nu\rho\sigma}G^{\rho\sigma} = \frac{1}{2}G^{\rho\sigma}\varepsilon_{\rho\sigma\mu\nu}F^{\mu\nu} = \tilde{F}G.$

3. For any symmetric matrix B $(B = B^T)$ and arbitrary matrix Gs

$$\frac{\partial(GB\tilde{G})}{\partial\chi} = 2\frac{\partial G}{\partial\chi}B\tilde{G}, \qquad \Leftrightarrow \qquad \frac{\partial G}{\partial\chi}B\tilde{G} = \frac{1}{2}\frac{\partial(GB\tilde{G})}{\partial\chi}$$

Proof:

$$GB\frac{\partial\tilde{G}}{\partial\chi} \equiv G^{\mu\nu}B\frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}\frac{\partial G^{\rho\sigma}}{\partial\chi} = \frac{\partial G^{\mu\nu}}{\partial\chi}B\frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}G^{\rho\sigma} = \frac{\partial G}{\partial\chi}B\tilde{G}$$

so

$$\frac{1}{2}\frac{\partial(GB\tilde{G})}{\partial\chi} = \frac{1}{2}\left(\frac{\partial G}{\partial\chi}B\tilde{G} + GB\frac{\partial\tilde{G}}{\partial\chi}\right) = \frac{\partial G}{\partial\chi}B\tilde{G}$$

4.

$$\partial(\delta L) = \partial(\delta)L + \delta\partial L.$$

Example (??????????????) (it didn't work, so I just believe in this, I'll try later again and ask about it!)

5. For (?? which matrices???)

$$\tilde{F}BG = \tilde{G}B^T F.$$

Proof:

$$\tilde{F}BG = \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}F^{a\rho\sigma}B^{ab}G^{b\mu\nu} = \frac{1}{2}\varepsilon_{\rho\sigma\mu\nu}F^{a\mu\nu}B^{ab}G^{b\rho\sigma} = \frac{1}{2}\varepsilon_{\rho\sigma\mu\nu}G^{b\rho\sigma}B^{T^{ba}}F^{a\mu\nu} = \tilde{G}B^T F.$$

6. (!??! for potentials lemma for obtaining the current)

.3.2 Hodge duality of forms

(??? позже обсуждение, как они конкретно нужны в теорполе???)

By Proyen's book

Since p - and q-forms have the same number of components when $p + q = D$, it is possible to define a $1:1$ map between them. This map is the Hodge duality map from $\Lambda^p(M) \to \Lambda^q(M)$, and it is quite useful in the physics of supergravity. The map is denoted by $\Omega^{(q)} = *\omega^{(p)}$

Since the map is linear we can define it on a basis of p-forms and then extend to a general form. It is convenient to use the local frame basis initially and define

$$*e^{a_1} \wedge \cdots \wedge e^{a_p} = \frac{1}{q!}e^{b_1} \wedge \cdots \wedge e^{b_q}\varepsilon_{b_1\cdots b_q}{}^{a_1\cdots a_p}.$$

A general p-form can be expressed in this basis, and we can proceed to define its dual via

$$\begin{aligned}\Omega^{(q)} = {}^*\omega^{(p)} &= *\left(\frac{1}{p!}\omega_{a_1\cdots a_p}e^{a_1} \wedge \cdots \wedge e^{a_p}\right) \\ &= \frac{1}{p!}\omega_{a_1\cdots a_p}{}^*e^{a_1} \wedge \cdots \wedge e^{a_p}\end{aligned}$$

Exercise 7.15 Show that the frame components of $\Omega^{(q)}$ are given by

$$\Omega_{b_1\cdots b_q} = ({}^*\omega)_{b_1\cdots b_q} = \frac{1}{p!}\varepsilon_{b_1\cdots b_q}{}^{a_1\cdots a_p}\omega_{a_1\cdots a_p}$$

These formulas are far less complicated than they look since there is only one independent term in each sum. For example, for $D = 4$ the dual of a 3-form is a 1-form. For basis elements we have $*e^1 \wedge e^2 \wedge e^3 = e^0$ and ${}^*e^0 \wedge e^1 \wedge e^2 = e^3$. For components, $({}^*\omega)_0 = \omega_{123}$ and $({}^*\omega)_3 = \omega_{012}$.

The duality has an important involutive property, which can be inferred from the following sequence of operations on basis elements:

$$\begin{aligned} *{}^{*}e^{a_1}\wedge\cdots\wedge e^{a_p}) &= \frac{1}{q!}e^{b_1}\wedge\cdots\wedge e^{b_q}\varepsilon_{b_1\cdots b_q}{}^{a_1\cdots a_p} \\ &= \frac{1}{p!q!}e^{c_1}\wedge\cdots\wedge e^{c_p}\varepsilon_{c_1\cdots c_p}{}^{b_1\cdots b_q}\varepsilon_{b_1\cdots b_q}{}^{a_1\cdots a_p} \\ &= -(-)^{pq}e^{c_1}\wedge\cdots\wedge e^{c_p}\delta^{a_1\cdots a_p}_{c_1\cdots c_p} \\ &= -(-)^{pq}e^{a_1}\wedge\cdots\wedge e^{a_p} \end{aligned}$$

This leads to the general relation ${}^{*}\left({}^{*}\omega^{(p)}\right) = -(-)^{pq}\omega^{(p)}$. This is the correct relation for a Lorentzian signature manifold. For Euclidean signature the involution property is $*\left({}^{*}\omega^{(p)}\right) = (-)^{pq}\omega^{(p)}$.

For even dimension $D = 2m$, it is possible to impose the constraint of self-duality (or anti-self-duality) on forms of degree m, i.e. $\Omega^{(m)} = \pm{}^{*}\Omega^{(m)}$. In a given dimension this condition is consistent only if duality is a strict involution, i.e. $-(-)^{m^2} = -(-)^m = +1$ for Lorentzian signature and $(-)^m = +1$ for Euclidean signature. Thus it is possible to have self-dual Yang-Mills instantons in four Euclidean dimensions. A self-dual $F^{(5)}$ is possible in $D = 10$ Lorentzian signature, and it indeed appears in Type IIB supergravity.

The duality relations defined above in a frame basis are easily transformed to a coordinate basis using the relations $e^a = e^a_\mu(x)\mathrm{d}x^\mu$ and $\mathrm{d}x^\mu = e^\mu_a(x)e^a$. For coordinate basis elements the duality map is

$$*\left(\,\mathrm{d}x^{\mu_1}\wedge\cdots\wedge\mathrm{d}x^{\mu_p}\right) = \frac{1}{q!}eg^{\mu_1\rho_1}\cdots g^{\mu_p\rho_p}\,\mathrm{d}x^{\nu_1}\wedge\cdots\wedge\mathrm{d}x^{\nu_p}\varepsilon_{\nu_1\cdots\nu_q\rho_1\cdots\rho_p}.$$

For antisymmetric tensor components, we have

$$\left({}^{*}\omega\right)_{\mu_1\cdots\mu_q} = \frac{1}{p!}e\varepsilon_{\mu_1\cdots\mu_q\rho_1\cdots\rho_p}g^{\nu_1\rho_1}\ldots g^{v_p\rho_p}\omega_{\nu_1\cdots v_p}$$

Following the discussion in Sec. 7.5, we may take as a volume form on M the wedge product ${}^{*}\omega^{(p)}\wedge\omega^{(p)}$ of any p-form and its Hodge dual. The integral of this volume form is simply the standard invariant norm of the tensor components of $\omega^{(p)}$, i.e.

$$\int{}^{*}\omega^{(p)}\wedge\omega^{(p)} = \frac{1}{p!}\int\mathrm{d}^Dx\sqrt{-g}\omega^{\mu_1\cdots\mu_p}\omega_{\mu_1\cdots\mu_p}$$

Exercise 7.16 Prove (7.56). Use the definitions above and those in Sec. 7.5 and the fact that

$$e^{a_1}\wedge\cdots\wedge e^{a_q}\wedge e^{b_1}\wedge\cdots\wedge e^{b_p} = -\varepsilon^{a_1\cdots a_q b_1\cdots b_p}\,\mathrm{d}V$$

where $\mathrm{d}V$ is the canonical volume element of (7.48).

Exercise 7.17 Show that the volume form $\mathrm{d}V$ can also be written as $*1$.

Exercise 7.18 Compare these definitions with Sec. 4.2.1, to obtain

$$\tilde{F}_{\mu\nu} = -\mathrm{i}\left({}^{*}F\right)_{\mu\nu}.$$

Show that the factor i ensures that the tilde operation squares to the identity. Self-duality is then possible for complex 2-forms.

Exercise 7.19 For applications to gauge field theories it is useful to record the relation between the components of the field strength 2-form and its dual:

$${}^{*}F_{\mu\nu} \equiv \frac{1}{2}\sqrt{-g}\varepsilon_{\mu\nu\rho\sigma}F^{\rho\sigma}, \qquad {}^{*}F^{\mu\nu} = \frac{1}{2\sqrt{-g}}\varepsilon^{\mu\nu\rho\sigma}F_{\rho\sigma}$$

Verify the second relation. Since both $F_{\mu\nu}$ and ${}^*F_{\mu\nu}$ are tensors, their indices are raised by $g^{\mu\nu}$.

.3.3 Properties of the non-compact real symplectic group

Theory

We recall briefly some basic properties of the non-compact real symplectic group $\mathrm{Sp}(2n,\mathrm{R})$. It can be defined as the group of real $2n \times 2n$ matrices S which satisfy

$$S^{\mathrm{T}}\Omega S = \Omega$$

where Ω is a $2n \times 2n$ non-singular antisymmetric matrix which can be taken to be

$$\Omega = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Writing

$$S = 1 + X$$

where X is an infinitesimal matrix, (A.1) gives

$$X^{\mathrm{T}}\Omega = -\Omega X$$

If we write

$$X = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

where A, B, C and D are $n \times n$ real matrices, (A.4) is equivalent to

$$B^{\mathrm{T}} = B, \quad C^{\mathrm{T}} = C, \quad A^{\mathrm{T}} = -D.$$

It is often convenient to rewrite the infinitesimal transformation of $\mathrm{Sp}(2n,\mathrm{R})$,

$$\delta \begin{pmatrix} F \\ G \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} F \\ G \end{pmatrix}$$

in the complex basis $F \pm iG$. One obtains

$$\delta \begin{pmatrix} F + iG \\ F - IG \end{pmatrix} = \begin{pmatrix} T & V^* \\ V & T^* \end{pmatrix} \begin{pmatrix} F + iG \\ F - iG \end{pmatrix}$$

where

$$2T = A + D + i(C - B),$$
$$2V = A - D - i(C + B),$$

so that T is antihermitian and V is symmetric:

$$T^{\dagger} = -T, \quad V^{\mathrm{T}} = V$$

In the complex basis a finite transformation of $\mathrm{Sp}(2n,\mathbf{R})$ takes the form

$$S = \begin{pmatrix} \phi_0 & \phi_1^* \\ \phi_1 & \phi_0^* \end{pmatrix}$$

where the $n \times n$ complex matrices ϕ_0 and ϕ_1 satisfy the constraint

$$\phi_0^\dagger \phi_0 - \phi_1^\dagger \phi_1 = 1$$

The inverse matrix is

$$S^{-1} = \begin{pmatrix} \phi_0^\dagger & -\phi_1^\dagger \\ -\phi_1^{\mathrm{T}} & \phi_0^{\mathrm{T}} \end{pmatrix}$$

The maximal compact subgroup $\mathrm{U}(n)$ of $\mathrm{Sp}(2n, \mathrm{R})$ is easily recognized in (A.8) for $V = 0$ and in (A.11), (A.12) for $\phi_1 = 0$.

A Bibliography

[1] L. Andrianopoli, M. Bertolini, A. Ceresole, R. D'Auria, S. Ferrara, P. Fré, and T. Magri. N = 2 supergravity and n = 2 super yang-mills theory on general scalar manifolds: Symplectic covariance gaugings and the momentum map. *Journal of Geometry and Physics*, 23(2):111–189, September 1997.

[2] JM Figueroa-O'Farrill. Electromagnetic duality for children, 1998.

[3] Daniel Z. Freedman and Antoine Van Proeyen. *Supergravity*. Cambridge Univ. Press, Cambridge, UK, 5 2012.

[4] Pietro Fré. Lectures on special kähler geometry and electric—magnetic duality rotations. *Nuclear Physics B - Proceedings Supplements*, 45(2–3):59–114, February 1996.

[5] Mary K. Gaillard and Bruno Zumino. Duality rotations for interacting fields. *Nuclear Physics B*, 193(1):221–244, 1981.

[6] Mary K. Gaillard and Bruno Zumino. Nonlinear electromagnetic self-duality and legendre transformations, 1997.

From 5d/4d U-dualities and $\mathcal{N}=8$ black holes Ceresole, Ferrara and Gnecchi

B Bibliography

[1] E. Cremmer and B. Julia, "The SO(8) Supergravity," Nucl. Phys. B **159**, 141 (1979).

[2] E. Cremmer, "Supergravities In 5 Dimensions," Invited paper at the Nuffield Gravity Workshop, Cambridge, Eng., Jun 22 - Jul 12, 1980. Published in Cambridge Workshop 1980:267.

[3] N. Arkani-Hamed, F. Cachazo and J. Kaplan, "What is the Simplest Quantum Field Theory?," arXiv:0808.1446 [hep-th].

[4] Z. Bern, J. J. M. Carrasco and H. Johansson, "Progress on Ultraviolet Finiteness of Supergravity," Invited talk at International School of Subnuclear Physics 2008: 46th Course: Predicted and Totally Unexpected in the Energy Frontier Opened by LHC, Erice, Italy, 29 Aug - 7 Sep 2008. arXiv:0902.3765 [hep-th].

[5] G. W. Gibbons and P. K. Townsend, "Vacuum interpolation in supergravity via super p-branes," Phys. Rev. Lett. **71**, 3754 (1993) [arXiv:hep-th/9307049].

[6] M. J. Duff, R. R. Khuri and J. X. Lu, "String solitons," Phys. Rept. **259**, 213 (1995) [arXiv:hep-th/9412184].

[7] S. Ferrara, R. Kallosh and A. Strominger, "$\mathcal{N}=2$ extremal black holes", Phys. Rev. D **52**, R5412 (1995), [arXiv:hep-th/9508072]; A. Strominger, "Macroscopic entropy of $\mathcal{N}=2$ extremal black holes", Phys. Lett. **B383**, 39 (1996), [arXiv:hep-th/9602111]; S. Ferrara and R. Kallosh, "Supersymmetry and attractors", Phys. Rev. D **54**, 1514 (1996), [arXiv:hep-th/9602136]; S. Ferrara, G. W. Gibbons and R. Kallosh, "Black holes and critical points in moduli space", Nucl. Phys. B **500**, 75 (1997), [arXiv:hep-th/9702103].

[8] S. Ferrara and R. Kallosh, "Universality of Supersymmetric Attractors," Phys. Rev. D **54**, 1525 (1996) [arXiv:hep-th/9603090];

[9] B. Bates and F. Denef, "Exact solutions for supersymmetric stationary black hole composites," arXiv:hep-th/0304094.

[10] F. Denef and G. W. Moore, "Split states, entropy enigmas, holes and halos," arXiv:hep-th/0702146.

[11] H. Ooguri, A. Strominger and C. Vafa, "Black hole attractors and the topological string," Phys. Rev. D **70**, 106007 (2004) [arXiv:hep-th/0405146].

[12] see for instance A. Sen, "Black Hole Entropy Function, Attractors and Precision Counting of Microstates," Gen. Rel. Grav. **40**, 2249 (2008) [arXiv:0708.1270 [hep-th]].

[13] M. Cvetic and D. Youm, "BPS saturated dyonic black holes of N = 8 supergravity vacua," Nucl. Phys. Proc. Suppl. **46**, 56 (1996) [arXiv:hep-th/9510098].

[14] R. R. Khuri and T. Ortin, "Supersymmetric Black Holes in N=8 Supergravity," Nucl. Phys. B **467**, 355 (1996) [arXiv:hep-th/9512177].

[15] G. Arcioni, A. Ceresole, F. Cordaro, R. D'Auria, P. Fre, L. Gualtieri and M. Trigiante, "N = 8 BPS black holes with 1/2 or 1/4 supersymmetry and solvable Lie algebra decompositions," Nucl. Phys. B **542**, 273 (1999) [arXiv:hep-th/9807136].

[16] L. Andrianopoli, R. D'Auria and S. Ferrara, "U-duality and central charges in various dimensions revisited," Int. J. Mod. Phys. A **13**, 431 (1998) [arXiv:hep-th/9612105].

[17] S. Ferrara and R. Kallosh, "On $\mathcal{N} = 8$ attractors", Phys. Rev. D **73** (2006) 125005 [arXiv:hep-th/0603247].

[18] D. Shih, A. Strominger and X. Yin, "Counting dyons in N = 8 string theory," JHEP **0606**, 037 (2006) [arXiv:hep-th/0506151].

[19] S. Ferrara, E. G. Gimon and R. Kallosh, "Magic supergravities, $\mathcal{N} = 8$ and black hole composites", Phys. Rev. D **74** (2006) 125018 [arXiv:hep-th/0606211].

[20] A. Sen, "N=8 Dyon Partition Function and Walls of Marginal Stability," JHEP **0807**, 118 (2008) [arXiv:0803.1014 [hep-th]].

[21] L. Andrianopoli, R. D'Auria, S. Ferrara and M. Trigiante, "Extremal black holes in supergravity," Lect. Notes Phys. **737**, 661 (2008) [arXiv:hep-th/0611345].

[22] A. Sen, "Arithmetic of N=8 Black Holes," arXiv:0908.0039 [hep-th].

[23] S. Ferrara and J. M. Maldacena, "Branes, central charges and U-duality invariant BPS conditions", Class. Quant. Grav. **15**, 749 (1998) [arXiv:hep-th/9706097].

[24] S. Ferrara and M. Günaydin, "Orbits of exceptional groups, duality and BPS states in string theory", Int. J. Mod. Phys. A **13** (1998) 2075 [arXiv:hep-th/9708025].

[25] E. Cartan, "Oeuvres Complètes", (Editions du Centre National de la Recherche Scientifique, Paris, 1984).

[26] R. Kallosh and B. Kol, $E_{7(7)}$ "Symmetric Area of the Black Hole Horizon", Phys. Rev. D **53** (1996) R5344 [arXiv:hep-th/9602014].

[27] C. M. Hull and P. K. Townsend, "Unity of superstring dualities", Nucl. Phys. B **438**, 109 (1995) [arXiv:hep-th/9410167].

[28] H. Lu, C. N. Pope and K. S. Stelle, "Multiplet structures of BPS solitons," Class. Quant. Grav. **15**, 537 (1998) [arXiv:hep-th/19708109].

[29] A. Ceresole, S. Ferrara and A. Marrani, "$4d/5d$ Correspondence for the Black Hole Potential and its Critical Points" , Class. Quant. Grav. **24**, 5651 (2007), arXiv:0707.0964 [hep-th].

[30] L. Andrianopoli, R. D'Auria, S. Ferrara and M. A. Lledó, "Gauging of flat groups in four dimensional supergravity", JHEP **0207** (2002) 010 [arXiv:hep-th/0203206].

[31] B. de Wit and H. Nicolai, "$\mathcal{N} = 8$ Supergravity", Nucl. Phys. B **208** (1982) 323.

[32] E. Sezgin and P. van Nieuwenhuizen, "Renormalizability Properties Of Spontaneously Broken $\mathcal{N} = 8$ Supergravity", Nucl. Phys. B **195** (1982) 325.

[33] A. Ceresole, S. Ferrara, A. Gnecchi and A. Marrani, "More on N=8 Attractors," arXiv:0904.4506 [hep-th], to appear on Phys. Rev. D.

[34] A. Ceresole, R. D'Auria and S. Ferrara, "The Symplectic Structure of $N = 2$ Supergravity and Its Central Extension", Nucl. Phys. Proc. Suppl. **46** (1996), [arXiv:hep-th/9509160].

[35] S. Ferrara and M. Günaydin, "Orbits and attractors for $\mathcal{N} = 2$ Maxwell-Einstein supergravity theories in five dimensions", Nucl. Phys. B **759** (2006) 1 [arXiv:hep-th/0606108].

[36] R. D'Auria, S. Ferrara and M. A. Lledó, "On central charges and Hamiltonians for 0-brane dynamics," Phys. Rev. D **60**, 084007 (1999) [arXiv:hep-th/9903089].

[37] S. Bellucci, S. Ferrara, M. Günaydin and A. Marrani, "Charge orbits of symmetric special geometries and attractors", Int. J. Mod. Phys. A **21** (2006) 5043 [arXiv:hep-th/0606209].

[38] L. Andrianopoli, R. D'Auria and S. Ferrara, "Five dimensional U-duality, black-hole entropy and topological invariants", Phys. Lett. B **411**, 39 (1997) [arXiv:hep-th/9705024].

[39] M. Gunaydin, G. Sierra and P. K. Townsend, "The Geometry Of N=2 Maxwell-Einstein Supergravity And Jordan Algebras," Nucl. Phys. B **242**, 244 (1984).

[40] M. K. Gaillard and B. Zumino, "Duality Rotations For Interacting Fields," Nucl. Phys. B **193** (1981) 221.

[41] P. Aschieri, S. Ferrara and B. Zumino, "Duality Rotations in Nonlinear Electrodynamics and in Extended Supergravity," Riv. Nuovo Cim. **31**, 625 (2009) [Riv. Nuovo Cim. **031**, 625 (2008)] [arXiv:0807.4039 [hep-th]].

[42] A. Ceresole, S. Ferrara and A. Gnecchi, in preparation.

From Special geometry and symplectic transformations by de Wit, Van Proeyen

C Bibliography

[1] B. de Wit, P.G. Lauwers, R. Philippe, Su S.-Q. and A. Van Proeyen, Phys. Lett. **134B** (1984) 37;
B. de Wit and A. Van Proeyen, Nucl. Phys. **B245** (1984) 89.

[2] A. Strominger, Commun. Math. Phys. **133** (1990) 163.

[3] N. Seiberg, Nucl. Phys. **B303** (1988) 286.

[4] S. Cecotti, S. Ferrara and L. Girardello, Int. J. Mod. Phys. **A4** (1989) 2457.

[5] S. Ferrara and A. Strominger, in *Strings '89*, eds. R. Arnowitt, R. Bryan, M.J. Duff, D.V. Nanopoulos and C.N. Pope (World Scientific, 1989), p. 245.

[6] L.J. Dixon, V.S. Kaplunovsky and J. Louis, Nucl. Phys. **B329** (1990) 27.

[7] P. Candelas and X. C. de la Ossa, Nucl. Phys. **B355** (1991) 455,
P. Candelas, X. C. de la Ossa, P. Green and L. Parkes, Phys. Lett. **258B** (1991) 118; Nucl. Phys. **B359** (1991) 21.

[8] P. Frè, contribution to these proceedings.

[9] G. Sierra and P.K. Townsend, in *Supersymmetry and Supergravity 1983*, ed. B. Milewski (World Scientific, Singapore, 1983);
S. J. Gates, Nucl. Phys. B238 (1984) 349.

[10] B. de Wit, in *Supergravity '81*, eds. S. Ferrara and J.G. Taylor (Cambridge Univ. Press, 1982);
A. Van Proeyen, in *Supersymmetry and Supergravity 1983*, ed. B. Milewski (World Scientific, Singapore, 1983).

[11] B. de Wit, P. Lauwers and A. Van Proeyen, Nucl. Phys. **B255** (1985) 569.

[12] L. Castellani, R. D' Auria and S. Ferrara, Phys. Lett. **B241** (1990) 57; Cl.Q. Grav. **7** (1990) 1767,
R. D'Auria, S. Ferrara and P. Fré, Nucl. Phys. **B359** (1991) 705.

[13] E. Cremmer, C. Kounnas, A. Van Proeyen, J.P. Derendinger, S. Ferrara, B. de Wit and L. Girardello, Nucl. Phys. **B250** (1985) 385.

[14] E. Cremmer and A. Van Proeyen, Class. Quantum Grav. **2** (1985) 445.

[15] C.M. Hull and A. Van Proeyen, Phys. Lett. **B351** (1995) 188, hep-th/9503022.

[16] A. Sen, contribution to these proceedings.

[17] B. de Wit and A. Van Proeyen, Phys. Lett. **B293** (1992) 94.

[18] A. Das, Phys. Rev. **D15**, 2805 (1977); E. Cremmer, J. Scherk and S. Ferrara, Phys. Lett. **68B**, 234 (1977); E. Cremmer and J. Scherk, Nucl. Phys. **B127**, 259 (1977).

[19] A. Ceresole, R. D'Auria, S. Ferrara and A. Van Proeyen, Nucl. Phys. **B444** (1995) 92, hep-th/9502072.

[20] E. Cremmer, J. Scherk and S. Ferrara Phys. Lett. **74B**, 61 (1978).

[21] S. Ferrara and A. Van Proeyen, Class. Quantum Grav. **6** (1989) 124.

[22] A. Sen, Nucl. Phys. **B388** (1992) 457 and Phys. Lett. **B303** (1993) 22;
A. Sen and J. H. Schwarz, Nucl. Phys. **B411** (1994) 35; Phys. Lett. **B312** (1993) 105.

[23] P. Fré and P. Soriani, Nucl. Phys. **B371** (1992) 659;
S. Ferrara, P. Fré and P. Soriani, Class. Quantum Grav. **9** (1992) 1649.

[24] B. de Wit, V Kaplunovsky, J. Louis and D. Lüst, Nucl. Phys. **B451** (1995) 53.

[25] A. Ceresole, R. D'Auria and S. Ferrara, Phys. Lett. **B339** (1994) 71, hep-th/9408036.

[26] B. de Wit and A. Van Proeyen, *Isometries of special manifolds*, to be published in the proceedings of the Meeting on Quaternionic Structures in Mathematics and Physics, Trieste, September 1994, preprint THU-95/13, KUL-TF-95/13, hep-th/9505097.

[27] A. Ceresole, R. D'Auria, S. Ferrara, W. Lerche and J. Louis, Int. J. Mod. Phys. **A8** (1993) 79, hep-th/9204035.

[28] M. Günaydin, G. Sierra and P.K. Townsend, Phys. Lett. **B133** (1983) 72; Nucl. Phys. **B242** (1984) 244, **B253** (1985) 573.

[29] B. de Wit, F. Vanderseypen and A. Van Proeyen, Nucl. Phys. **B400** (1993) 463.

[30] B. de Wit and A. Van Proeyen, Commun. Math. Phys. **149** (1992) 307.

[31] N. Seiberg and E. Witten, Nucl. Phys. **B426** (1994) 19; **B431** (1994) 484.

[32] P. Frè and P. Soriani, "The N=2 Wonderland: from Calabi–Yau manifolds to topological field theories", World Scientific, Singapore, 1995.

From Duality Rotations in Nonlinear Electrodynamics and in Extended Supergravity by Aschieri, Ferrara, Zumino

D Bibliography

[1] M. Born, L. Infeld, *Foundations of the New Field Theory*, Proc. Roy. Soc. (London) A144 (1934) 425

[2] E. Schrödinger, *Contributions to Born's New Theory of the Electromagnetic Field*, Proc. Roy. Soc. (London) A150 (1935) 465

[3] S. Ferrara, J. Scherk, B. Zumino, *Algebraic Properties of Extended Supergravity Theories*, Nucl. Phys. B121 (1977) 393

[4] E. Cremmer, J. Scherk and S. Ferrara, *SU(4) Invariant Supergravity Theory,* Phys. Lett. B **74**, 61 (1978).

[5] E. Cremmer, B. Julia, *The SO(8) Supergravity*, Nucl. Phys. B159 (1979) 141, E. Cremmer and B. Julia, *The N=8 Supergravity Theory. 1. The Lagrangian,* Phys. Lett. B **80** (1978) 48.

[6] M. K. Gaillard, B. Zumino, *Duality Rotations for Interacting Fields*, Nucl. Phys. B193 (1981) 221

[7] B. Zumino, *Quantum Structure of Space and Time*, Eds. M. J. Duff and C. J. Isham (Cambridge University Press) (1982) 363

[8] L. Castellani, R. D'Auria and P. Fré, *Supergravity and Superstrings: a Geometric Perspective* World Scientific 1991

[9] A. Ceresole, R. D'Auria, S. Ferrara and A. Van Proeyen, *Duality transformations in supersymmetric Yang-Mills theories coupled to supergravity,* Nucl. Phys. B **444** (1995) 92 [arXiv:hep-th/9502072].

[10] L. Andrianopoli, R. D'Auria, S. Ferrara, P. Fré and M. Trigiante, *R-R Scalars, U–Duality and Solvable Lie Algebras*, hep-th/9611014, Nucl.Phys.B496:617-629,1997

[11] L. Andrianopoli, R. D'Auria and S. Ferrara, *U-duality and central charges in various dimensions revisited,* Int. J. Mod. Phys. A **13** (1998) 431 [arXiv:hep-th/9612105].

[12] G. W. Gibbons, D. A. Rasheed *Electric-Magnetic Duality Rotations in Non-Linear Electrodynamics*, Nucl. Phys. B454 (1995) 185; hep-th/9506035

[13] G. W. Gibbons, D. A. Rasheed, $SL(2,R)$ *Invariance of Nonlinear Electrodynamics Coupled to an Axion and Dilaton*, Phys. Lett. B365 (1996) 46; hep-th/9509141

[14] M. K. Gaillard, B. Zumino, *Self-Duality in Nonlinear Electromagnetism*, Supersymmetry and Quantum Field Theory, Kharkov 1997, J. Wess and V. P. Akulov, eds., Springer-Verlag 1998; hep-th/9705226

[15] M. K. Gaillard, B. Zumino, *Nonlinear Electromagnetic Self-Duality and Legendre Transformations*, Duality and Supersymmetric Theories, D. I. Olive and P. C. West, eds., Cambridge University Press 1999; hep-th/9712103

[16] M. Araki, Y. Tanii, *Duality Symmetries in Nonlinear Gauge Theories*, Int. J. Mod. Phys. A14 (1999) 1139; hep-th/9808029

[17] D. Brace, B. Morariu, B. Zumino, *Duality Invariant Born-Infeld Theory*, Yuri Golfand memorial volume; hep-th/9905218

[18] P. Aschieri, D. Brace, B. Morariu and B. Zumino, *Nonlinear self-duality in even dimensions,* Nucl. Phys. B **574**, 551 (2000) [arXiv:hep-th/9909021].

[19] P. Aschieri, D. Brace, B. Morariu and B. Zumino, *Proof of a symmetrized trace conjecture for the Abelian Born-Infeld Lagrangian,* Nucl. Phys. B **588**, 521 (2000) [arXiv:hep-th/0003228].

[20] E. A. Ivanov and B. M. Zupnik, *New representation for Lagrangians of self-dual nonlinear electrodynamics,* Talk given at International Workshop on Supersymemtry and Quantum Symmetry (SQS '01): 16th Max Born Symposium, Karpacz, Poland, 21-25 Sep 2001. Published in Karpacz 2001, Supersymmetries and quantum symmetries, 235-250, arXiv:hep-th/0202203.
E. A. Ivanov and B. M. Zupnik, *New approach to nonlinear electrodynamics: Dualities as symmetries of interaction,* Phys. Atom. Nucl. **67** (2004) 2188 [Yad. Fiz. **67** (2004) 2212] [arXiv:hep-th/0303192].

[21] I. Antoniadis, H. Partouche, T. R. Taylor, *Spontaneous Breaking of N=2 Global Supersymmetry*, Phys. Lett. B372 (1996); 83 hep-th/9512006

[22] M. Rocek, A. A. Tseytlin, *Partial breaking of global D=4 supersymmetry, constrained superfields, and 3-brane actions*, Phys. Rev. D59 (1999) 106001; hep-th/9811232

[23] S. M. Kuzenko and S. Theisen, *Supersymmetric duality rotations,* JHEP **0003** (2000) 034 [arXiv:hep-th/0001068].

[24] S. M. Kuzenko and S. Theisen, *Nonlinear self-duality and supersymmetry,* Fortsch. Phys. **49** (2001) 273 [arXiv:hep-th/0007231].

[25] S. M. Kuzenko and S. A. McCarthy, *Nonlinear self-duality and supergravity,* JHEP **0302** (2003) 038 [arXiv:hep-th/0212039]. S. M. Kuzenko and S. A. McCarthy, *On the component structure of N = 1 supersymmetric nonlinear electrodynamics,* JHEP **0505** (2005) 012 [arXiv:hep-th/0501172].

[26] Y. Tanii, *N=8 Supergravity in Six-Dimensions*, Phys. Lett. 145B (1984) 197

[27] S. Cecotti, S. Ferrara, L. Girardello, *Hidden Noncompact Symmetries in String Theory*, Nucl. Phys. B308 (1988) 436

[28] B. L. Julia, *Dualities in the classical supergravity limits*, in Strings Branes and Dualities, Cargese (1997), 121 hep-th/9805083

[29] C. M. Hull and P. K. Townsend, *Unity of superstring dualities,* Nucl. Phys. B **438**, 109 (1995) [arXiv:hep-th/9410167].

[30] P. Fre, *Lectures on Special Kahler Geometry and Electric–Magnetic Duality Rotations,* Nucl. Phys. Proc. Suppl. **45BC** (1996) 59 [arXiv:hep-th/9512043].

[31] E. Cremmer, B. L. Julia, H. Lü, C. N. Pope, *Dualisation of Dualities, I & II*, Nucl.Phys. B523 (1998) 73, B535 (1998) 242; hep-th/9710119, hep-th/9806106

[32] J. Schwinger, *A Magnetic Model of Matter,* Science 22 **165** (1969) 757.

[33] B. de Wit and H. Nicolai, *N=8 Supergravity,* Nucl. Phys. B **208** (1982) 323.

[34] L. Andrianopoli, R. D'Auria, S. Ferrara and M. A. Lledo, *Gauging of flat groups in four dimensional supergravity,* JHEP **0207** (2002) 010 [arXiv:hep-th/0203206].

[35] J. Scherk and J. H. Schwarz, *How To Get Masses From Extra Dimensions,* Nucl. Phys. B **153**, 61 (1979).

[36] E. Cremmer, J. Scherk and J. H. Schwarz, *Spontaneously Broken N=8 Supergravity,* Phys. Lett. B **84** (1979) 83.

[37] S. Ferrara, R. Kallosh and A. Strominger, *N=2 extremal black holes,* Phys. Rev. D **52** (1995) 5412 [arXiv:hep-th/9508072].

[38] S. Ferrara and R. Kallosh, *Supersymmetry and Attractors,* Phys. Rev. D **54** (1996) 1514 [arXiv:hep-th/9602136].
S. Ferrara and R. Kallosh, *Universality of Supersymmetric Attractors,* Phys. Rev. D **54** (1996) 1525 [arXiv:hep-th/9603090].

[39] A. Strominger, *Macroscopic Entropy of* $N = 2$ *Extremal Black Holes,* Phys. Lett. B **383** (1996) 39 [arXiv:hep-th/9602111].

[40] S. Bellucci, S. Ferrara, R. Kallosh and A. Marrani, *Extremal Black Hole and Flux Vacua Attractors,* arXiv:0711.4547 [hep-th].

[41] L. Andrianopoli, R. D'Auria, S. Ferrara and M. Trigiante, *Extremal black holes in supergravity,* Lect. Notes Phys. **737**, 661 (2008) [arXiv:hep-th/0611345].

[42] S. Ferrara, K. Hayakawa and A. Marrani, *Erice Lectures on Black Holes and Attractors,* arXiv:0805.2498 [hep-th].

[43] S. Ferrara, G. W. Gibbons and R. Kallosh, *Black holes and critical points in moduli space,* Nucl. Phys. B **500** (1997) 75 [arXiv:hep-th/9702103].

[44] R. Kallosh and B. Kol, *E(7) Symmetric Area of the Black Hole Horizon,* Phys. Rev. D **53** (1996) 5344 [arXiv:hep-th/9602014].

[45] S. Ferrara and J. M. Maldacena, *Branes, central charges and U-duality invariant BPS conditions,* Class. Quant. Grav. **15** (1998) 749 [arXiv:hep-th/9706097].
S. Ferrara and M. Gunaydin, *Orbits of exceptional groups, duality and BPS states in string theory,* Int. J. Mod. Phys. A **13** (1998) 2075 [arXiv:hep-th/9708025].

[46] S. Deser, R. Puzalowski, *Supersymmetric Nonpolynomial Vector Multiplets and Causal Propagation*, J. Phys. A13 (1980) 2501

[47] S. Cecotti, S. Ferrara, *Supersymmetric Born-Infeld Lagrangians*, Phys. Lett. 187B (1987) 335

[48] J, Bagger, A, Galperin, *A New Goldstone Multiplet for a Partially Broken Supersymmetry*, Phys. Rev. D55 (1997) 1091; hep-th/9608177

[49] D. S. Freed, *Special Kaehler manifolds,* Commun. Math. Phys. **203** (1999) 31 [arXiv:hep-th/9712042].

[50] M. A. Lledo, O. Macia, A. Van Proeyen and V. S. Varadarajan, *Special geometry for arbitrary signatures,* Contribution to the handbook on pseudo-Riemannian geometry and super- symmetry, ed. V. Cortés, Eur. Math. Society series "IRMA Lectures in Mathematics and Theoretical Physics". arXiv:hep-th/0612210, mu

[51] V. Cortes, *Special Kaehler manifolds: a survey,* in Proceedings of the 21st Winter School on "Geometry and Physics" (Srni 2001), Eds. J. Slovak and Martin Cadek, Rend. Circ. Mat. Palermo (2) Suppl. no. 66 (2001), 11, arXiv:math/0112114.

[52] P. Aschieri, *Duality rotations and BPS monopoles with space and time noncommutativity,* Nucl. Phys. B **617** (2001) 321 [arXiv:hep-th/0106281].

[53] P. Aschieri, *Monopoles in space-time noncommutative Born-Infeld theory,* Fortsch. Phys. **50** (2002) 680 [arXiv:hep-th/0112186].

[54] P. Aschieri, L. Castellani and A. P. Isaev, *Discretized Yang-Mills and Born-Infeld actions on finite group geometries,* Int. J. Mod. Phys. A **18** (2003) 3555 [arXiv:hep-th/0201223].

[55] I. Bialynicki-Birula and Z. Bialynicka-Birula, *Quantum Electrodynamics*, Pergamon, Oxford 1975, (also available at http://www.cft.edu.pl/ birula/publ.php).

[56] S. Deser and C. Teitelboim, *Duality Transformations Of Abelian And Nonabelian Gauge Fields,* Phys. Rev. D **13** (1976) 1592.

[57] I. Bialynicki-Birula, *Nonlinear Electrodynamics: Variations On A Theme By Born And Infeld, In *Jancewicz, B. (Ed.), Lukierski, J. (Ed.): Quantum Theory Of Particles and Fields*, 31-48* (also available at http://www.cft.edu.pl/ birula/publ.php).

[58] S. Deser and O. Sarioglu, *Hamiltonian electric/magnetic duality and Lorentz invariance,* Phys. Lett. B **423** (1998) 369 [arXiv:hep-th/9712067].
Put reference to Deser

[59] S. R. Coleman, J. Wess and B. Zumino, *Structure of phenomenological Lagrangians. 1,* Phys. Rev. **177** (1969) 2239.

[60] C. G. . Callan, S. R. Coleman, J. Wess and B. Zumino, *Structure of phenomenological Lagrangians. 2,* Phys. Rev. **177** (1969) 2247.

[61] P. C. Argyres, C. R. Nappi, *Spin 1 Effective Actions from Open Strings*, Nucl. Phys. B330 (1990) 151

[62] A. A. Tseytlin, *On Non-abelian Generalisation of Born-Infeld Action in String Theory*, Nucl. Phys. B501 (1997) 41; hep-th/9701125

[63] L. Andrianopoli, R. D'Auria and S. Ferrara, *U-invariants, black-hole entropy and fixed scalars,* Phys. Lett. B **403** (1997) 12 [arXiv:hep-th/9703156].

[64] L. Castellani, A. Ceresole, S. Ferrara, R. D'Auria, P. Fre and E. Maina, *The Complete N=3 Matter Coupled Supergravity,* Nucl. Phys. B **268** (1986) 317.

[65] E. Cremmer in *Supergravity '81*, ed. by S. Ferrara and J. G. Taylor, Pag. 313; B. Julia in *Superspace & Supergravity*, ed. by S. Hawking and M. Rocek, Cambridge (1981) pag. 331

[66] R. Kallosh and M. Soroush, *Explicit Action of E7(7) on N=8 Supergravity Fields,* Nucl. Phys. B **801** (2008) 25 [arXiv:0802.4106 [hep-th]].

[67] M. Trigiante, *Dual Gauged Supergravities,* arXiv:hep-th/0701218.

[68] B. Craps, F. Roose, W. Troost and A. Van Proeyen, *What is special Kaehler geometry?,* Nucl. Phys. B **503**, 565 (1997) [arXiv:hep-th/9703082].

[69] P. Aschieri, D. Svrtan, *Convergence of perturbative solutions of matrix equations,* in preparation.

[70] A. S. Schwarz, *Noncommutative Algebraic Equations and Noncommutative Eigenvalue Problem,* Lett. Math. Phys. **52** (2000) 177 [arXiv:hep-th/0004088].

[71] B. L. Cerchiai and B. Zumino, *Properties of perturbative solutions of unilateral matrix equations,* Lett. Math. Phys. **54**, 33 (2000) [arXiv:hep-th/0009013].
B. L. Cerchiai and B. Zumino, *Some remarks on unilateral matrix equations,* Mod. Phys. Lett. A **16**, 191 (2001) [arXiv:hep-th/0105065].

[72] B. Zumino, *Supersymmetry and Kähler Manifolds*, Phys. Lett. 87B (1979) 203

[73] P. Binetruy and M.K. Gaillard, *S-Duality Constraints on Effective Potentials for Gaugino Condensation*, Phys. Lett. B365 (1996) 87; hep-th/9506207

[74] L. Andrianopoli, R. D'Auria, S. Ferrara and M. Trigiante, *Black-hole attractors in $N = 1$ supergravity,* JHEP **0707** (2007) 019 [arXiv:hep-th/0703178].

[75] S. Ferrara and A. Strominger, *$N = 2$ space-time supersymmetry and Calabi Yau Moduli Space*, in Proceedings of College Station Workshop "Strings '89", pag. 245, eds. Arnowitt et al., World scientific 1989. P. Candelas and X. de la Ossa, *Moduli Space of Calabi Yau Manifolds,* Nucl. Phys. B **355**, 455 (1991). B. de Wit and A. Van Proeyen, *Potentials And Symmetries Of General Gauged N=2 Supergravity - Yang-Mills Models,* Nucl. Phys. B **245**, 89 (1984). E. Cremmer, C. Kounnas, A. Van Proeyen, J. P. Derendinger, S. Ferrara, B. de Wit and L. Girardello, *Vector Multiplets Coupled To N=2 Supergravity: Superhiggs Effect, Flat Potentials And Geometric Structure,* Nucl. Phys. B **250**, 385 (1985). B. de Wit,

P. G. Lauwers and A. Van Proeyen, *Lagrangians Of N=2 Supergravity - Matter Systems,* Nucl. Phys. B **255**, 569 (1985). S. Ferrara, C. Kounnas, D. Lust and F. Zwirner, *Duality invariant partition functions and automorphic superpotentials for (2,2) string compactifications,* Nucl. Phys. B **365**, 431 (1991). L. Castellani, R. D'Auria and S. Ferrara, *Special geometry without special coordinates,* Class. Quant. Grav. **7**, 1767 (1990). L. Castellani, R. D'Auria and S. Ferrara, *Special Kahler Geometry; an intrinsic Formulation from $N = 2$ Space-Time Supersymmetry,* Phys. Lett. B **241**, 57 (1990).

[76] A. Strominger, *Special Geometry,* Commun. Math. Phys. **133** (1990) 163.

[77] L. Andrianopoli, M. Bertolini, A. Ceresole, R. D'Auria, S. Ferrara, P. Fré and T. Magri, *$N = 2$ supergravity and $N = 2$ super Yang-Mills theory on general scalar manifolds: Symplectic covariance, gaugings and the momentum map,* J. Geom. Phys. **23** (1997) 111 [arXiv:hep-th/9605032]. L. Andrianopoli, M. Bertolini, A. Ceresole, R. D'Auria, S. Ferrara and P. Fré, *General Matter Coupled N=2 Supergravity,* Nucl. Phys. B **476** (1996) 397 [arXiv:hep-th/9603004].

[78] E. Cremmer and A. Van Proeyen, *Classification Of Kahler Manifolds In N=2 Vector Multiplet Supergravity Couplings,* Class. Quant. Grav. **2** (1985) 445.

[79] N. Seiberg and E. Witten, *String theory and noncommutative geometry,* JHEP**9909** (1999) 032 [hep-th/9908142].

[80] J. Madore, S. Schraml, P. Schupp and J. Wess, *Gauge theory on noncommutative spaces,* Eur. Phys. J. C **16** (2000) 161 [arXiv:hep-th/0001203].

[81] B. Jurco, P. Schupp and J. Wess, *Nonabelian noncommutative gauge theory via noncommutative extra dimensions,* Nucl. Phys. B **604** (2001) 148 [hep-th/0102129].

[82] N. Seiberg, L. Susskind and N. Toumbas, *Space/time non-commutativity and causality,* JHEP**0006** (2000) 044 [hep-th/0005015].
J. L. Barbon and E. Rabinovici, *Stringy fuzziness as the custodian of time-space noncommutativity,* Phys. Lett. B **486** (2000) 202 [hep-th/0005073].
J. Gomis and T. Mehen, *Space-time noncommutative field theories and unitarity,* Nucl. Phys. B **591** (2000) 265 [hep-th/0005129].

[83] O. Aharony, J. Gomis and T. Mehen, *On theories with light-like noncommutativity,* JHEP**0009** (2000) 023 [hep-th/0006236].

[84] D. Bahns, S. Doplicher, K. Fredenhagen and G. Piacitelli, *On the unitarity problem in space/time noncommutative theories,* Phys. Lett. B **533** (2002) 178 [arXiv:hep-th/0201222].

[85] P. Aschieri, *On duality rotations in light-like noncommutative electromagnetism,* Mod. Phys. Lett. A **16** (2001) 163 [hep-th/0103150].

[86] O. J. Ganor, G. Rajesh and S. Sethi, *Duality and non-commutative gauge theory,* Phys. Rev. D **62** (2000) 125008 [hep-th/0005046].

[87] R. Gopakumar, J. Maldacena, S. Minwalla and A. Strominger, *S-duality and noncommutative gauge theory,* JHEP**0006** (2000) 036 [hep-th/0005048].

[88] S. Rey and R. von Unge, *S-duality, noncritical open string and noncommutative gauge theory,* Phys. Lett. B **499** (2001) 215 [hep-th/0007089].

[89] J. X. Lu, S. Roy and H. Singh, *SL(2,Z) duality and 4-dimensional noncommutative theories,* Nucl. Phys. B **595** (2001) 298 [hep-th/0007168].

[90] M. Alishahiha, Y. Oz and J. G. Russo, *Supergravity and light-like non-commutativity,* JHEP **0009** (2000) 002 [hep-th/0007215].

[91] A. Schwarz, *Morita equivalence and duality,* Nucl. Phys. B **534** (1998) 720 [hep-th/9805034].
P. Ho, *Twisted bundle on quantum torus and BPS states in matrix theory,* Phys. Lett. B **434** (1998) 41 [hep-th/9803166].
D. Brace, B. Morariu and B. Zumino, *Dualities of the matrix model from T-duality of the type II string,* Nucl. Phys. B **545** (1999) 192 [hep-th/9810099].

[92] C. Chu and P. Ho, *Noncommutative open string and D-brane,* Nucl. Phys. B **550** (1999) 151 [hep-th/9812219].

From Duality Transformations in Supersymmetric Yang-Mills Theories coupled to Supergravity by Ceresole, Ferrara, Proeyen

[1] N. Seiberg and E. Witten, Nucl. Phys. B426 (1994) 19, hep-th/9407087; Nucl. Phys. B431 (1994) 484, hep-th/9408099.

[2] A. Klemm, W. Lerche, S. Yankielowicz and S. Theisen, Simple Singularities and $N=2$ Supersymmetric Yang-Mills Theory, preprint CERN-TH.7495/94, LMUTPW 94/16, hep-th/9411048 and On the monodromies of $N=2$ supersymmetric Yang-Mills theory, to be published in the proceedings of the Workshop on Physics from the Planck Scale to Electromagnetic Scale, Warsaw, Poland, Sep 21-24, 1994 , and for 28th International Symposium on Particle Theory, Wendisch-Rietz, Germany, 30 Aug - 3 Sep 1994, preprint CERN-TH-7538-94, hep-th/9412158; P. C. Argyres and A. E. Faraggi, The Vacuum Structure and Spectrum of $N=2$ Supersymmetric $SU(n)$ Gauge Theory, preprint IASSNSHEP-94/94, hep-th/9411057

[3] B. de Wit and A. Van Proeyen, Nucl. Phys. B245 (1984) 89; E. Cremmer, C. Kounnas, A. Van Proeyen, J. P. Derendinger, S. Ferrara, B. De Wit and L. Girardello, Nucl. Phys. B250 (1985) 385; B. de Wit, P. G. Lauwers and A. Van Proeyen, Nucl. Phys. $\underline{B255}$ (1985) 569; S. Cecotti, S. Ferrara and L. Girardello, Int. J. Mod. Phys. $\underline{A4}$(1989)2475.

[4] A. Ceresole, R. D'Auria and S. Ferrara, Phys. Lett. 339 B (1994) 71, hepth/9408036.

[5] D. Amati, K. Konishi, Y. Meurice, G. C. Rossi and G. Veneziano, Phys. Rep. $\underline{162}$(1988)169.

[6] D. Anselmi and P. Frè, Nucl. Phys. $\underline{B404}$ (1993) 288; and Gauged Hyperinstantons and Monopole equations , preprint HUTP-94/A041, SISSA 182/94/EP, hepth/9411205.

[7] E. Witten, Monopoles and Four-Manifolds, preprint IASSNS-HEP-94-96, hepth/9411102.

[8] S. Ferrara, L. Girardello, C. Kounnas and M. Porrati, Phys. Lett. $\underline{B192}$ (1987) 368.

[9] N. Seiberg, Nucl. Phys. $\underline{B303}$ (1988) 206.

[10] S. Ferrara and M. Porrati, Phys. Lett. B216 (1989) 289.

[11] For a review on target space duality see A. Giveon, M. Porrati and E. Rabinovici, Phys. Rep. $\underline{244}$ (1994) 77, hep-th/9401139.

[12] A. Sen, Nucl. Phys. B388 (1992) 457 and Phys. Lett. 303B (1993) 22; A. Sen and J. H. Schwarz, Nucl. Phys. $\underline{B411}$ (1994) 35; Phys. Lett. $\underline{312B}$ (1993) 105.

[13] M. J. Duff and R. R. Khuri, Nucl. Phys. $\underline{B411}$ (1994) 473.

[14] J. Gauntlett and J. A. Harvey, S-Duality and the Spectrum of Magnetic Monopoles in Heterotic String Theory, preprint EFI-94-36, hep-th/9407111.

[15] L. Girardello, A. Giveon, M. Porrati and A. Zaffaroni, Phys. Lett. B334 (1994) 331, hep-th/9406128.

[16] J. H. Schwarz, Evidence for Non-perturbative String Symmetries, proceedings of the Conference on Topology, Strings and Integrable Models (Satellite Colloquium to the ICMP-11 18-23 Jul 1994), Paris, France, 25-28 Jul 1994, preprint CALT68-1965, hep-th/9411178.

[17] S. Ferrara, J. Scherk and B. Zumino, Nucl. Phys. *B*121 (1977) 393.

[18] M. K. Gaillard and B. Zumino, Nucl. Phys. *B*193 (1981) 221.

[19] P. Di Vecchia, R. Musto, F. Nicodemi and R. Pettorino, Nucl. Phys. *B*252 (1985) 635.

[20] N. Seiberg, Phys. Lett. 206*B* (1988) 75.

[21] S. Ferrara and A. Van Proeyen, Class. Quantum Grav. 6 (1989) 124.

[22] E. Cremmer and A. Van Proeyen, Class. Quantum Grav. 2 (1985) 445.

[23] P. Candelas and X. de la Ossa, Nucl. Phys. B355 (1991) 455; P. Candelas, X. de la Ossa, P. S. Green and L. Parkes, Phys. Lett. 258*B* (1991) 118; Nucl. Phys. B359 (1991) 21.

[24] A. C. Cadavid and S. Ferrara, Phys. Lett. 267*B* (1991) 193; W. Lerche, D. Smit and N. Warner, Nucl. Phys. B372 (1992) 87.

[25] S. Ferrara and A. Strominger, in Strings ' 89 , eds. R. Arnowitt, R. Bryan, M.J. Duff, D.V. Nanopoulos and C.N. Pope(World Scientific, 1989), p. 245; A. Strominger, Comm. Math. Phys. 133 (1990) 163.

[26] L. Castellani, R. D'Auria and S. Ferrara, Phys. Lett. 241*B* (1990) 57; Class. Quantum Grav. 1 (1990) 317.

[27] R. D'Auria, S. Ferrara and P. Frè, Nucl. Phys. B359 (1991) 705.

[28] S. Ferrara and J. Louis, Phys. Lett. 278*B* (1992) 240.

[29] A. Ceresole, R. D'Auria, S. Ferrara, W. Lerche and J. Louis, Int. J. Mod. Phys. *A*8(1993)79

[30] E. Cremmer, S. Ferrara, L. Girardello and A. Van Proeyen, Nucl. Phys. *B*212 (1983) 413.

[31] B. de Wit, J.W. van Holten and A. Van Proeyen, Nucl. Phys. *B*184 (1981) 77

[32] B. de Wit, F. Vanderseypen and A. Van Proeyen, Nucl. Phys. *B*400 (1993) 463.

[33] A. Das, Phys. Rev. D15 (1977) 2805; E. Cremmer, J. Scherk and S. Ferrara, Phys. Lett. 68*B* (1977) 234; E. Cremmer and J. Scherk, Nucl. Phys. *B*127 (1977) 259.

[34] E. Cremmer, J. Scherk and S. Ferrara, Phys. Lett. 74*B* (1978) 61.

[35] R. Kallosh, A. Linde, T. Ortín, A. Peet and A. Van Proeyen, Phys. Rev. D46 (1992) 5278.

[36] E. S. Fradkin, M. A. Vasiliev, Nuovo Cim. Lett. 25 (1979) 79; B. de Wit and J. W. van Holten, Nucl. Phys. *B*155 (1979) 530; B. de Wit, J. W. van Holten and A. Van Proeyen, Nucl. Phys. *B*167 (1980) 186.

[37] E. Cremmer, B. Julia, J. Scherk, S. Ferrara, L. Girardello and P. van Nieuwenhuizen, Phys. Lett. 79*B* (1978) 231, Nucl. Phys. *B*147 (1978) 105.

[38] M. Villasante, Phys. Rev. *D*45 (1992) 831.

[39] K. Narain, Phys. Lett. 169*B* (1986) 41; K. Narain, M. Sarmadi and E. Witten, Nucl. Phys. B279 (1987) 369.

[40] A. Shapere and F. Wilczek, Nucl. Phys. B320 (1989) 669.

[41] I. Antoniadis, C. Bachas and C. Kounnas, Nucl. Phys. B289 (1987) 87; H. Kawai, D. C. Lewellen and S. H. Tye, Nucl. Phys. *B*288 (1987)1.

[42] S. Ferrara, C. Kounnas, D. Lust and F. Zwirner, Nucl. Phys. B365 (1991) 431.

[43] L. J. Dixon, V. S. Kaplunovsky and J. Louis, Nucl. Phys. B355 (1990) 27.

[44] G. L. Cardoso and B. A. Ovrut, Nucl. Phys. B369 (1992) 351.

[45] J. P. Derendinger, S. Ferrara, C. Kounnas and F. Zwirner, Nucl. Phys. B372 (1992) 145.

[46] I. Antoniadis, E. Gava, K. S. Narain and T. R. Taylor, Nucl. Phys. *B*413 (1994) 162.

[47] P. Frè and P. Soriani, Nucl. Phys. *B*371 (1992) 659.

[48] see for instance A. Salam and E. Sezgin, Supergravities in diverse dimensions, World Scientific (1989); L. Castellani, R. D'Auria and P. Frè, Supergravity and Superstrings, a geometric perspective, World Scientific (1991).

[49] C. M. Hull and P. K. Townsend, Unity of Superstring Dualities, preprint QMW 94-30, hep-th/9410167.

[50] S. Ferrara, C. Kounnas, Nucl. Phys. B328 (1989) 406; S. Ferrara, P. Fre Int. J. Mod. Phys. *A5*(1990)989.

[51] R. Dijkgraaf, E. Verlinde, H. Verlinde, Comm. Math. Phys. 115 (1988) 649.

[52] M. K. Prasad and C. M. Sommerfield, Phys. Rev. Lett. 35 (1975) 760; E. B. Bogomol'nyi, Sov. J. Nucl. Phys. 24 (1976) 449; E. Witten and D. Olive, Phys. Lett. *78B* (1978) 97; C. Montonen and D. Olive, Phys. Lett. *72B* (1977) 117; P. Goddard, J. Nuyts and D. Olive, Nucl. Phys. *B125* (1977) 1.

[53] G. Gibbons and C. Hull, Phys. Lett. *B109* (1982) 190; G. Gibbons and N. Manton, Nucl. Phys. *B274*(1986)183.

[54] A. Sen, Int. J. Mod. Phys. *A9* (1994) 3707, hep-th/9402002 and references there in.

[55] J. Harvey and J. Liu, Phys. Lett. B268 (1991) 40.

[56] R.R. Khuri, Phys. Lett. *B259* (1991) 261; Phys. Lett. *B294* (1992) 325.

[57] R. Kallosh and T. Ortin, Phys. Rev. D48 (1993) 742, hep-th/9302109.

from Guill Zum D in Non lin ED

E Bibliography

[1] S. Ferrara, J. Scherk and B. Zumino, *Nucl. Phys.* **B121:** 393 (1977); E. Cremmer and B. Julia, *Nucl. Phys.* **B159:** 141 (1979).

[2] M.K. Gaillard and B. Zumino, *Nucl. Phys.* **B193:** 221 (1981).

[3] B. Zumino, *Quantum Structure of Space and Time*, Eds. M.J. Duff and C.J. Isham (Cambridge University Press) p. 363 (1982).

[4] G.W. Gibbons and D.A. Rasheed, *Nucl. Phys.* **B454:** 185 (1995).

[5] G.W. Gibbons and D.A. Rasheed, *Phys. Lett.* **B365:** 46 (1996).

[6] M. Born and L. Infeld, *Proc. Roy. Soc.* (London) **A144:** 425 (1934).

[7] E. Schrödinger, *Proc. Roy. Soc.* (London) **A150:** 465 (1935).

[8] P. Binétruy and M.K. Gaillard, *Phys. Rev.* **D32:** 931 (1985).

[9] P. Binétruy and M.K. Gaillard, *Phys. Lett.* **B365:** 87 (1996).

[10] J.H. Schwarz and A. Sen, *Phys. Lett.* **B312:** 105 (1993) and *Nucl. Phys.* **B411:** 35 (1994); M. Duff, *Nucl. Phys.* **B442:** 47 (1995); E. Witten, *Nucl. Phys.* **B443:** 85 (1995).

[11] S. Cecotti, S. Ferrara and L. Girardello, *Nucl. Phys.* **B308:** 436 (1988).

F Bibliography

[1] S. Ferrara, J. Scherk and B. Zumino, *Nucl. Phys.* **B121:** 393 (1977); E. Cremmer and B. Julia, *Nucl. Phys.* **B159:** 141 (1979).

[2] M.K. Gaillard and B. Zumino, *Nucl. Phys.* **B193:** 221 (1981).

[3] B. Zumino, *Quantum Structure of Space and Time*, Eds. M.J. Duff and C.J. Isham (Cambridge University Press) p. 363 (1982).

[4] G.W. Gibbons and D.A. Rasheed, *Nucl. Phys.* **B454:** 185 (1995).

[5] G.W. Gibbons and D.A. Rasheed, *Phys. Lett.* **B365:** 46 (1996).

[6] M.K. Gaillard and B. Zumino, Berkeley preprint LBNL-40370, UCB-PTH-97/29, hep-th/9705226 (1997), to be published in the memorial volume for D.V. Volkov.

[7] M. Born and L. Infeld, *Proc. Roy. Soc.* (London) **A144:** 425 (1934).

[8] E. Schrödinger, *Proc. Roy. Soc.* (London) **A150:** 465 (1935).

[9] P. Binétruy and M.K. Gaillard, *Phys. Rev.* **D32:** 931 (1985).

[10] R. Courant and D. Hilbert, *Methods of Mathematical Physics* **Vol. II**, Interscience (1962), p. 93 and Chapters I and II *passim.*

[11] M. Perry and J.H. Schwarz, *Nucl. Phys.* **B489:** 47 (1997).

[12] E. Witten, *Phys. Lett.* **B86:** 283 (1979).

[13] N. Seiberg and E. Witten, *Nucl. Phys.* **B426:** 19 (1994).

[14] See, *e.g.* D.I. Olive, *Nucl. Phys.* **B** (Proc. Suppl.) **46:** 1 (1996).

[15] S. Deser and R. Puzalowski, *J. Phys.* **A13:** 2501 (1980).

[16] S. Cecotti and S. Ferrara, *Phys. Lett.* **B187:** 335 (1987).

[17] P. Binétruy and M.K. Gaillard, *Phys. Lett.* **B365:** 87 (1996).

[18] J.H. Schwarz and A. Sen, *Phys. Lett.* **B312:** 105 (1993) and *Nucl. Phys.* **B411:** 35 (1994); M. Duff, *Nucl. Phys.* **B442:** 47 (1995); E. Witten, *Nucl. Phys.* **B443:** 85 (1995).

[19] S. Cecotti, S. Ferrara and L. Girardello, *Nucl. Phys.* **B308:** 436 (1988).

From duality in Quantum Field Theory and String Theory by Álvarez-Gaumé Zamora

G Bibliography

[1] S. Deser and C. Teitelboim, Phys. Rev. **D13** (1976), 1592;
S. Deser, A. Gomberoff, M. Henneaux and C. Teitelboim, Phys. Lett. **B400** (1997), 80.

[2] P.A.M. Dirac, Proc. Roy. Soc. **A133** (1931), 60.

[3] J. Schwinger, Phys. Rev. **144** (1966) 1087; **173** (1968) 1536;
D. Zwanziger, Phys. Rev. **176** (1968) 1480.

[4] C. Vafa and E. Witten, Nucl. Phys. **B431** (1994), 3;
F. Ferrari, hep-th/9702166.

[5] H.B. Nielsen and P. Olesen, Nucl. Phys. **B61**(1973), 45-61.

[6] G. 't Hooft, Nucl. Phys. **B79** (1974) 276.

[7] A. M. Polyakov, JETP Lett. **20** (1974) 194.

[8] B. Julia and A. Zee, Phys. Rev. **D11** (1975) 2227.

[9] E. B. Bogomol'nyi, Sov. J. Nucl. Phys. **24** (1976) 449.

[10] M. K. Prasad and C. M. Sommerfield, Phys. Rev. Lett. **35** (1975) 760.

[11] E. Witten, Phys. Lett. **86B** (1979) 283.

[12] G. 't Hooft, Nucl. Phys. **B190**[FS3](181), 455, and *1981 Cargese Summer School Lecture Notes on Fundamental Interactions*, in 'Under the Spell of the Gauge Principle', World Scientific, 1994.

[13] N. Seiberg and E. Witten, Nucl. Phys. **B426** (1994), 19.

[14] L. Álvarez-Gaumé, J. Distler, C. Kounnas and M. Mariño, Int. J. Mod. Phys. **A11** (1996) 4745;
L. Álvarez-Gaumé and M. Mariño, Int. J. Mod. Phys. **A12** (1997) 975.

[15] L. Álvarez-Gaumé, M. Mariño and F. Zamora, hep-th/9703072, hep-th/9707017, to appear in Int. J. Mod. Phys. **A**.

[16] N. Evans, S.D.H. Hsu and M. Schwetz, Nucl. Phys. **B484**, 124 (1997).

[17] K.G. Wilson, Phys. Rev. **D10** (1974), 2445.

[18] E. Fradkin and S.H. Shenker, Phys. Rev. **D12** (1979), 3682;
T. Banks and E. Rabinovici, Nucl. Phys. **B160** (1979), 349.

[19] J. Wess and J. Bagger, 'Supersymmetry and Supergravity', Princeton University Press, 2nd ed., 1992; and references therein.

[20] A. Salam and J. Strathdee, Nucl. Phys. **B76** (1974), 477.

[21] S. Weinberg, 'The Quantum Theory of Fields I', Cambridge University Press, 1996.

[22] B. Zumino, Phys. Lett. **87B**, 203 (1979).

[23] M.T. Grisaru, W. Siegel and M. Rocek, Nucl. Phys. **B159** (1979), 429.

[24] N. Seiberg, 'The Power of Holomorphy: Exact results in 4-D SUSY Gauge Theories.', in PASCOS 94, pg.357, hep-th/9408013.

[25] K. Konishi and K. Shizuya, Nuovo Cim. **90A** (1985), 111.

[26] N. Arkani-Hamed and H. Murayama, hep-th/9705189, hep-th/9707133.

[27] B. de Wit, P.G. Lauwers and A. van Proeyen, Nucl. Phys. **B255** (1985), 569.

[28] K. Intriligator and N. Seiberg, Nucl. Phys. Proc. Suppl. **45BC** (1996) 1, hep-th/9509066;
M.E. Peskin, TASI 96 lectures, hep-th/9702094;
M. Shifman, hep-th/9704114.

[29] E. Witten, *Nucl. Phys.* **B202** (1982), 253.

[30] A.C. Davis, M. Dine and N. Seiberg, Phys. Lett. **125B** (1983), 487;
I. Affleck, M. Dine and N. Seiberg, Phys. Rev. Lett **51**, 1026 (1983).

[31] I. Affleck, M. Dine and N. Seiberg, Nucl. Phys. **B241** (1984), 493.

[32] D. Finnell and P. Pouliot, *Nucl. Phys.* **B453** (1995), 225.

[33] K. Intriligator, R.G. Leigh and N. Seiberg, *Phys. Rev.* **D50** (1994), 1092.

[34] K. Intriligator, *Phys. Lett.* **336B** (1994), 409.

[35] G. Veneziano and S. Yankielowicz, *Phys. Lett.* **113B** (1982), 231; T. Taylor, G. Veneziano and S. Yankielowicz, *Nucl. Phys.* **B218** (1983), 439.

[36] N. Seiberg, Phys. Rev. **D49** (1994), 6857.

[37] G. 't Hooft, 'Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking', Cargese 1979, in 'Under the Spell of the Gauge Principle', World Scientific, 1994.

[38] N. Seiberg, Nucl. Phys. **B435** (1995), 129.

[39] T. Banks and A. Zaks, *Nucl. Phys.* **B196** (1982), 189.

[40] R. Haag, J. T. Lopuszanski and M. Sohnius, Nucl. Phys. **B88** (1975) 257.

[41] S.J. Gates, Jr., Nucl. Phys. **B238** (1984), 349.

[42] L. Álvarez-Gaumé and S.F. Hassan, Fortsch. Phys. **45**, 159 (1997);
A. Bilal, hep-th/9601077;
C. Gómez and R. Hernández, Advanced School on Effective Theories, Almuñecar 1995, hep-th/9510023;
W. Lerche, Nucl. Phys. Proc. Suppl. **55B**, 83 (1997).

[43] E. Witten and D. Olive, Phys. Lett. **78B** (1978) 97.

[44] N. Seiberg and E. Witten, Nucl. Phys. **B431** (1994), 484.

[45] S. Mandelstam, Phys. Rept. **C23** (1976), 245.

[46] L. Girardello and M.T. Grisaru, Nucl. Phys. **B194** (1982), 65.

[47] N. Evans, S.D.H. Hsu and M. Schwetz, Phys. Lett **355B** 475 (1995), Nucl. Phys. **B 456** 205 (1995);
O. Aharony, J. Sonnenschein, M.E. Peskin and S. Yankielowicz, Phys. Rev. **D52** 6157 (1995);
E. D'Hoker, Y. Mimura and N. Sakai, Phys. Rev. **D54** 7724 (1996).

[48] M.B. Green, J.H. Schwarz and E.Witten, 'Superstring Theory', Cambridge University Press (1987);
and references therein.

[49] A. Giveon, M. Porrati and E. Rabinovici, Phys. Rept. **244**, 77 (1994); and references therein.

[50] L. Álvarez-Gaumé and E. Witten, Nucl. Phys. **B234**, 269 (1984).

[51] S.H. Shenker, 'The Strength of Nonperturbative Effects in String Theory', in *Random Surfaces and Quantum Gravity*, eds. O.Alvarez et al. (1991).

[52] M.J. Duff, R.R. Khuri and J.X. Lu, Phys. Rept. **259**, 213 (1995);
M.J. Duff, 'Supermembranes', TASI 96 lectures, hep-th/9611203;
and references therein.

[53] C.M. Hull and P.K. Townsend, Nucl. Phys. **B438**(1995), 109;
E. Witten, Nucl. Phys. **B443**(1995), 83;
A. Strominger, Nucl. Phys. **B451**(1995), 96.

[54] A. Font, L. Ibañez, D. Lust and F. Quevedo, Phys. Lett. **249**, 35 (1990).

[55] J. Polchinski, TASI 96 lectures, hep-th/9611050;
C. Bachas, 'Half a Lecture on D-branes', hep-th/9701019;
and references therein.

[56] J. Dai, R.G. Leigh and J. Polchinski, Mod. Phys. Lett. **A4**, 2073 (1989);
R.G. Leigh, Mod. Phys. Lett. **A4**, 2767 (1989).

[57] J. Polchinski, Phys. Rev. Lett. **75** (1995), 4724.

[58] E. Witten, Nucl. Phys. **B460** (1996), 335.

[59] J. Polchinski and E. Witten, Nucl. Phys. **B460** (1996), 525.

[60] P.K. Townsend, Phys. Lett. **B350** (1995), 184;
E. Witten, Nucl. Phys. **B443** (1995), 85.

[61] E. Cremmer, B. Julia and J. Scherk, Phys. Lett. **76B**, 409.

[62] P. Aspinwall, TASI 96 lectures, hep-th/9611137;
J. Schwarz, TASI 96 lectures, hep-th/9607202;
P.K. Townsend, Trieste 95 lectures, hep-th/9612121;
E. Witten, Nucl. Phys. **B460** (1996) 335;
P. Horava and E. Witten, Nucl. Phys. **B460** (1996) 506;
T. Banks, W. Fischler, S.H. Shenker and L. Susskind, Phys. Rev. **D55** 5112 (1997);
A. Hanany and E. Witten, Nucl. Phys. **B492** (1997) 152;
E. Witten, hep-th/9703166;
R. Dijkgraaf, E. Verlinde and H. Verlinde, hep-th/9709107.

Bibliography by JM Figueroa-O'Farrill

[Ada69] J. F. Adams, Lectures on Lie groups, W. A. Benjamin, 1969.

[AH85] M. F. Atiyah and N. J. Hitchin, Low energy scattering of nonabelian monopoles, Phys. Lett. 107A (1985), 21-25.

[AH88], The geometry and dynamics of magnetic monopoles, Princeton University Press, 1988.

[Bes86] A. L. Besse, Einstein manifolds, Springer-Verlag, 1986.

[Blu94] J. D. Blum, Supersymmetric quantum mechanics of monopoles in $N = 4$ Yang-Mills theory, Phys. Lett. B333 (1994), 92-97.

[Bog76] E. B. Bogomol'nyi, The stability of classical solutions, Soviet J. Nuc. Phys. 24 (1976), 449-454.

[BT81] R. Bott and L. W. Tu, Differential forms in algebraic topology, Springer-Verlag, 1981.

[BtD85] T. Bröcker and T. tom Dieck, Representations of compact Lie groups, Springer-Verlag, 1985.

[Cal78] C. Callias, Axial anomalies and index theorems on open spaces, Comm. Math. Phys. 62 (1978), 213.

[Col77] S. Coleman, Classical lumps and their quantum descendants, New Phenomena in Subnuclear Physics (A. Zichichi, ed.), Plenum, 1977, Erice lectures on 't Hooft Polyakov monopoles, larger gauge groups, homotopy considerations; comparison with sine-Gordon.

[CPNS76] S. Coleman, S. Parke, A. Neveu, and C. M. Sommerfield, Can one dent a dyon?, Phys. Rev. D15 (1976), 544-545.

[DHdV78] A. D'Adda, R. Horsley, and P. di Vecchia, Supersymmetric magnetic monopoles and dyons, Phys. Lett. 76B (1978), 298-302.

[Dir31] P. A. M. Dirac, Quantised singularities in the electromagnetic field, Proc. R. Soc. A133 (1931), 60-72.

[Gau94] J. P. Gauntlett, Low energy dynamics of $N = 2$ supersymmetric monopoles, Nuc. Phys. 411B (1994), 433-460.

[GH78] P. Griffiths and J. Harris, Principles of algebraic geometry, Wiley, 1978.

[GM86] G. W. Gibbons and N. S. Manton, Classical and quantum dynamics of BPS monopoles, Nuc. Phys. 274B (1986), 183-224.

[GNO77] P. Goddard, J. Nuyts, and D. Olive, Gauge theories and magnetic charge, Nuc. Phys. B125 (1977), 1-28.

[Gol62] S. I. Goldberg, Curvature and homology, Academic Press, 1962.

[HKLR87] N. J. Hitchin, A. Karlhede, U. Lindström, and M. Rocek, Hyperkähler metrics and supersymmetry, Comm. Math. Phys. 108 (1987), 535-589.

[Hum72] J. E. Humphreys, Introduction to lie algebras and representation theory, vol. 9, Springer-Verlag, 1972.

[JT80] A. Jaffe and C. Taubes, Vortices and monopoles, Birkhäuser, 1980.

[JZ75] B. Julia and A. Zee, Poles with both magnetic and electric charges in non-abelian gauge theories, Phys. Rev. D11 (1975), 2227-2232.

[KN63] S. Kobayashi and K. Nomizu, Foundations of differential geometry, vol. 1, Interscience, 1963.

[KN69], Foundations of differential geometry, vol. 2, Interscience, 1969.

[Lic76] A. Lichnerowicz, Global theory of connections and holonomy groups, Noordhoff International Pub., 1976.

[Man82] N. S. Manton, A remark on the scattering of BPS monopoles, Phys. Lett. 110B (1982), 54-56.

[MO77] C. Montonen and D. Olive, Magnetic monopoles as gauge particles, Phys. Lett. 72B (1977), 117-120.

[Oli79] D. Olive, The electric and magnetic charges as extra components of four-momentum, Nuc. Phys. B153 (1979), 1-12.

[Osb79] H. Osborn, Topological charges for $N = 4$ supersymmetric gauge theories and monopoles of spin 1, Phys. Lett. 83B (1979), 321326.

[Pol74] A. M. Polyakov, Particle spectrum in quantum field theory, JETP Letters 20 (1974), 194-95.

[PS75] M. K. Prasad and C. M. Sommerfield, Exact classical solution for the 't Hooft monopole and the Julia-Zee dyon, Phys. Rev. Lett. 35 (1975), $760-762$.

[PS86] A. Pressley and G. Segal, Loop groups, Clarendon, 1986.

[Sal89] S. Salamon, Riemannian geometry and holonomy groups, Longman Scientific and Technical, 1989.

[Ser73] J. P. Serre, A course in arithmetic, Springer-Verlag, 1973.

[Soh85] M. F. Sohnius, Introducing supersymmetry, Phys. Rep. 128 (1985), 39-204.

[Ste51] N. E. Steenrod, The topology of fibre bundles, Princeton University Press, 1951.

[Tau83] C. H. Taubes, Stability in Yang-Mills theories, Comm. Math. Phys. 91 (1983), 473-540.

[tH74] G. 't Hooft, Magnetic monopoles in unified gauge theories, Nuc. Phys. B79 (1974), 276-284.

[War83] F. W. Warner, Foundations of differentiable manifolds and lie groups, vol. 94, Springer-Verlag, 1983.

[Wei79] E. J. Weinberg, Parameter counting for multimonopole solutions, Phys. Rev. D20 (1979), 936-944.

[Wel80] R. O. Wells, Jr., Differential analysis on complex manifolds, Springer-Verlag, 1980.

[Wit79] E. Witten, Dyons of charge e $0/2\pi$, Phys. Lett. 86B (1979), 283287.

[WO78] E. Witten and D. Olive, Supersymmetry algebras that include topological charges, Phys. Lett. 78B (1978), 97-101.

[Zum77] B. Zumino, Euclidean supersymmetry and the many-instanton problem, Phys. Lett. 69B (1977), 369-371.